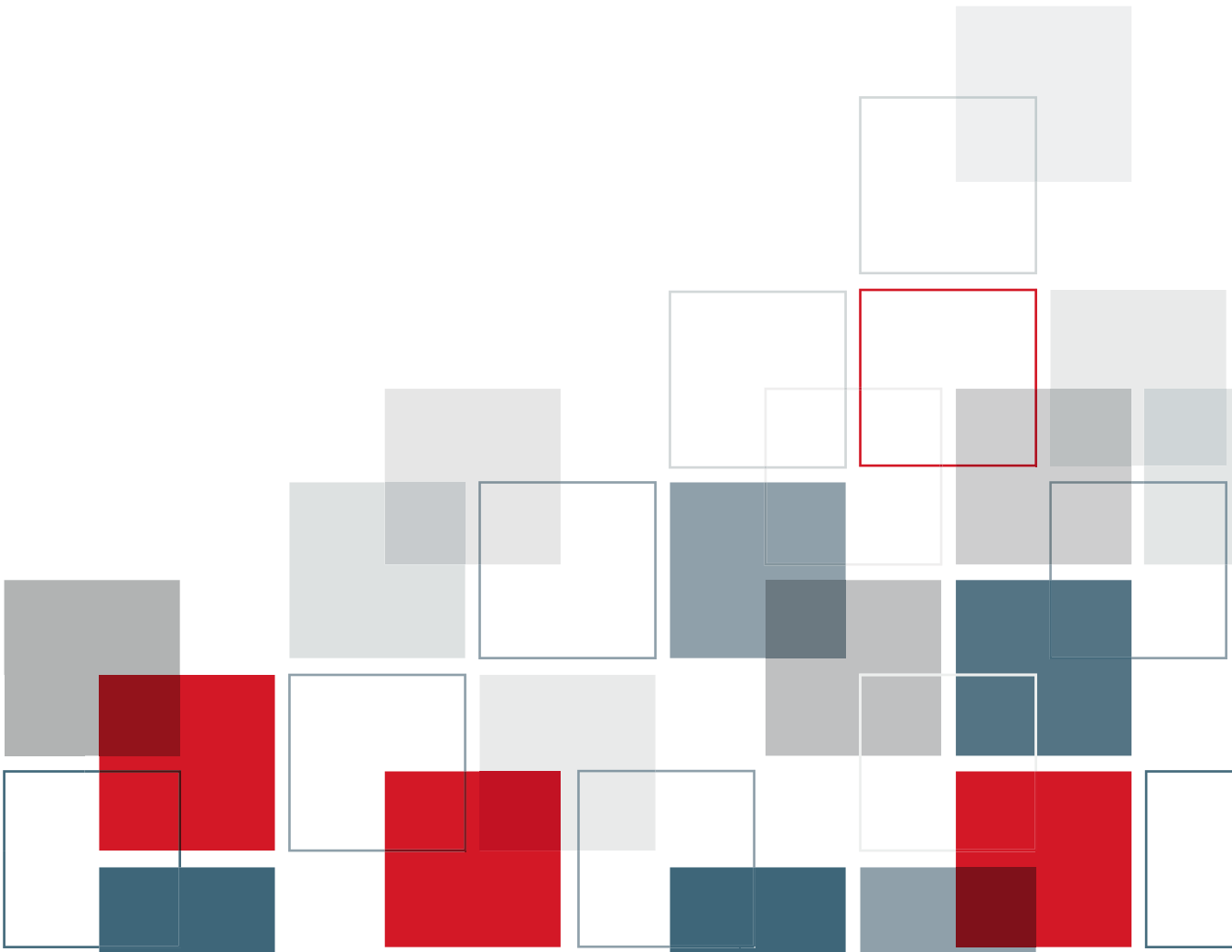


> SPSS® 13.0 Base Manual del usuario



Si desea obtener más información sobre los productos de software de SPSS®, visite nuestra página Web en <http://www.spss.com> o póngase en contacto con

SPSS Inc.
233 South Wacker Drive, 11th Floor
Chicago, IL 60606-6412 EE.UU.
Tel: (312) 651-3000
Fax: (312) 651-3668

SPSS es una marca registrada; los demás nombres de productos son marcas comerciales de SPSS Inc. para los programas de software de su propiedad. El material descrito en este software no puede ser reproducido ni distribuido sin la autorización expresa por escrito por parte de los propietarios de la marca registrada y de los derechos de la licencia en el software y en los copyrights de los materiales publicados.

El SOFTWARE y la documentación se proporcionan con DERECHOS LIMITADOS. Su uso, duplicación o revelación por parte del Gobierno están sujetos a las restricciones establecidas en la subdivisión (c)(1)(ii) de la cláusula Rights in Technical Data and Computer Software en 52.227-7013. El fabricante es SPSS Inc., 233 South Wacker Drive, 11th Floor, Chicago, IL 60606-6412, EE.UU.

Aviso general: El resto de los nombres de productos mencionados en este documento se utilizan sólo con fines identificativos y pueden ser marcas comerciales de sus respectivas empresas.

TableLook es una marca comercial de SPSS Inc.

Windows es una marca comercial registrada de Microsoft Corporation.

DataDirect, DataDirect Connect, INTERSOLV y SequeLink son marcas comerciales registradas de DataDirect Technologies. Algunas partes de este producto se han creado utilizando LEADTOOLS © 1991–2000, LEAD Technologies, Inc. TODOS LOS DERECHOS RESERVADOS.

LEAD, LEADTOOLS y LEADVIEW son marcas comerciales registradas de LEAD Technologies, Inc.

Sax Basic es una marca comercial de Sax Software Corporation. Copyright © 1993–2004 de Polar Engineering and Consulting. Todos los derechos reservados.

Algunas partes de este producto están basadas en el trabajo de FreeType Team (<http://www.freetype.org>).

Una parte del software de SPSS contiene tecnología zlib. Copyright © 1995–2002 de Jean-loup Gailly y Mark Adler. El software zlib se proporciona “tal cual”, sin ningún tipo de garantía implícita o explícita.

Una parte del software de SPSS contiene bibliotecas de Sun Java Runtime. Copyright © 2003 de Sun Microsystems, Inc. Reservados todos los derechos. Las bibliotecas de Sun Java Runtime incluyen código con licencia de RSA Security, Inc. Algunas partes de las bibliotecas tienen licencia de IBM y están disponibles en <http://oss.software.ibm.com/icu4j/>.

SPSS® Base 13.0 Manual del usuario

Copyright © 2004 de SPSS Inc.

Todos los derechos reservados.

Impreso en Irlanda.

Queda prohibida la reproducción, el almacenamiento en sistemas de recuperación o la transmisión de cualquier parte de esta publicación en cualquier forma y por cualquier medio (electrónico o mecánico, fotocopia, grabación o cualquier otro) sin previa autorización expresa y por escrito de parte del editor.

1 2 3 4 5 6 7 8 9 0 06 05 04

ISBN 1-56827-974-4

Prefacio

SPSS 13.0

SPSS 13.0 es un sistema global para el análisis de datos. SPSS puede adquirir datos de casi cualquier tipo de archivo y utilizarlos para generar informes tabulares, gráficos y diagramas de distribuciones y tendencias, estadísticos descriptivos y análisis estadísticos complejos.

Este manual, *SPSS® Base 13.0 Manual del usuario*, proporciona la documentación sobre la interfaz gráfica de usuario de SPSS para Windows. Los ejemplos que utilizan procedimientos estadísticos que se encuentran en SPSS Base 13.0 se suministran en el sistema de ayuda que se instala con el software. Los algoritmos utilizados en los procedimientos estadísticos están disponibles en el CD-ROM del producto.

Asimismo, además de los menús y los cuadros de diálogos, SPSS utiliza un lenguaje de comandos. Algunas de las funciones avanzadas del sistema sólo son accesibles a través de la sintaxis de comandos. (Dichas funciones no están disponibles en la versión para estudiantes.) La sintaxis de comandos completa aparece documentada en la referencia de sintaxis de comandos (*SPSS 13.0 Command Syntax Reference*), disponible en el menú Ayuda.

SPSS Opciones

Las siguientes opciones están disponibles como mejoras adicionales de la versión completa (no la versión para estudiantes) del sistema SPSS Base:

SPSS Modelos de regresión™ proporciona técnicas para analizar los datos que no se ajusten a los tradicionales modelos estadísticos. Incluye procedimientos para el análisis probit, la regresión logística, la estimación ponderada, la regresión de mínimos cuadrados en dos fases y la regresión no lineal general.

SPSS Modelos avanzados™ se centra en las técnicas utilizadas con más frecuencia en la investigación experimental y biomédica sofisticada. Incluye procedimientos para los modelos lineales generales (MLG), los modelos lineales mixtos, el análisis de componentes de la varianza, el análisis loglineal, la regresión ordinal, las tablas de mortalidad actuariales, el análisis de supervivencia de Kaplan-Meier y las regresiones de Cox básica y extendida.

SPSS Tablas™ crea una amplia variedad de informes tabulares de gran calidad, como por ejemplo tablas sofisticadas y presenta datos de respuestas múltiples.

SPSS Tendencias™ realiza análisis de predicción y de series temporales muy completos, con diversos modelos de ajuste de curvas, modelos de suavizado y métodos para la estimación de funciones autorregresivas.

SPSS Categorías® ejecuta procedimientos de escalamiento óptimo, incluidos los análisis de correspondencias.

SPSS Análisis conjunto™ ejecuta el análisis conjunto.

SPSS Pruebas exactas™ calcula los valores p exactos (valores de significación) para las pruebas estadísticas en aquellos casos en los que las muestras son pequeñas o se distribuyen de forma poco uniforme y puedan hacer que las pruebas habituales resulten poco precisas.

SPSS Análisis de valores perdidos™ describe los patrones de los datos perdidos, realiza una estimación de las medias y otros estadísticos y permite imputar los valores a las observaciones perdidas.

SPSS Mapas™ transforma los datos distribuidos geográficamente en mapas de alta calidad, con símbolos, colores, gráficos de barras, gráficos de sectores y combinaciones de temas, para representar no sólo lo que está ocurriendo sino también dónde tiene lugar.

SPSS Muestras complejas™ permite a los analistas de encuestas, mercado, salud y opinión pública, así como a los sociólogos que utilizan una metodología de encuesta de ejemplo, incorporar los diseños de muestras complejas al análisis de datos.

SPSS Árboles de clasificación™ crea un modelo de clasificación basado en árboles. Clasifica los casos en grupos o pronostica los valores de la variable dependiente (criterio) de acuerdo con los valores de las variables independientes (predictores). El procedimiento proporciona herramientas de validación para análisis de clasificación exploratorios y confirmatorios.

La familia de productos de SPSS también incluye aplicaciones para la introducción de datos, análisis de textos, clasificación, redes neuronales y diagramas de flujo.

Instalación

Para instalar Base system, ejecute el Asistente para autorización de licencia utilizando el código de autorización que le envió SPSS Inc. Para obtener más información, consulte las instrucciones de instalación proporcionadas con el sistema Base de SPSS.

Compatibilidad

SPSS está diseñado para ejecutarse en gran cantidad de sistemas de ordenadores. Consulte las instrucciones de instalación entregadas con su sistema para obtener información específica acerca de los requisitos mínimos y los recomendados.

Números de serie

El número de serie es su número de identificación con SPSS Inc. Necesitará este número cuando se ponga en contacto con SPSS Inc. para recibir información sobre asistencia, formas de pago o actualización del sistema. El número de serie se incluye en el sistema Base de SPSS.

Servicio al cliente

Si tiene cualquier duda referente a la forma de envío o pago, póngase en contacto con su oficina local, que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>. Tenga preparado su número de serie para identificarse.

Cursos de preparación

SPSS Inc. ofrece cursos de preparación, tanto públicos como in situ. En todos los cursos habrá talleres prácticos. Estos cursos tendrán lugar periódicamente en las principales capitales. Si desea obtener más información sobre estos cursos, póngase en contacto con su oficina local que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>.

Asistencia técnica

El servicio de asistencia técnica de SPSS está a disposición de todos los clientes registrados. Los clientes podrán ponerse en contacto con este servicio de asistencia técnica si desean recibir ayuda sobre la utilización de SPSS o sobre la instalación en alguno de los entornos de hardware admitidos. Para ponerse en contacto con el servicio de asistencia técnica, consulte la página Web de SPSS en <http://www.spss.com>, o póngase en contacto con la oficina más cercana, que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>. Tenga preparada la información necesaria para identificarse personalmente, a su organización y el número de serie de su sistema.

Publicaciones adicionales

Puede adquirir copias adicionales de los manuales de los productos de SPSS directamente de SPSS Inc. Visite la sección Store de la página Web de SPSS en <http://www.spss.com/estore> o póngase en contacto con su oficina de SPSS local que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>. Para pedidos telefónicos en Estados Unidos y Canadá, llame a SPSS Inc. al 800-543-2185. Para pedidos telefónicos desde otros países, póngase en contacto con la oficina más cercana que encontrará en la página Web de SPSS.

El libro *SPSS Statistical Procedures Companion*, de Marija Norušis, ha sido publicado por Prentice Hall. Se prevé una nueva versión de este libro, actualizado para SPSS 13.0. El libro *SPSS Advanced Statistical Procedures Companion*, que también se basa en SPSS 13.0, se publicará muy pronto. El libro *SPSS Guide to Data Analysis* para SPSS 13.0 también se está elaborando. Las publicaciones anunciadas de forma exclusiva por Prentice Hall estarán disponibles en la página Web de SPSS en <http://www.spss.com/estore> (seleccione su país de origen y pulse en Books).

Díganos su opinión

Sus comentarios son importantes. Háganos saber su experiencia con los productos SPSS. Nos interesa especialmente recibir noticias sobre aplicaciones nuevas e interesantes para el sistema SPSS. Envíenos un correo electrónico a suggest@spss.com o escriba a SPSS Inc., Attn.: Director of Product Planning, 233 South Wacker Drive, 11th Floor, Chicago, IL 60606-6412, EE.UU.

Acerca de este manual

Este manual es la documentación de la interfaz gráfica del usuario para los procedimientos incluidos en Base system. Las ilustraciones de los cuadros de diálogo están tomadas de SPSS para Windows. Los cuadros de diálogo de los demás sistemas operativos son similares. La sintaxis de comandos completa se documenta de forma detallada en la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*), disponible en el menú Ayuda.

Cómo ponerse en contacto con SPSS

Si desea pertenecer a nuestra lista de correo, póngase en contacto con nuestras oficinas que encontrará en la página Web en <http://www.spss.com/worldwide>.

Contenido

1	Resumen	1
	¿Cuáles son las novedades de SPSS 13.0?	2
	Ventanas	4
	Menús	7
	Barra de estado	7
	Cuadros de diálogo	8
	Nombres y etiquetas de variable en las listas de los cuadros de diálogo	9
	Controles del cuadro de diálogo	10
	Subcuadros de diálogo	10
	Selección de variables	11
	Para obtener información sobre las variables en un cuadro de diálogo	11
	Para obtener información sobre los controles del cuadro de diálogo	12
	Pasos básicos en el análisis de datos	13
	Asesor estadístico	13
	Información adicional sobre SPSS	13
2	Obtención de ayuda	15
	Uso de la tabla de contenido de la Ayuda	17
	Uso del índice de la ayuda	18
	Obtención de ayuda en los controles de un cuadro de diálogo	18
	Obtención de ayuda sobre los términos de los resultados	19
	Uso de los estudios de casos	20
	Copia del texto de ayuda de una ventana emergente	20

3 Archivos de datos

21

Apertura de un archivo de datos	21
Para abrir archivos de datos.	21
Tipos de archivos de datos	22
Opciones de apertura de archivos	23
Lectura de archivos de Excel	23
Lectura de archivos de Excel de versiones anteriores y otras hojas de cálculo en el Editor de datos Hojas de cálculo	24
Lectura de archivos de dBASE en el Editor de datos	25
Lectura de archivos de bases de datos.	25
Selección de una fuente de datos.	26
Acceso a la base de datos	28
Selección de campos de datos	29
Creación de una consulta con parámetros	37
Adición de Datos	38
Definición de variables	40
Ordenación de casos	41
Resultados	42
Asistente para la importación de texto	44
Información sobre el archivo	57
Almacenamiento de archivos de datos	57
Almacenamiento de archivos de datos modificados	58
Almacenamiento de archivos de datos en formato de Excel	58
Almacenamiento de archivos de datos en formato de SAS	59
Para guardar un archivo de datos nuevo o guardar los datos con un formato distinto	61
Almacenamiento de datos: Tipos de archivos de datos.	62
Almacenamiento de subconjuntos de variables	64

Almacenamiento de archivos: Opciones	65
Protección de datos originales	65
Archivo activo virtual	66
4 <i>Análisis en modo distribuido</i>	71
Análisis distribuido respecto al análisis local	71
5 <i>Editor de datos</i>	83
Vista de datos	84
Vista de variables	85
Introducción de datos	97
Edición de datos	100
Ir a caso	103
Estado de selección de casos en el Editor de datos	104
Editor de datos: Opciones de presentación	104
Impresión en el Editor de datos	106
6 <i>Preparación de datos</i>	107
Definición de propiedades de variables	108
Copia de propiedades de datos	116
Identificación de casos duplicados	125
Categorizador visual	128
Categorización de variables	131
Generación automática de categorías	134

Copia de categorías	137
Valores perdidos definidos por el usuario en el Categorizador visual.	139

7 Transformaciones de los datos 141

Cálculo de variables	141
Funciones.	144
Valores perdidos en funciones	145
Generadores de números aleatorios	146
Contar apariciones de valores dentro de los casos	147
Recodificación de valores	149
Recodificar en las mismas variables	149
Recodificar en distintas variables	152
Asignar rangos a los casos	156
Recodificación automática	160
Asistente para fecha y hora	163
Transformaciones de los datos de serie temporal	181
Puntuación de datos con modelos predictivos	189

8 Gestión y transformación de los archivos 191

Ordenar casos	192
Transponer	193
Fusión de archivos de datos	193
Añadir casos	194
Añadir variables	197
Agregar datos.	200
Segmentar archivo	205

Seleccionar casos	206
Ponderar casos	211
Reestructuración de los datos	213

9 Trabajo con resultados 239

Visor.	239
Uso de los resultados en otras aplicaciones	248
Pegado de objetos en el Visor.	252
Pegado especial.	252
Para pegar un objeto de otra aplicación en el Visor	253
Exportar resultados	253
Impresión de documentos del Visor	266
Almacenamiento de resultados.	273

10 Visor de borrador 275

Para crear resultados de borrador	276
Control del formato de los resultados de borrador.	277
Atributos de fuente.	282
Para imprimir resultados de borrador	282
Almacenamiento de los resultados del Visor de borrador	284

11 Tablas pivote 285

Gestión de una tabla pivote.	285
Trabajo con capas	290

Señalizadores	294
Presentación y ocultación de casillas	295
Edición de los resultados	297
Cambio del aspecto de las tablas	298
Propiedades de tabla	300
Para modificar las propiedades de las tablas pivote	300
Propiedades de tabla: General	301
Para modificar las propiedades generales de las tablas	301
Propiedades de tabla: Notas al pie	302
Para modificar las propiedades de los marcadores de las notas al pie	302
Propiedades de tabla: Formatos de casilla	303
Para cambiar los formatos de casilla	304
Propiedades de tabla: Bordes.	305
Para cambiar los bordes de una tabla.	305
Para mostrar los bordes ocultos de una tabla pivote	306
Propiedades de tabla: Imprimiendo	306
Para controlar la impresión de las tablas pivote	307
Fuente	307
Anchos de casillas de datos	308
Propiedades de casilla	309
Para modificar las propiedades de una casilla	309
Propiedades de casilla: Valor	310
Para cambiar los formatos del valor de una casilla	310
Para cambiar los formatos de los valores de una columna	311
Propiedades de casilla: Alineación.	311
Para cambiar la alineación de las casillas	312
Propiedades de casilla: Márgenes	313
Para cambiar los márgenes de las casillas	313
Propiedades de casilla: Sombreado	314
Para cambiar el sombreado de las casillas	314

Marcador de nota al pie	315
Selección de filas y columnas en las tablas pivote	316
Para seleccionar una fila o una columna de una tabla pivote	316
Modificación de los resultados de las tablas pivote	316
Impresión de tablas pivote	317
Para imprimir las capas ocultas de una tabla pivote	318
Control de las rupturas de tabla para tablas anchas y largas.	318

12 Trabajar con sintaxis de comandos 321

Reglas de la sintaxis	322
Pegar sintaxis desde cuadros de diálogo	323
Copia de la sintaxis desde las anotaciones de los resultados	324
Edición de la sintaxis en un archivo diario.	326
Para ejecutar la sintaxis de comandos	328
Varios comandos Ejecutar	328

13 Frecuencias 331

Frecuencias: Estadísticos.	334
Frecuencias: Gráficos	336
Frecuencias: Formato.	337

14 Descriptivos 339

Descriptivos: Opciones.	341
---------------------------------	-----

15 Explorar 343

Explorar: Estadísticos 347
Explorar: Gráficos. 348
Explorar: Opciones. 350

16 Tablas de contingencia 351

Capas de las tablas de contingencia. 354
Gráficos de barras agrupadas de las tablas de contingencia 354
Tablas de contingencia: Estadísticos 354
Tablas de contingencia: Mostrar en las casillas 358
Tablas de contingencia: Formato de tabla 360

17 Resumir 361

Resumir: Opciones 363
Resumir: Estadísticos 364

18 Medias 367

Medias: Opciones 370

19 Cubos OLAP 373

Cubos OLAP: Estadísticos 376

Cubos OLAP: Diferencias	379
Cubos OLAP: Título	380

20 Pruebas T 381

Prueba T para muestras independientes	381
Prueba T para muestras relacionadas	385
Prueba T para una muestra	389

21 ANOVA de un factor 393

ANOVA de un factor: Contrastes	396
ANOVA de un factor: Contrastes post hoc	397
ANOVA de un factor: Opciones	400

22 MLG Análisis univariante 403

MLG: Modelo	407
MLG: Contrastes	410
MLG: Gráficos de perfil	412
MLG: Comparaciones post hoc	413
MLG: Guardar	416
Opciones MLG	418
Funciones adicionales de los comandos UNIANOVA	419

23 Correlaciones bivariadas 421

Correlaciones bivariadas: Opciones 423
Funciones adicionales de los comandos CORRELATIONS y NONPAR
CORR 424

24 Correlaciones parciales 425

Correlaciones parciales: Opciones 428

25 Distancias 429

Distancias: Medidas de disimilaridad 431
Distancias: Medidas de similaridad 432

26 Regresión lineal 435

Métodos de selección de variables en el análisis de regresión lineal 439
Regresión lineal: Establecer regla 441
Regresión lineal: Gráficos. 441
Regresión lineal: Para guardar variables nuevas 443
Regresión lineal: Estadísticos. 446
Regresión lineal: Opciones. 448

27 Estimación curvilínea **451**

Modelos del procedimiento Estimación curvilínea	455
Estimación curvilínea: Guardar	456

28 Análisis discriminante **459**

Análisis discriminante: Definir rango	463
Análisis discriminante: Seleccionar casos	463
Análisis discriminante: Estadísticos	464
Análisis discriminante: Método de inclusión por pasos	465
Análisis discriminante: Clasificar	467
Análisis discriminante: Guardar	468

29 análisis factorial **471**

Selección de casos en el análisis factorial	477
Análisis factorial: Descriptivos	478
Análisis factorial: Extracción	479
Análisis factorial: Rotación	481
Análisis factorial: Puntuaciones factoriales	483
Análisis factorial: Opciones	484

30 Selección de procedimientos para la conglomeración	485
31 Análisis de conglomerados en dos fases	487
Opciones del análisis de conglomerados en dos fases	491
Gráficos de análisis de conglomerados en dos fases	494
Resultados de análisis de conglomerados en dos fases	495
32 Análisis de conglomerados jerárquico	497
Análisis de conglomerados jerárquico: Método	501
Análisis de conglomerados jerárquico: Estadísticos	502
Análisis de conglomerados jerárquico: Gráficos.	503
Análisis de conglomerados jerárquico: Guardar variables nuevas	503
33 Análisis de conglomerados de K-medias	505
Eficacia del análisis de conglomerados de K-medias	510
Análisis de conglomerados de K-medias: Iterar	511
Análisis de conglomerados de K-medias: Guardar	512
Análisis de conglomerados de K-medias: Opciones	512

34 Pruebas no paramétricas **515**

Prueba de chi-cuadrado	516
Prueba binomial	521
Prueba de rachas	523
Prueba de Kolmogorov-Smirnov para una muestra	526
Pruebas para dos muestras independientes	529
Pruebas para dos muestras relacionadas	533
Pruebas para varias muestras independientes	537
Pruebas para varias muestras relacionadas	541

35 Análisis de respuestas múltiples **545**

Definir conjuntos de respuestas múltiples	546
Frecuencias de respuestas múltiples	548
Tablas de contingencia de respuestas múltiples	550
Tablas de respuestas múltiples: Definir rangos de las variables	552
Tablas de contingencia de respuestas múltiples: Opciones	553
Funciones adicionales del comando MULT RESPONSE	554

36 Informes de los resultados **555**

Informe de estadísticos en filas	555
Informe de estadísticos en columnas	564
Funciones adicionales del comando REPORT	570

37 *Análisis de fiabilidad* 571

Análisis de fiabilidad: Estadísticos 573
Funciones adicionales del comando RELIABILITY 575

38 *Escalamiento multidimensional* 577

Escalamiento multidimensional: Forma de los datos 579
Escalamiento multidimensional: Crear la medida a partir de los datos. 580
Escalamiento multidimensional: Modelo 581
Escalamiento multidimensional: Opciones 582
Funciones adicionales del comando de ALSCAL 583

39 *Estadísticos de la razón* 585

Estadísticos de la razón 587

40 *Conceptos básicos de la utilidad de gráficos* 589

Creación y modificación de gráficos 589
Opciones de definición de gráfico 594

41 *Curvas COR* 603

Curvas COR: Opciones 606

42 Utilidades 607

Información sobre la variable	607
Comentarios del archivo de datos.	608
Conjuntos de variables	609
Definir conjuntos de variables	609
Usar conjuntos	610
Reordenación de listas de variables de destino.	611

43 Opciones 613

General: Opciones	614
Opciones del Visor	616
Opciones del Visor de borrador.	617
Etiqueta de los resultados: Opciones	619
Opciones de gráfico	621
Opciones de gráficos interactivos.	626
Tabla pivote: Opciones	627
Datos: Opciones.	630
Moneda: Opciones.	631
Procesos: Opciones	633

44 Personalización de menús y barras de herramientas 635

Editor de menús	635
Personalización de las barras de herramientas.	636
Mostrar barras de herramientas.	637

Para personalizar las barras de herramientas	637
--	-----

45 Unidad de producción 643

Uso de la Unidad de producción	645
Opciones de exportación	646
Entradas del usuario	648
Solicitud de entrada del usuario en los trabajos de producción con macros	650
Opciones de producción	651
Control del formato en los trabajos de producción	652
Ejecución de trabajos de producción desde una línea de comandos	655
Publicar en Web	657
Inicio de sesión del servidor Web de SmartViewer	658

46 Utilidad de procesamiento de SPSS 659

Para ejecutar un proceso	659
Procesos incluidos en SPSS	660
Autoprocesos	661
Creación y edición de procesos	663
Para editar un proceso	663
Ventana de proceso	664
Procesos de iniciación	666
Creación de autoprosesos	668
Funcionamiento de los procesos	672
Clases de objetos de SPSS y convenciones de denominación de variables	675
Procedimiento nuevo (Proceso)	680

Adición de una descripción a un proceso	683
Procesamiento de cuadros de diálogo personalizados	683
Depuración de procesos.	688
Archivos de procesos y archivos de sintaxis.	691

47 Sistema de gestión de resultados 695

Tipos de objetos de resultados	699
Identificadores de comandos y subtipos de tabla	701
Etiquetas de tabla.	702
Opciones de SGR	704
Registro	710
Exclusión de presentación de resultados del Visor	710
Envío de resultados a archivos de datos de SPSS.	711
Estructura de tablas OXML.	721
Identificadores SGR	725

Apéndices

A Administrador de acceso a bases de datos 729

B Personalización de documentos HTML 731

Para añadir código HTML personalizado a los documentos de resultados exportados	731
Contenido y formato del archivo de texto para HTML personalizado	732

Para utilizar otro archivo o ubicación para el código HTML personalizado. 732

Índice

735

Resumen

SPSS para Windows proporciona un poderoso sistema de análisis estadístico y de gestión de datos en un entorno gráfico, utilizando menús descriptivos y cuadros de diálogo sencillos que realizan la mayor parte del trabajo. La mayoría de las tareas se pueden llevar a cabo simplemente situando el puntero del ratón en el lugar deseado y pulsando en el botón.

Además de la simple interfaz de los análisis estadísticos consistente en situar el puntero y pulsar, SPSS para Windows proporciona:

Editor de datos. Sistema versátil, similar a una hoja de cálculo, para definir, introducir, editar y presentar datos.

Visor. El Visor permite examinar los resultados, mostrarlos y ocultarlos de forma selectiva, modificar el orden de presentación en la pantalla y desplazar tablas y gráficos de gran calidad entre SPSS y otras aplicaciones.

Tablas de pivote multidimensionales. Sus resultados cobrarán vida en las tablas pivote multidimensionales. Explore las tablas reorganizando las filas, las columnas y las capas. Haga importantes descubrimientos que suelen quedar velados en los informes convencionales. Compare fácilmente los grupos dividiendo la tabla de manera que aparezca solamente un grupo cada vez.

Gráficos de alta resolución. Como funciones básicas de SPSS se incluyen gráficos de sectores, gráficos de barras, histogramas, diagramas de dispersión y gráficos 3-D de alta resolución y a todo color, entre muchos otros.

Acceso a bases de datos. Obtenga información de bases de datos mediante el Asistente para bases de datos en lugar de utilizar consultas SQL de gran complejidad.

Transformaciones de los datos. Las funciones de transformación le ayudarán a preparar sus datos para el análisis. Puede crear fácilmente subconjuntos de datos, combinar categorías y añadir, agregar, fusionar, segmentar y transponer archivos, entre muchas otras posibilidades.

Distribución electrónica. Se pueden enviar informes por correo electrónico pulsando en un botón o exportar tablas y gráficos en formato HTML para distribuirlos por Internet o dentro de una Intranet.

Ayuda en pantalla. Los tutoriales le ofrecerán una introducción global ampliamente detallada, los temas de la Ayuda sensible al contexto de los cuadros de diálogo le guiarán a través de la ejecución de tareas específicas, las definiciones de las ventanas emergentes para los resultados de las tablas pivote le explicarán los términos estadísticos y Estudios de casos le proporcionará ejemplos prácticos sobre cómo utilizar los procedimientos estadísticos y cómo interpretar los resultados.

Lenguaje de comandos. Aunque la mayoría de las tareas se pueden llevar a cabo simplemente situando el puntero del ratón en el lugar deseado y pulsando en el botón, SPSS proporciona además un potente lenguaje de comandos que permite guardar y automatizar muchas tareas comunes. El lenguaje de comandos también proporciona algunas funcionalidades no incluidas en los menús y cuadros de diálogo.

La documentación completa sobre la sintaxis de comandos se instala automáticamente al mismo tiempo que SPSS. Para acceder a la documentación sobre la sintaxis:

- ▶ Elija en los menús
 - ?
 - Referencia de sintaxis de comandos

¿Cuáles son las novedades de SPSS 13.0?

Hay muchas funciones nuevas en SPSS 13.0.

Gestión de datos

- El Asistente de fecha y hora facilita la realización de muchos cálculos con fechas y horas, incluyendo el cálculo de la diferencia entre las fechas y la adición o sustracción de una duración a partir de una fecha. Si desea obtener más información, consulte “Asistente para fecha y hora” en Capítulo 7 en p. 163.

- Puede añadir resultados agregados al archivo de datos de trabajo. Si desea obtener más información, consulte “Agregar datos” en Capítulo 8 en p. 200.
- Puede crear esquemas de recodificación automática coherentes para diversas variables y archivos de datos. Si desea obtener más información, consulte “Recodificación automática” en Capítulo 7 en p. 160.
- Los valores de cadena pueden tener hasta 32.767 bytes. Previamente, el límite era de 255.
- Puede crear varios paneles en la Vista de datos del Editor de datos. Si desea obtener más información, consulte “Editor de datos: Opciones de presentación” en Capítulo 5 en p. 104.

Gráficos

- gráficos de barras 3-D.
- Pirámides de población.
- Gráficos de puntos.
- Gráficos con paneles.

Mejoras en los estadísticos

- Nueva opción Árbol de clasificación para generar modelos de árbol.
- Nuevo procedimiento MLG en la opción Muestras complejas.
- Nuevo procedimiento Regresión logística en la opción Muestras complejas.
- Nuevo procedimiento Correspondencia múltiple en la opción Categorías.
- Los estadísticos AIC y BIC se añaden a Regresión logística multinomial en la opción Modelos de regresión y existe además la posibilidad de especificar el tipo de estadístico utilizado para determinar la adición o eliminación de términos de modelo al utilizar varios métodos de pasos sucesivos.

Resultados

- Panel de control del Sistema de gestión de resultados. Si desea obtener más información, consulte “Sistema de gestión de resultados” en Capítulo 47 en p. 695.
- Exportar resultados del Visor a PowerPoint. Si desea obtener más información, consulte “Exportar resultados” en Capítulo 9 en p. 253.

- Más opciones de ordenación de resultados y la posibilidad de ocultar las categorías de subtotal en Tablas personalizadas (opción Tablas).
- Resultados de las tablas pivote para Estimación curvilínea y Respuesta múltiple en el sistema Base, todos los procedimientos de series temporales (incluyendo la opción Tendencias), Kaplan-Meier y Tablas de mortalidad de la opción Modelos avanzados y Regresión no lineal de la opción Modelos de regresión.

Sintaxis de comandos

- Puede controlar el directorio de trabajo con el nuevo comando `CD`.
- Puede controlar el tratamiento de errores en los archivos de sintaxis de comandos incluidos con el nuevo comando `INSERT`.
- Puede utilizar el comando `FILE HANDLE` para definir las rutas de directorio.

Si desea obtener más información sobre la sintaxis de comandos, consulte el archivo PDF de la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*), al que puede acceder desde el menú Ayuda.

Servidor de SPSS

- Puede puntuar los datos basados en modelos generados con muchos procedimientos de SPSS. Si desea obtener más información, consulte “Puntuación de datos con modelos predictivos” en Capítulo 7 en p. 189.
- Puede trabajar con fuentes de bases de datos de forma más eficaz si agrega u ordena previamente los datos en la base de datos antes de leerlos en SPSS. Si desea obtener más información, consulte “Adición de Datos” en Capítulo 3 en p. 38.

Ventanas

Existen diversos tipos de ventanas en SPSS:

Editor de datos. Esta ventana muestra el contenido del archivo de datos. Con el Editor de datos puede crear nuevos archivos de datos o modificar los existentes. La ventana Editor de datos se abre automáticamente cuando se inicia una sesión de SPSS. No se puede tener más de un archivo de datos abierto al mismo tiempo.

Visor. Todas las tablas, los gráficos y los resultados estadísticos se muestran en el Visor. Puede editar los resultados y guardarlos para utilizarlos posteriormente. La ventana del Visor se abre automáticamente la primera vez que se ejecuta un procedimiento que genera resultados.

Visor de borrador. Los resultados pueden mostrarse como texto normal (en lugar de como tablas pivote interactivas) en el Visor de borrador.

Editor de tablas pivote. Con el Editor de tablas pivote es posible modificar los resultados mostrados en este tipo de tablas de diversas maneras. Puede editar el texto, intercambiar los datos de las filas y las columnas, añadir colores, crear tablas multidimensionales y ocultar y mostrar los resultados de manera selectiva.

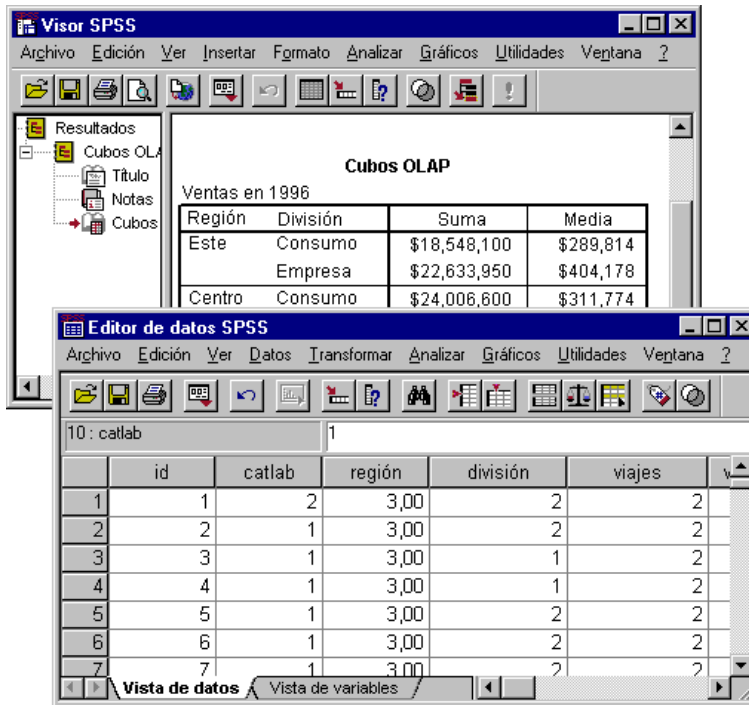
Editor de gráficos. Puede modificar los gráficos y diagramas de alta resolución en las ventanas de los gráficos. Es posible cambiar los colores, seleccionar diferentes tipos de fuentes y tamaños, intercambiar los ejes horizontal y vertical, rotar diagramas de dispersión 3-D e incluso cambiar el tipo de gráfico.

Editor de resultados de texto. Los resultados de texto que no aparecen en las tablas pivote pueden modificarse con el Editor de resultados de texto. Puede editar los resultados y cambiar las características de las fuentes (tipo, estilo, color y tamaño).

Editor de sintaxis. Puede pegar las selecciones del cuadro de diálogo en una ventana de sintaxis, donde aparecerán en forma de sintaxis de comandos. A continuación puede editar esta sintaxis para utilizar las funciones especiales de SPSS que no se encuentran disponibles en los cuadros de diálogo. También puede guardar los comandos en un archivo para utilizarlos en sesiones de SPSS posteriores.

Editor de procesos. El procesamiento y la automatización OLE permiten personalizar y automatizar muchas tareas en SPSS. Utilice el Editor de procesos para crear y modificar los procesos básicos.

Figura 1-1
Editor de datos y Visor



Ventana designada frente a ventana activa

Si tiene abiertas varias ventanas del Visor, los resultados se dirigirán hacia la ventana **designada** del Visor. Si tiene abierta más de una ventana del Editor de sintaxis, la sintaxis de comandos se pegará en la ventana designada del Editor de sintaxis. Las ventanas designadas se distinguen por un signo de exclamación (!) en la barra de estado y es posible cambiarlas en cualquier momento.

La ventana designada no debe confundirse con la ventana **activa**, que es la ventana actualmente seleccionada. Si tiene ventanas superpuestas, la ventana activa es la que aparece en primer plano. Si abre una ventana, esa ventana se convertirá automáticamente en la ventana activa y en la ventana designada.

Para cambiar la ventana designada del Visor o del Editor de sintaxis

- ▶ Convierta la ventana que desee designar en la ventana activa (pulse en cualquier punto de la ventana).
- ▶ Pulse en la herramienta Designar ventana de la barra de herramientas (la que tiene el signo de exclamación).

o

- ▶ Elija en los menús:
Utilidades
Designar ventana

Figura 1-2
Herramienta Designar ventana

Cambiar la ventana designada



Menús

Muchas de las tareas que desea realizar en SPSS comienzan por las selecciones de menús. Cada ventana de SPSS tiene su propia barra de menús con los menús adecuados para ese tipo de ventana.

Los menús Analizar y Gráficos aparecen en todas las ventanas, lo cual permite generar fácilmente nuevos resultados sin tener que cambiar de ventana.

Barra de estado

La barra de estado que aparece en la parte inferior de cada ventana de SPSS proporciona la siguiente información:

Estado del comando. En cada procedimiento o comando que se ejecuta, un recuento de casos indica el número de casos procesados hasta el momento. En los procedimientos estadísticos que requieren procesamientos iterativos, se muestra el número de iteraciones.

Estado del filtro. Si ha seleccionado una muestra aleatoria o un subconjunto de casos para el análisis, el mensaje Filtrado indica que existe algún tipo de filtrado activado actualmente y por tanto en el análisis no se incluyen todos los casos del archivo de datos.

Estado de ponderación. El mensaje Ponderado indica que se está utilizando una variable de ponderación para ponderar los casos para el análisis.

Estado de división del archivo. El mensaje Segmentado indica que el archivo de datos se ha segmentado en diferentes grupos para su análisis en función de los valores de una o más variables de agrupación.

Para mostrar u ocultar la barra de estado

- ▶ Elija en los menús:
 - Ver
 - Barra de estado

Cuadros de diálogo

La mayoría de las opciones de los menús acceden a un cuadro de diálogo cuando se seleccionan. Los cuadros de diálogo se utilizan para seleccionar variables y opciones para el análisis.

Los cuadros de diálogo para los procedimientos estadísticos tienen normalmente dos componentes básicos:

Lista de variables de origen. Lista de variables del archivo de datos de trabajo. En la lista de origen sólo aparecen los tipos de variables que permiten el procedimiento seleccionado. La utilización de variables de cadena corta y de cadena larga está restringida en muchos procedimientos.

Listas de variables de destino. Una o varias listas que indican las variables elegidas para el análisis; un ejemplo son las listas de variables dependientes e independientes.

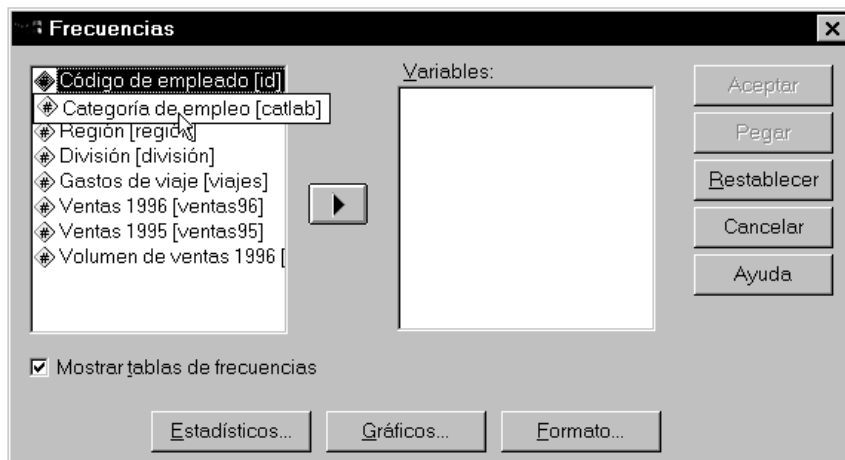
Nombres y etiquetas de variable en las listas de los cuadros de diálogo

Es posible elegir entre la presentación de nombres de variable o de etiquetas de variable en las listas de los cuadros de diálogo.

- Para determinar la presentación de nombres o de etiquetas de variable, seleccione Opciones en el menú Edición de cualquier ventana.
- Para definir o modificar las etiquetas de variable, elija Vista de variables en el Editor de datos.
- Para los datos importados de orígenes de bases de datos se utilizan los nombres de los campos como etiquetas de variable.
- Si las etiquetas son largas, sitúe el puntero sobre la etiqueta de la lista para verla completa.
- Si no se define ninguna etiqueta de variable se mostrará el nombre de la variable.

Figura 1-3

Etiquetas de variable mostradas en un cuadro de diálogo



Controles del cuadro de diálogo

Existen cinco controles estándares en la mayoría de los cuadros de diálogo:

Aceptar. Ejecuta el procedimiento. Después de seleccionar las variables y elegir las especificaciones adicionales, pulse en Aceptar para ejecutar el procedimiento. Además produce el cierre del cuadro de diálogo.

Pegar. Genera la sintaxis de comandos a partir de las selecciones del cuadro de diálogo y la pega en la ventana de sintaxis. A continuación, puede personalizar los comandos con funciones adicionales que no se encuentran disponibles en los cuadros de diálogo.

Restablecer. Desactiva las variables en las listas de variables seleccionadas y restablece todas las especificaciones del cuadro de diálogo y los subcuadros de diálogo al estado por defecto.

Cancelar. Cancela los cambios de las selecciones del cuadro de diálogo desde la última vez que se abrió y lo cierra. Durante una sesión se mantienen las selecciones del cuadro de diálogo. El cuadro de diálogo retiene el último conjunto de especificaciones hasta que se anulan.

Ayuda. Ayuda contextual. Este botón le lleva a una ventana de Ayuda con información sobre el cuadro de diálogo actual. También puede obtener ayuda sobre los controles individuales del cuadro de diálogo pulsando en el botón derecho del ratón.

Subcuadros de diálogo

Puesto que la mayoría de los procedimientos ofrecen un gran número de posibilidades, un solo cuadro de diálogo no puede contener todas las opciones disponibles. El cuadro de diálogo principal ofrece normalmente la información mínima que se requiere para ejecutar un procedimiento. Las especificaciones adicionales se realizan en los subcuadros de diálogo.

En el cuadro de diálogo principal, los controles con puntos suspensivos (...) después del nombre indican que se mostrará un subcuadro de diálogo.

Selección de variables

Para seleccionar una sola variable, es suficiente con resaltarla en la lista de variables de origen y pulsar en el botón de flecha derecha situado junto a la lista de variables de destino. Si sólo existe una lista de variables de destino, puede pulsar dos veces en las variables individuales para desplazarlas desde la lista de origen hasta la de destino.

También pueden seleccionar diversas variables:

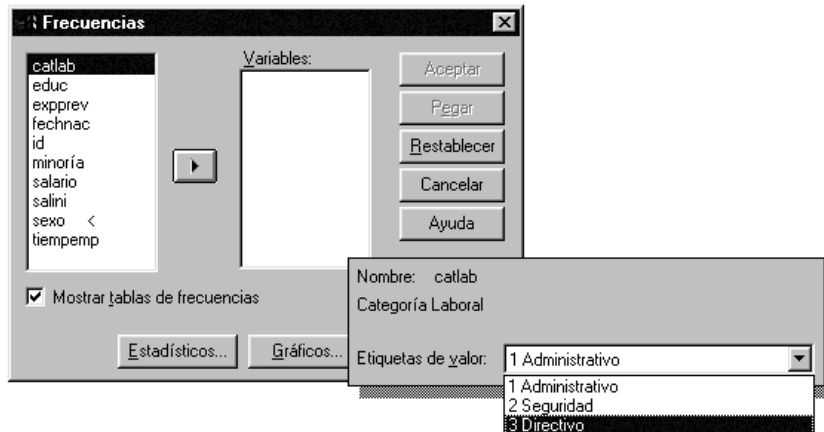
- Para seleccionar diversas variables que estén agrupadas en la lista, pulse en la primera de ellas y, después, en la última del grupo mientras mantiene pulsada la tecla Mayús.
- Para resaltar diversas variables que no estén agrupadas en la lista, utilice el método siguiente: pulse en la primera variable, luego pulse en la siguiente variable mientras mantiene la tecla Ctrl pulsada, y así sucesivamente.

Para obtener información sobre las variables en un cuadro de diálogo

- ▶ Pulse con el botón derecho del ratón en la lista de variables de origen o de destino.
- ▶ Pulse en Información de las variables en el menú contextual emergente.

Figura 1-4

Acceso a la información sobre la variable con el botón derecho del ratón



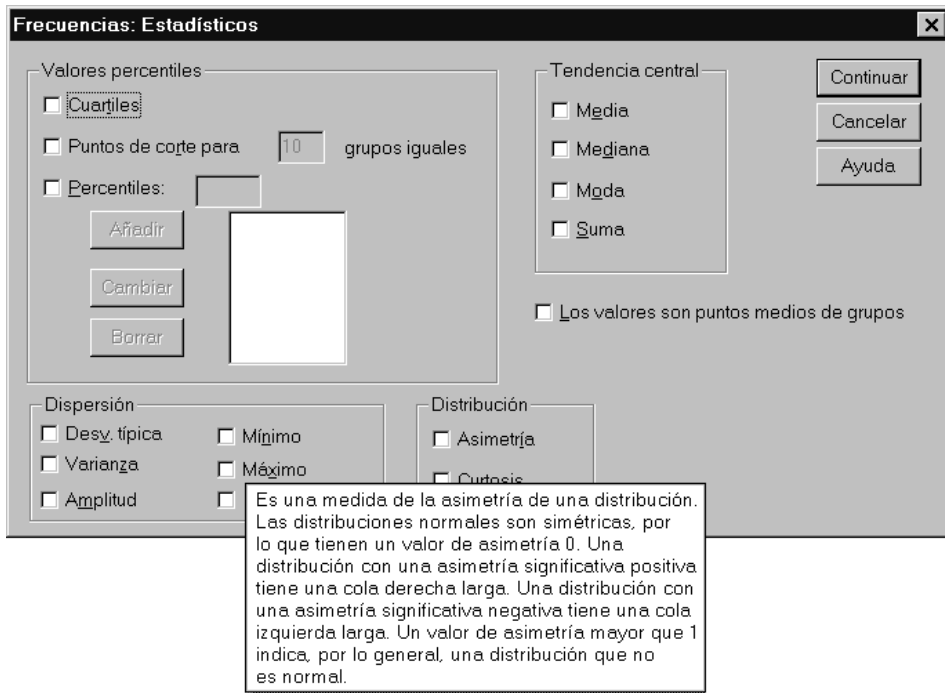
Para obtener información sobre los controles del cuadro de diálogo

- ▶ Pulse con el botón derecho del ratón en el control sobre el que desee obtener información.
- ▶ Elija ¿Qué es esto? en el menú contextual emergente.

Aparecerá una ventana emergente que presentará información acerca del control.

Figura 1-5

Ayuda emergente ¿Qué es esto? del botón derecho del ratón sobre los controles de los cuadros de diálogo



Pasos básicos en el análisis de datos

Analizar datos con SPSS es fácil. Simplemente tiene que:

Introducir los datos en SPSS. Es posible abrir un archivo de datos de SPSS previamente guardado, leer una hoja de cálculo o un archivo de datos de texto, o introducir los datos directamente en el Editor de datos.

Seleccionar un procedimiento. Seleccione un procedimiento de los menús para calcular estadísticos o crear un gráfico.

Seleccionar las variables para el análisis. Las variables del archivo de datos se muestran en un cuadro de diálogo para el procedimiento.

Ejecute el procedimiento y observe los resultados. Los resultados se muestran en el Visor.

Asesor estadístico

Si no está familiarizado con SPSS o con los procedimientos estadísticos que ofrece, el asesor estadístico (Statistics Coach) puede ayudarle solicitándole información mediante preguntas simples, utilizando un lenguaje no técnico y ejemplos visuales que le ayudarán a seleccionar las funciones básicas de los estadísticos y los gráficos más apropiados para sus datos.

Para utilizar el Asesor estadístico, elija en los menús de cualquier ventana de SPSS:

?

Asesor estadístico

El Asesor estadístico cubre sólo un subconjunto selecto de procedimientos del sistema Base de SPSS. Está diseñado para proporcionar una asistencia general para muchas de las técnicas estadísticas básicas que se utilizan habitualmente.

Información adicional sobre SPSS

Si desea obtener una introducción global más detallada sobre los aspectos básicos de SPSS, consulte el tutorial en pantalla. En cualquier menú de SPSS elija:

?

Tutorial

Obtención de ayuda

La ayuda se proporciona de diversas formas:

Menú Ayuda. En la mayoría de las ventanas de SPSS, el menú Ayuda proporciona acceso al sistema de ayuda principal además de a los tutoriales y al material de referencia técnica.

- **Temas.** Proporciona acceso a las pestañas Contenido, Índice y Buscar, que pueden usarse para buscar temas específicos de la Ayuda.
- **Tutorial.** Instrucciones ilustradas paso a paso sobre cómo utilizar muchas de las funciones básicas de SPSS. No necesita consultar el tutorial de principio a fin. Puede elegir los temas que desea consultar, pasar de un tema a otro, ver los temas en cualquier orden y utilizar el índice o la tabla de contenidos para buscar temas concretos.
- **Estudios de casos.** Ejemplos prácticos sobre cómo crear diferentes tipos de análisis estadísticos y cómo interpretar los resultados. También se proporcionan los archivos de datos de muestra utilizados en estos ejemplos para que pueda trabajar en dichos ejemplos y observar con exactitud cómo se generaron los resultados. Puede elegir los procedimientos concretos que desee aprender en la tabla de contenidos o buscar los temas correspondientes en el índice.
- **Asesor estadístico.** Método de asistencia para orientarle en el proceso de búsqueda del procedimiento que desea utilizar. Tras realizar una serie de selecciones, el Asesor estadístico abre el cuadro de diálogo para el procedimiento estadístico, de generación de informes o de creación de gráficos que cumple los criterios seleccionados. El Asesor estadístico proporciona acceso a la mayoría de los procedimientos estadísticos y de generación de informes en el sistema Base y en los procedimientos de creación de gráficos.
- **Referencia de sintaxis de comandos.** La sintaxis de comandos completa aparece documentada en la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*), disponible en el menú Ayuda.

Ayuda contextual. En muchos puntos de la interfaz de usuario, puede obtener ayuda sensible al contexto.

- **Botón Ayuda en los cuadros de diálogo.** La mayoría de los cuadros de diálogo disponen de un botón Ayuda que permite acceder directamente al tema de ayuda correspondiente. Este tema proporciona información general y vínculos a los temas relacionados.
- **Ayuda en el menú contextual de un cuadro de diálogo.** Muchos cuadros de diálogo proporcionan ayuda sensible al contexto para controles y funciones individuales. Pulse con el botón derecho del ratón en cualquier control de un cuadro diálogo y seleccione ¿Qué es esto? en el menú contextual que aparece para ver una descripción del control e instrucciones sobre su uso. (Si ¿Qué es esto? no aparece en el menú contextual, esta forma de ayuda no está disponible para dicho cuadro de diálogo.)
- **Ayuda en el menú contextual de una tabla pivote.** Pulse con el botón derecho del ratón en los términos de una tabla pivote activada en el Visor y seleccione ¿Qué es esto? en el menú contextual para ver las definiciones de los términos.
- **Estudios de casos.** Pulse con el botón derecho del ratón en una tabla pivote y seleccione Estudios de casos en el menú contextual para ir directamente a un ejemplo detallado del procedimiento que ha generado la tabla. (Si Estudios de casos no aparece en el menú contextual, esta forma de ayuda no está disponible para dicho procedimiento.)
- **Esquemas de sintaxis de comandos.** En una ventana de sintaxis de comandos, coloque el cursor en cualquier punto de un bloque de sintaxis para un comando y pulse F1 en el teclado. Se muestra un esquema de sintaxis de comandos completo para dicho comando.

Otros recursos. Si no encuentra la información que desea en el sistema de ayuda, estos otros recursos pueden ofrecerle las respuestas que necesita.

- **SPSS para la guía de desarrolladores de Windows.** Proporciona información y ejemplos para las herramientas de desarrollador incluidas con SPSS para Windows, incluyendo la automatización OLE, las API de otros fabricantes, las DLL de entrada/salida, la unidad de producción y la utilidad de procesamiento.

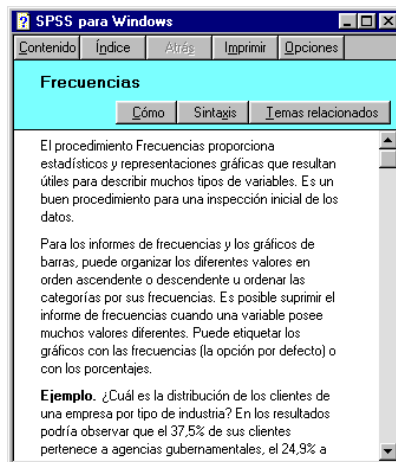
La guía de desarrolladores (*Developer's Guide*) está disponible en formato PDF en el directorio *SPSS\developer* del CD de instalación.

- **Sitio Web del servicio técnico.** Se pueden encontrar respuestas a muchos problemas comunes en <http://support.spss.com>. (El sitio Web de servicio técnico requiere un ID de inicio de sesión y una contraseña. La información sobre cómo obtener el ID y la contraseña se facilita en la dirección URL mencionada anteriormente.)

Uso de la tabla de contenido de la Ayuda

- ▶ Seleccione en cualquiera de las ventanas las siguientes opciones de los menús:
 - ?
Temas
- ▶ Pulse en la pestaña Contenido.
- ▶ Pulse dos veces en los elementos con un icono de libro para expandir o contraer el contenido.
- ▶ Pulse en un elemento para ir a ese tema de la Ayuda.

Figura 2-1
Pestaña Contenido de la ventana de ayuda

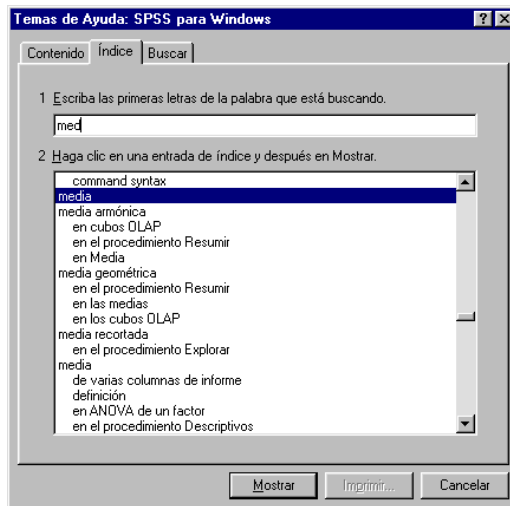


Uso del índice de la ayuda

- ▶ Seleccione en cualquiera de las ventanas las siguientes opciones de los menús:
 - ?
 - Temas
- ▶ Pulse en la pestaña Índice.
- ▶ Introduzca el término que desee buscar en el índice.
- ▶ Pulse dos veces en el tema que desee consultar.

El índice de la ayuda utiliza la búsqueda incremental para buscar el texto introducido y selecciona la coincidencia más aproximada del índice.

Figura 2-2
Pestaña Índice y búsqueda incremental



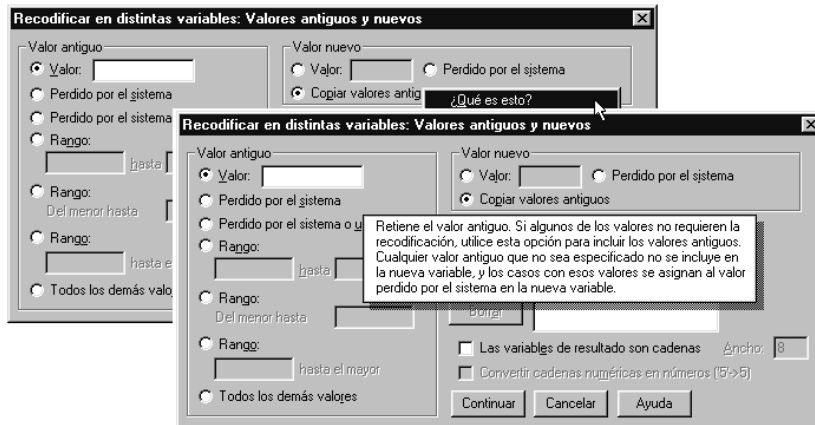
Obtención de ayuda en los controles de un cuadro de diálogo

- ▶ Pulse con el botón derecho del ratón en el control del cuadro de diálogo sobre el que desee obtener información.
- ▶ Seleccione ¿Qué es esto? en el menú contextual emergente.

Aparecerá una ventana emergente con una descripción del control y de cómo utilizarlo. También se puede acceder a la información general sobre el cuadro de diálogo mediante el botón Ayuda que contiene.

Figura 2-3

Ayuda sobre el control del cuadro de diálogo con el botón derecho del ratón



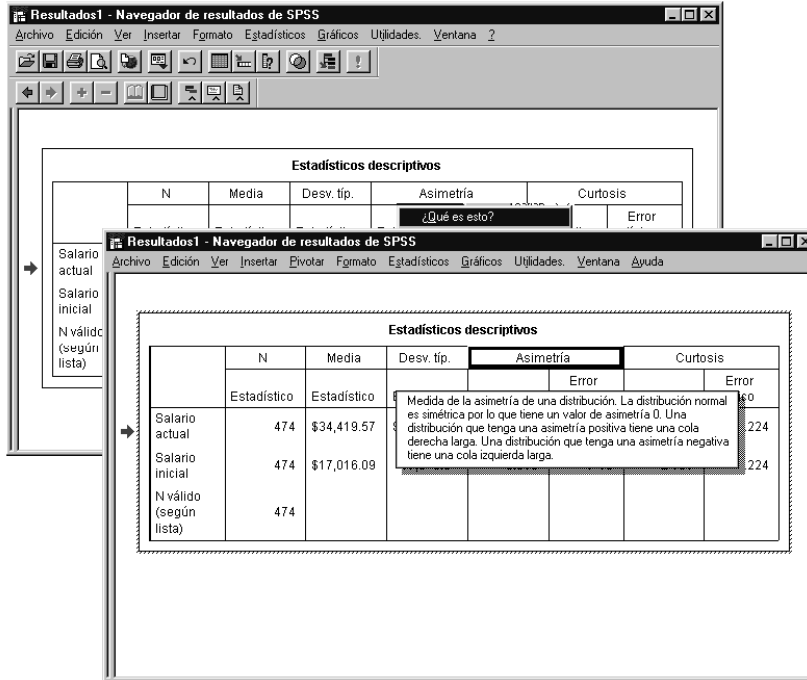
Obtención de ayuda sobre los términos de los resultados

- ▶ Pulse dos veces en la tabla pivote para activarla.
- ▶ Pulse con el botón derecho del ratón en el término del que desee obtener información.
- ▶ Seleccione ¿Qué es esto? en el menú contextual.

Aparecerá una definición del término en una ventana emergente.

Figura 2-4

Ayuda de glosario de la tabla pivote activada con el botón derecho del ratón



Uso de los estudios de casos

- ▶ Pulse con el botón derecho del ratón en una tabla pivote de la ventana del Visor.
- ▶ Elija Estudios de casos en el menú contextual emergente.

Copia del texto de ayuda de una ventana emergente

- ▶ Pulse con el botón derecho del ratón en cualquier parte de la ventana emergente.
- ▶ Elija Copiar en el menú contextual.

Se copiará todo el texto de la ventana emergente.

Archivos de datos

Los archivos de datos pueden tener formatos muy diversos, y este programa se ha sido diseñado para trabajar con muchos de ellos, incluyendo:

- Hojas de cálculo creadas con Lotus 1-2-3 y Excel.
- Archivos de bases de datos creados con dBase y varios formatos SQL.
- Archivos de texto delimitados por tabuladores y otros tipos de archivos de texto ASCII.
- Archivos de datos con formato SPSS creados en otros sistemas operativos.
- Archivos de datos de SYSTAT.
- archivos de datos de SAS

Apertura de un archivo de datos

Además de los archivos guardados en formato SPSS, puede abrir archivos de Excel, Lotus 1-2-3, dBASE y archivos delimitados por tabuladores sin necesidad de convertirlos a un formato intermedio ni de introducir información sobre la definición de los datos.

Para abrir archivos de datos

- ▶ Elija en los menús:
 - Archivo
 - Abrir
 - Datos...
- ▶ En el cuadro de diálogo Abrir archivo, seleccione el archivo que desea abrir.
- ▶ Pulse en Abrir.

Si lo desea, tiene la posibilidad de:

- Leer los nombres de las variables de la primera fila en las hojas de cálculo y en los archivos delimitados por tabuladores.
- Especificar el rango de casillas que desee leer en los archivos de hojas de cálculo.
- Especificar una hoja dentro de un archivo de Excel que desee leer (Excel 5 o versiones posteriores).

Tipos de archivos de datos

SPSS. Abre archivos de datos guardados con formato SPSS, incluyendo SPSS para Windows, Macintosh, UNIX y el producto SPSS/PC+ para DOS.

SPSS/PC+. Abre archivos de datos de SPSS/PC+.

SYSTAT. Abre archivos de datos de SYSTAT.

SPSS portátil. Abre archivos de datos guardados con formato SPSS portátil. El almacenamiento de archivos en este formato lleva mucho más tiempo que guardarlos en formato SPSS.

Excel. Abre archivos de Excel.

Lotus 1-2-3. Abre archivos de datos guardados en formato 1-2-3 en las versiones 3.0, 2.0 o 1A de Lotus.

SYLK. Abre archivos de datos guardados en formato SYLK (vínculo simbólico), un formato utilizado por algunas aplicaciones de hoja de cálculo.

dBASE. Abre archivos con formato dBASE para dBASE IV, dBASE III o III PLUS, o dBASE II. Cada caso es un registro. Las etiquetas de valor y de variable y las especificaciones de valores perdidos se pierden si se guarda un archivo en este formato.

Nombre de archivo largo de SAS. Versión de 7-9 para SAS Windows, extensión larga.

Nombre de archivo corto de SAS. Versión de 7-9 para SAS Windows, extensión corta.

SAS v6 para Windows. Versión 6.08 de SAS para Windows y OS2.

SAS v6 para UNIX. Versión 6 de SAS para UNIX (Sun, HP, IBM).

Transporte de SAS. Archivo de transporte de SAS.

Texto. Archivo de texto ASCII.

Opciones de apertura de archivos

Leer los nombres de variable. En las hojas de cálculo, puede leer los nombres de variable de la primera fila del archivo o de la primera fila del rango definido. Los valores se convertirán según sea preciso para crear nombres de variables válidos, incluyendo la conversión de espacios en subrayados.

Hoja de trabajo. Los archivos de Excel 5 o de versiones posteriores pueden contener varias hojas de trabajo. El Editor de datos lee por defecto la primera hoja. Para leer una diferente, seleccione la que desee en la lista desplegable.

Rango. En los archivos de hoja de cálculo, también puede leer un rango de casillas. Para especificar rangos de casillas utilice el mismo método que empleará en la aplicación de hoja de cálculo.

Lectura de archivos de Excel

Leer los nombres de variable. Puede leer los nombres de variable de la primera fila del archivo o de la primera fila del rango definido. Los valores que no cumplan las normas de denominación de variables se convertirán en nombres de variables válidos y los nombres originales se utilizarán como etiquetas de variable.

Hoja de trabajo. Los archivos de Excel pueden contener varias hojas de trabajo. El Editor de datos lee por defecto la primera hoja. Para leer una diferente, seleccione la que desee en la lista desplegable.

Rango. También puede leer un rango de casillas. Para especificar rangos de casillas utilice el mismo método que emplearía en Excel.

Lectura de archivos de Excel 5 o de versiones posteriores en el Editor de datos

Siga las siguientes normas para leer archivos de Excel 5 o de versiones posteriores:

Tipo y ancho de datos. Cada columna es una variable. El tipo de datos y el ancho de cada variable está determinado por el tipo de datos y el ancho en el archivo de Excel. Si la columna contiene más de un tipo de datos (por ejemplo, fecha y numérico), el

tipo de datos se define como cadena y todos los valores se leen como valores de cadena válidos.

Casillas en blanco. En las variables numéricas, las casillas en blanco se convierten en el valor perdido del sistema indicado por un punto (o una coma). En las variables de cadena, los espacios en blanco son valores de cadena válidos y las casillas en blanco se tratan como valores de cadena válidos.

Nombres de variable. Si lee la primera fila del archivo de Excel (o la primera fila del rango especificado) como nombres de variable, los valores que no cumplan las normas de denominación de variables se convertirán en nombres de variables válidos y los nombres originales se utilizarán como etiquetas de variable. Si no lee nombres de variable del archivo de Excel, se asignarán nombres de variable por defecto.

Lectura de archivos de Excel de versiones anteriores y otras hojas de cálculo en el Editor de datos Hojas de cálculo

Las siguientes normas se aplican para leer archivos de Excel de versiones anteriores a la versión 5 y otros datos de hoja de cálculo:

Tipo y ancho de datos. El tipo y el ancho de los datos para cada variable se determinan según el ancho de la columna y el tipo de datos de la primera casilla de la columna. Los valores de otro tipo se convierten en valor perdido por el sistema. Si la primera casilla de datos de la columna está en blanco, se utiliza el tipo de datos global por defecto para la hoja de cálculo (normalmente numérico).

Casillas en blanco. En las variables numéricas, las casillas en blanco se convierten en el valor perdido del sistema indicado por un punto (o una coma). En las variables de cadena, los espacios en blanco son valores de cadena válidos y las casillas en blanco se tratan como valores de cadena válidos.

Nombres de variable. Si no se leen los nombres de variable de la hoja de cálculo, se utilizan las letras de las columnas (*A, B, C, ...*) como nombres de variable de los archivos de Excel y de Lotus. Para los archivos de SYLK y de Excel guardados en el formato de presentación R1C1, el programa utiliza para los nombres de variable el número de la columna precedido por la letra *C* (*C1, C2, C3, ...*).

Lectura de archivos de dBASE en el Editor de datos

Los archivos de bases de datos son, lógicamente, muy similares a los archivos de datos con formato SPSS. Las siguientes normas generales se aplican a los archivos de dBASE:

- Los nombres de campo se convierten en nombres de variable válidos.
- Los dos puntos en los nombres de campo de dBASE se convierten en subrayado.
- Se incluyen los registros marcados para ser eliminados que aún no se han purgado. El programa crea una nueva variable de cadena, *D_R*, que incluye un asterisco en los casos marcados para su eliminación.

Lectura de archivos de bases de datos

Podrá leer los datos desde cualquier formato de base de datos para los que disponga de un controlador de base de datos. En el análisis en modo local, los controladores necesarios deben estar instalados en el ordenador local. En el análisis en modo distribuido (disponible con la versión de servidor), los controladores deben estar instalados en el servidor remoto. Si desea obtener más información, consulte “Análisis en modo distribuido” en Capítulo 4 en p. 71.

Para leer archivos de base de datos

- ▶ Elija en los menús:
 - Archivo
 - Abrir base de datos
 - Nueva consulta...
- ▶ Seleccione la fuente de datos.
- ▶ En función de la fuente de datos, puede ser necesario seleccionar el archivo de base de datos y/o introducir un nombre de acceso, contraseña y demás información.
- ▶ Seleccione las tablas y los campos.
- ▶ Especifique cualquier relación existente entre las tablas.

Si lo desea, tiene la posibilidad de:

- Especificar cualquier criterio de selección para los datos.
- Añadir un mensaje solicitando al usuario que introduzca datos para crear una consulta con parámetros.
- Guardar la consulta creada antes de ejecutarla.

Para editar una consulta de base de datos guardada

- ▶ Elija en los menús:
 - Archivo
 - Abrir base de datos
 - Editar consulta...
- ▶ Seleccione el archivo de consulta (*.spq) que desee editar.
- ▶ Siga las instrucciones para crear una consulta.

Para leer archivos de bases de datos con una consulta ODBC guardada

- ▶ Elija en los menús:
 - Archivo
 - Abrir base de datos
 - Ejecutar consulta...
- ▶ Seleccione el archivo de consulta (*.spq) que desee ejecutar.
- ▶ En algunos archivos de base de datos tendrá que introducir un nombre de acceso y una contraseña.
- ▶ Si la consulta tiene una solicitud incrustada, es posible que deba introducir datos adicionales (por ejemplo, el trimestre para el que desee obtener cifras de ventas).

Selección de una fuente de datos

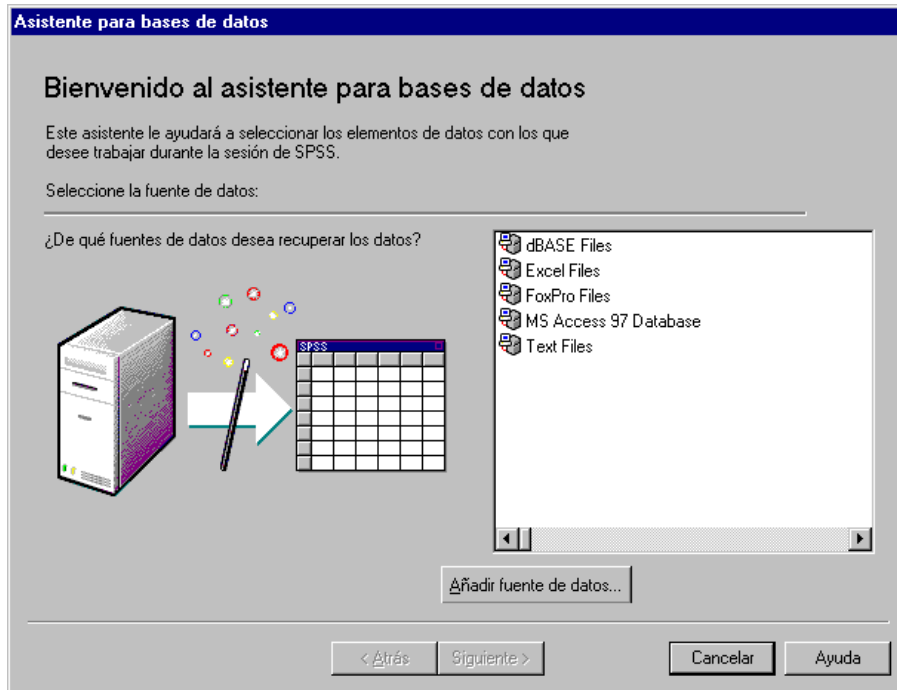
Utilice la primera pantalla para seleccionar el tipo de fuente de datos que leer. Una vez seleccionado el tipo de archivo, el Asistente para bases de datos solicitará la ruta de acceso del archivo de datos.

Si no tiene configurada ninguna fuente de datos o si desea añadir una nueva, pulse en Añadir fuente de datos. En el análisis en modo distribuido (disponible con la versión de servidor), este botón no está disponible. Para añadir fuentes de datos en el análisis en modo distribuido, consulte con el administrador del sistema.

Fuentes de datos. Una fuente de datos está compuesta por dos partes esenciales de información: el controlador que se utilizará para acceder a los datos y la ubicación de la base de datos a la que se desea acceder. Para especificar las fuentes de datos deberán estar instalados los controladores adecuados. Para el análisis en modo local se pueden instalar los controladores desde el CD-ROM de este producto.

- **Data Access Pack de SPSS.** Instala los controladores de una variedad de formatos de bases de datos. Está disponible en el menú de reproducción automática.
- **Data Access Pack de Microsoft.** Instala los controladores para productos Microsoft, incluido Microsoft Access. Para instalar el Data Access Pack de Microsoft, pulse dos veces sobre Microsoft Data Access Pack en la carpeta Microsoft Data Access Pack del CD-ROM.

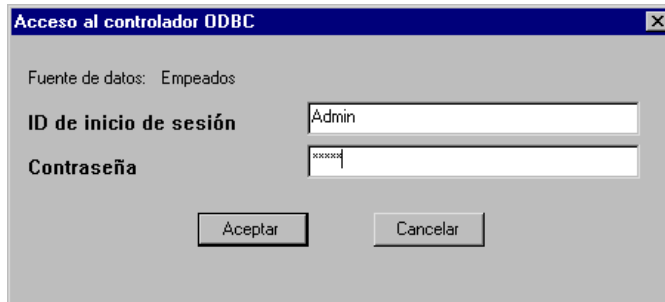
Figura 3-1
Cuadro de diálogo Asistente para bases de datos



Acceso a la base de datos

Si la base de datos precisa una contraseña, el Asistente para bases de datos la solicitará para abrir la fuente de datos.

Figura 3-2
Cuadro de diálogo de acceso



The image shows a Windows-style dialog box titled "Acceso al controlador ODBC". The dialog box has a blue title bar with a close button (X) in the top right corner. The main area is light gray and contains the following elements:

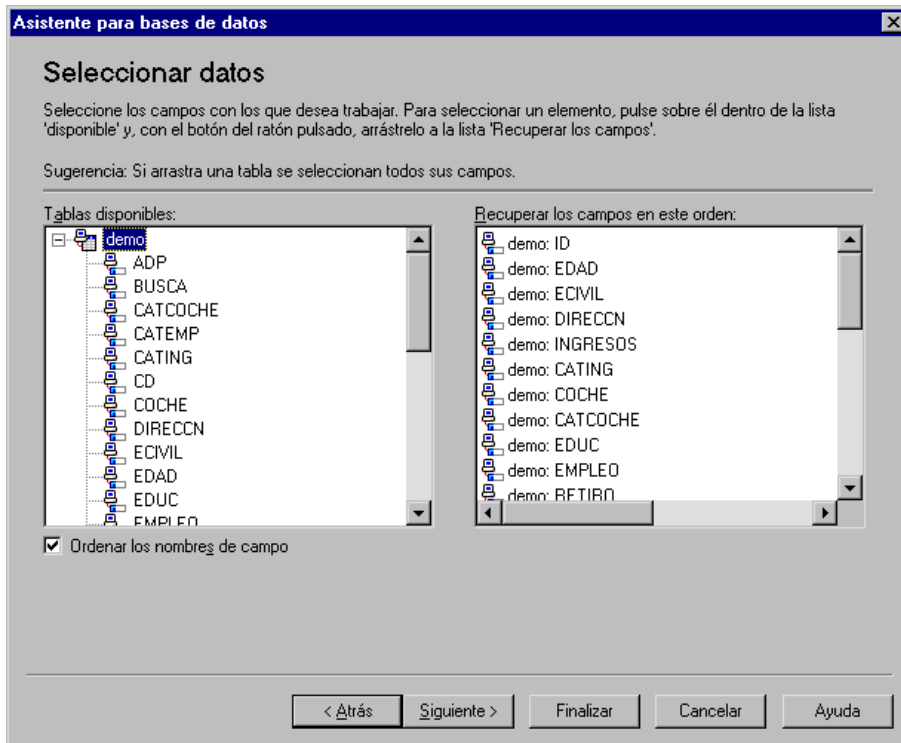
- Text: "Fuente de datos: Empeados"
- Label: "ID de inicio de sesión" followed by a text input field containing the text "Admin".
- Label: "Contraseña" followed by a text input field containing the text "xxxxxx".
- Two buttons at the bottom: "Aceptar" (Accept) and "Cancelar" (Cancel).

Selección de campos de datos

El paso de selección de datos controla las tablas y los campos que se deben leer. Los campos (las columnas) de la base de datos se leen como variables.

Si una tabla tiene un campo cualquiera seleccionado, todos sus campos serán visibles en las ventanas subsiguientes del Asistente para bases de datos, sin embargo, sólo se importarán como variables los campos seleccionados en este cuadro de diálogo. Esto le permitirá crear uniones entre tablas y especificar criterios empleando los campos que no esté importando.

Figura 3-3
Cuadro de diálogo Seleccionar datos



Presentación de los nombres de los campos. Para ver los campos de la tabla, pulse en el signo más (+) situado a la izquierda del nombre de una tabla. Para ocultar los campos pulse en el signo menos (-) situado a la izquierda del nombre de una tabla.

Para añadir un campo. Pulse dos veces en cualquier campo de la lista Tablas disponibles o arrástrelo hasta la lista Recuperar los campos en este orden. Los campos se pueden volver a ordenar arrastrándolos y colocándolos dentro de la propia lista de campos seleccionados.

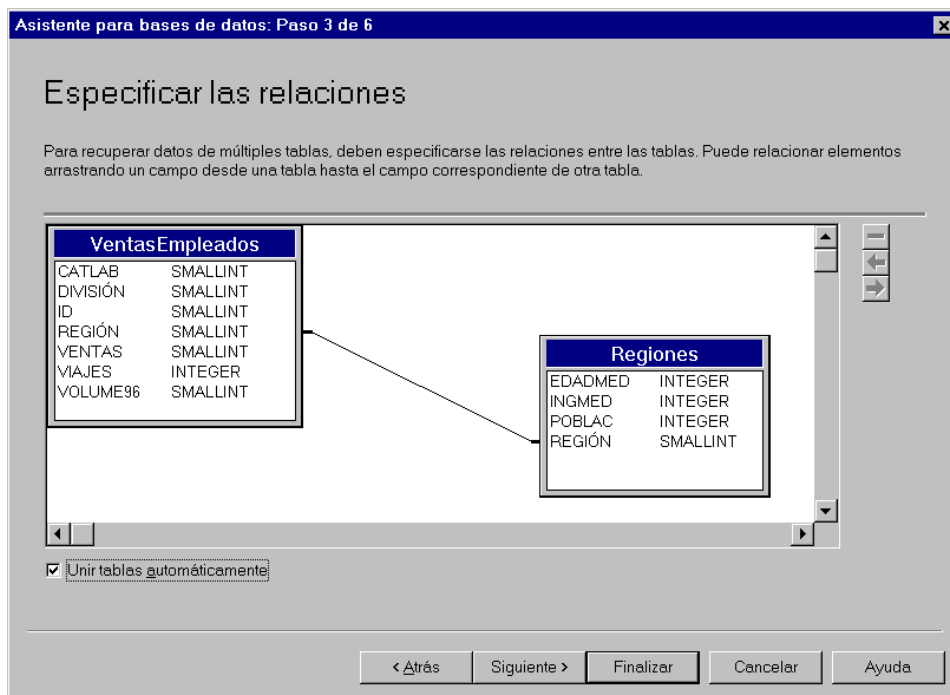
Para eliminar un campo. Pulse dos veces en cualquier campo de la lista Recuperar los campos en este orden, o bien arrástrelo hasta la lista Tablas disponibles.

Ordenar los nombres de campo. Si se selecciona, el Asistente para bases de datos mostrará los campos disponibles en orden alfabético.

Creación de una relación entre tablas

El cuadro de diálogo Especificar las relaciones permite definir relaciones entre las tablas. Si selecciona campos de más de una tabla, deberá definir al menos una unión.

Figura 3-4
Especificar cuadro de diálogo Relaciones



Establecimiento de relaciones. Para crear una relación, arrastre un campo desde cualquier tabla hasta el campo con el que quiera unirlo. El Asistente para bases de datos dibujará una **línea de unión** entre los dos campos que indica su relación. Estos campos deben ser del mismo tipo de datos.

Unir tablas automáticamente. Intenta unir las tablas automáticamente en función de las claves primarias/externas o de los nombres de campo y tipos de datos coincidentes.

Especificación de tipos de uniones. Si el controlador permite uniones exteriores, podrá especificar uniones interiores, uniones exteriores izquierdas o uniones exteriores derechas. Para seleccionar el tipo de unión, pulse dos veces en la línea de unión entre los campos y el asistente mostrará el cuadro de diálogo Propiedades de la relación.

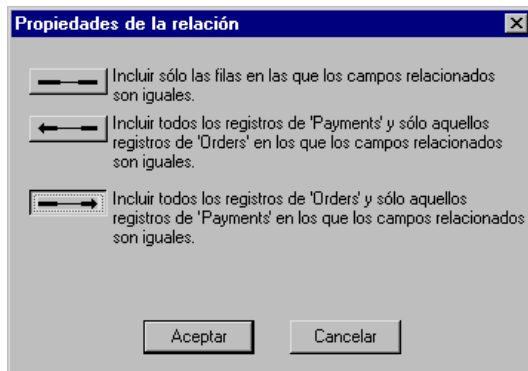
También puede servirse de los iconos situados en la esquina superior derecha del cuadro de diálogo para elegir el tipo de unión.

Propiedades de la relación

Este cuadro de diálogo permite especificar el tipo de relación que une las tablas.

Figura 3-5

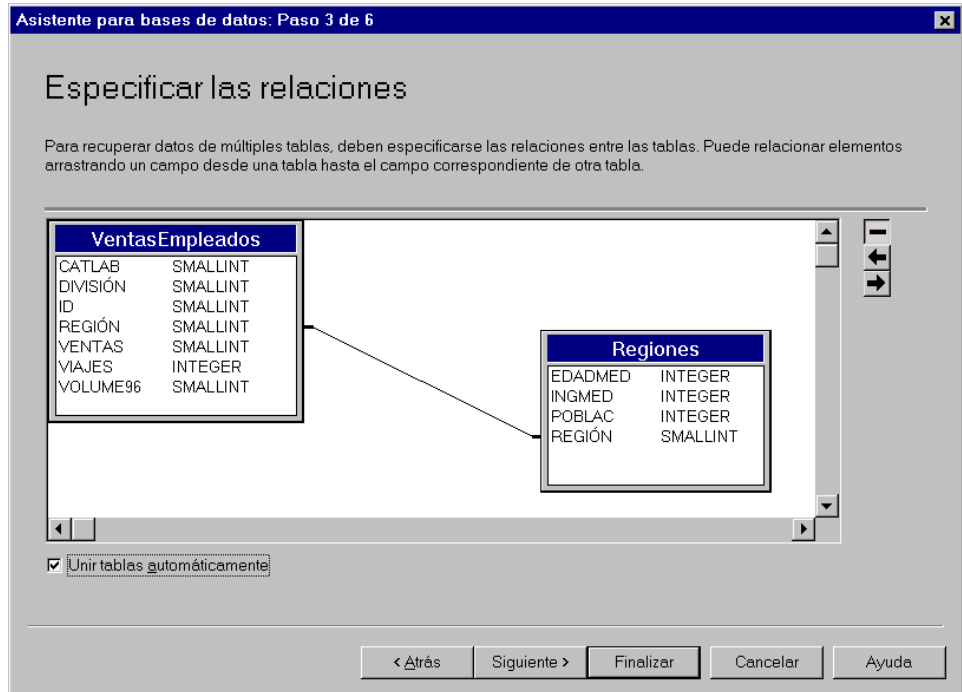
Cuadro de diálogo Propiedades de la relación



Uniones interiores. Una unión interior incluye sólo las filas donde los campos relacionados son iguales.

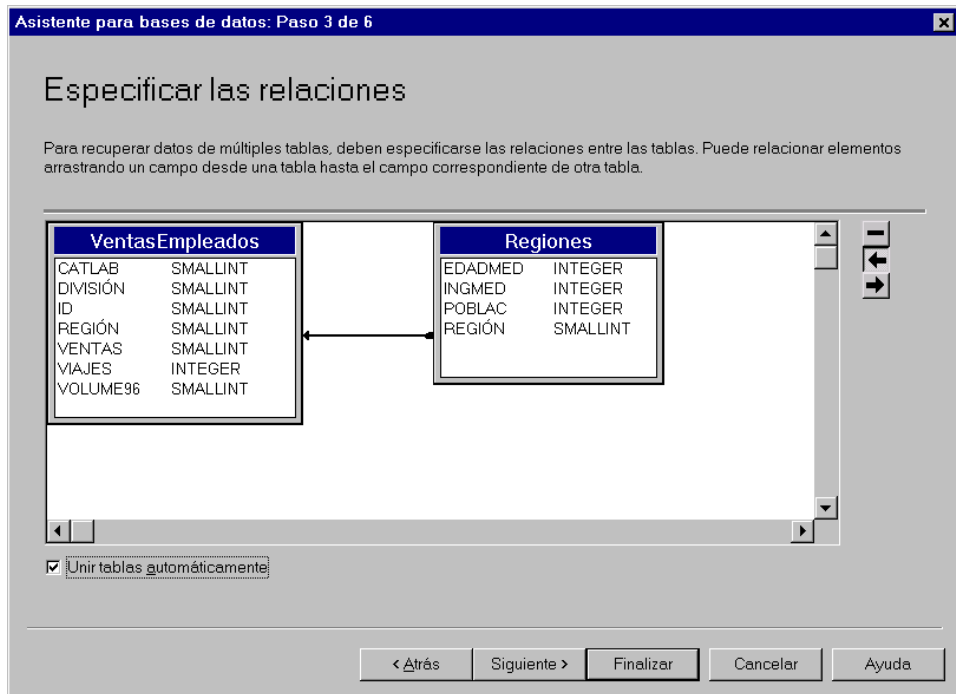
Al finalizar obtendrá un conjunto de datos que contiene las variables *ID*, *REGIÓN*, *VENTAS95* e *INGMEDIOS* por cada empleado que trabajó en una región determinada.

Figura 3-6
Creación de una unión interior



Uniones exteriores. Una unión exterior izquierda incluye todos los registros de la tabla izquierda y sólo aquellos registros de la tabla derecha en los que los campos relacionados son iguales. En una unión exterior derecha, esta relación se invierte, de forma que el programa importa todos los registros de la tabla derecha y sólo aquellos registros de la tabla izquierda en los que los campos relacionados son iguales.

Figura 3-7
Creación de una unión exterior derecha



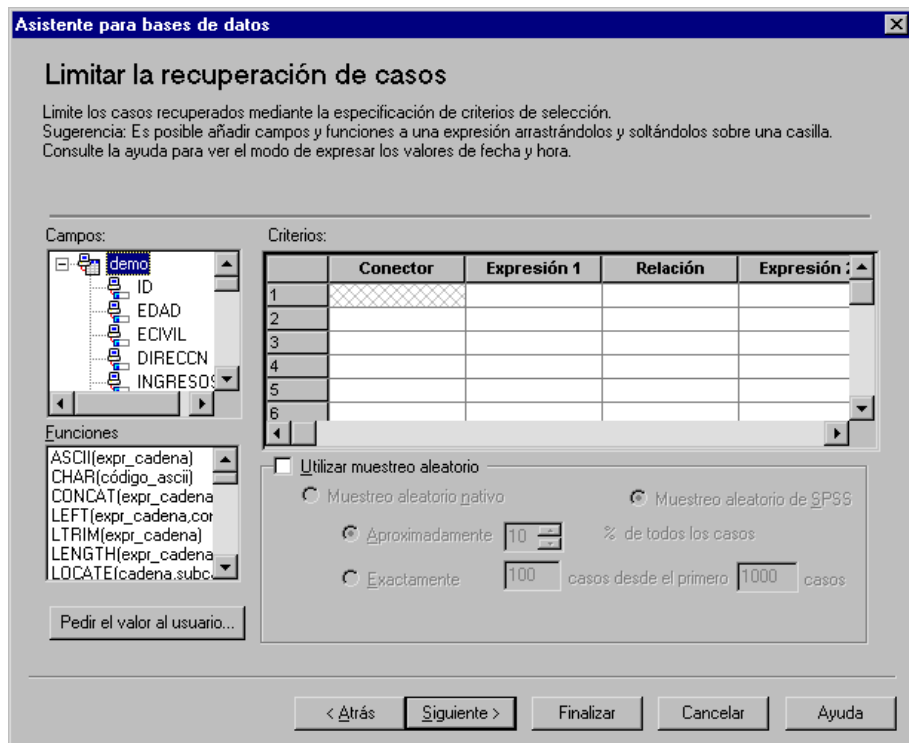
Limitar la recuperación de casos

Este cuadro de diálogo permite especificar el criterio para seleccionar subconjuntos de casos (filas). La limitación de los casos consiste generalmente en rellenar la rejilla de criterios con uno o varios criterios. Los criterios constan de dos expresiones y de alguna relación entre ellas, y devuelven un valor *verdadero*, *falso* o *perdido* para cada caso.

- Si el resultado es *verdadero*, se selecciona el caso.
- Si el resultado es *falso* o *perdido*, no se selecciona el caso.

- La mayoría de los criterios utiliza al menos uno de los seis operadores de relación (<, >, <=, >=, = y <>).
- Las expresiones pueden incluir nombres de campo, constantes, operadores aritméticos, funciones numéricas y de otros tipos, y variables lógicas. Puede utilizar como variables los campos que no vaya a importar.

Figura 3-8
Cuadro de diálogo Limitar la recuperación de casos



Para crear sus criterios necesita por lo menos dos expresiones y una relación para conectarlas.

- Para crear una expresión, coloque el cursor en una casilla de expresión. Puede escribir nombres de campo, constantes, operadores aritméticos, funciones numéricas y de otros tipos, y variables lógicas. Existen otros métodos para situar un campo en una de las casillas de los criterios, como pulsar dos veces en el campo en la lista de campos,

arrastrar el campo desde esta lista o seleccionar un campo en el menú desplegable que se encuentra disponible en cualquier casilla de expresión activa.

- ▶ Las dos expresiones suelen estar conectadas mediante un operador relacional, como pueden ser = o >. Para elegir la relación, sitúe el cursor en la casilla de relación y escriba el operador o selecciónelo en el menú desplegable.

Las fechas y horas de las expresiones se deben especificar de un modo especial (incluidas las llaves que se muestran en los ejemplos):

- Los literales de fecha se deben especificar usando el formato general: {d aaaa-mm-dd}.
- Los literales de hora se deben especificar usando el formato general: {t hh:mm:ss}.
- Los literales de fecha/hora (marcas de hora) se deben especificar usando el formato general: {dt aaaa-mm-dd hh:mm:ss}.

Funciones. SPSS ofrece una selección de funciones preincorporadas SQL aritméticas, lógicas, de cadena, de fecha y de hora. Puede seleccionar una función de la lista y arrastrarla hasta la expresión, o introducir una función SQL válida. Consulte la documentación de la base de datos para obtener funciones SQL válidas. Hay una lista de funciones estándar disponibles en:

http://msdn.microsoft.com/library/en-us/odbc/hm/odbcscalar_functions.asp

Utilizar muestreo aleatorio. Selecciona una muestra aleatoria de casos de la fuente de datos. En el caso de grandes fuentes de datos, puede limitar el número de casos a una pequeña muestra representativa. De este modo se puede reducir considerablemente el tiempo de ejecución de los procedimientos. Si el muestreo aleatorio original se encuentra disponible para la fuente de datos, resulta más rápido que el muestreo aleatorio de SPSS dado que este último aún debe leer toda la fuente de datos para extraer una muestra aleatoria.

- **Aproximadamente (Seleccionar, Muestra aleatoria).** Genera una muestra aleatoria con el porcentaje aproximado de casos indicado.
- **Exactamente (Seleccionar, Muestra aleatoria).** Selecciona una muestra aleatoria con el número de casos especificado a partir del número total de casos especificado. Si el número total de casos especificado excede el número total de casos presentes en el archivo de datos, la muestra contendrá un número menor de casos proporcional al número solicitado.

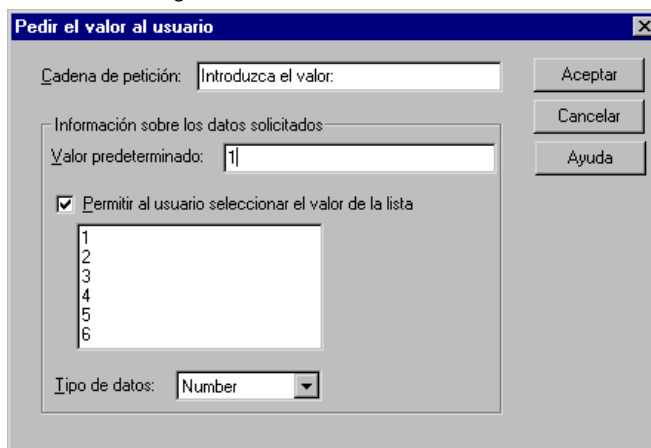
Pedir el valor al usuario. Permite insertar una solicitud en la consulta para crear una **consulta con parámetros**. Cuando un usuario ejecute la consulta se le solicitará que introduzca los datos aquí especificados. Puede interesarle esta opción si necesita obtener diferentes vistas de los mismos datos. Por ejemplo, es posible que desee ejecutar la misma consulta para ver las cifras de ventas de diversos trimestres fiscales. Sitúe el cursor en cualquier casilla de expresión y pulse en Pedir el valor al usuario para crear una petición.

Nota: Si utiliza el muestreo aleatorio, agregación (disponible en el modo distribuido con el servidor de SPSS) no estará disponible.

Creación de una consulta con parámetros

Utilice el cuadro de diálogo Pedir el valor al usuario para crear un cuadro de diálogo que solicite información al usuario cada vez que ejecute su consulta. Es especialmente útil para realizar consultas de una misma fuente de datos empleando criterios diferentes.

Figura 3-9
Cuadro de diálogo Pedir el valor al usuario



Para crear una solicitud es necesario introducir una cadena de petición y un valor por defecto. Esta cadena aparecerá cada vez que un usuario ejecute la consulta. La solicitud debe especificar el tipo de datos que debe introducir el usuario y, si

las opciones no se ofrecen en una lista, el formato de estos datos. Por ejemplo, “Introduzca un trimestre (Q1, Q2, Q3, ...)”.

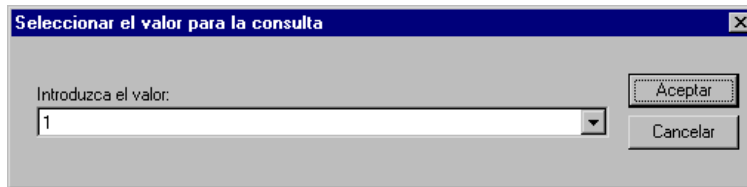
Permitir al usuario seleccionar el valor de la lista. Si selecciona esta opción, puede limitar las elecciones del usuario a los valores que incluya en esta lista, los cuales se separan con retornos.

Tipo de datos. Especifique aquí el tipo de datos entre las opciones Número, Cadena o Fecha.

El resultado final presenta este aspecto:

Figura 3-10

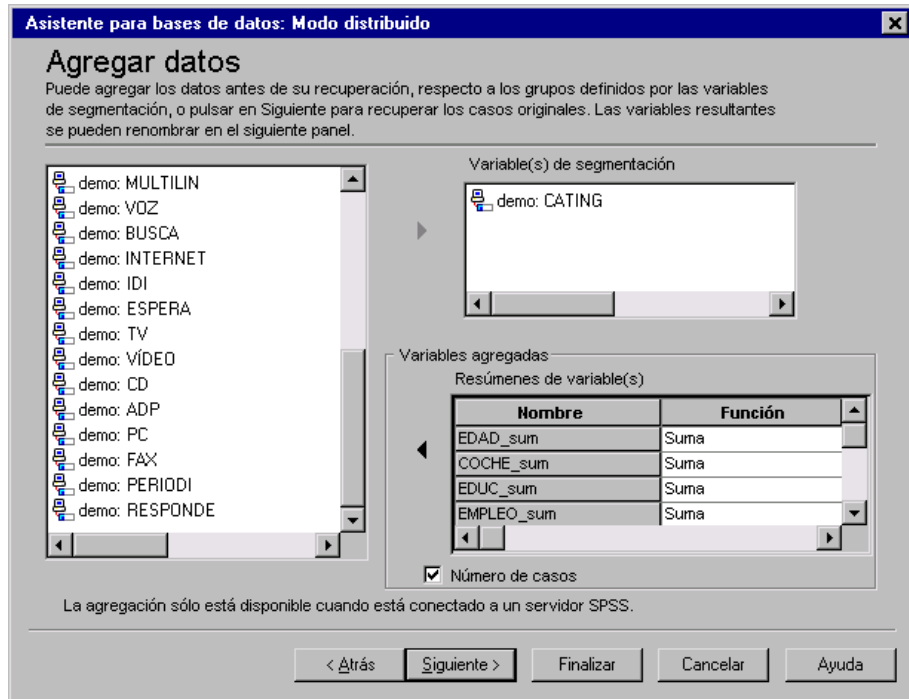
Cuadro de diálogo de petición de valor definido por el usuario



Adición de Datos

Si se encuentra en modo distribuido, conectado a un servidor remoto (disponible con el servidor SPSS), podrá agregar los datos antes de leerlos en SPSS.

Figura 3-11
Cuadro de diálogo Agregar datos



También se pueden agregar los datos después de leerlos en SPSS, pero si lo hace antes ahorrará tiempo en el caso de grandes fuentes de datos.

- ▶ Seleccione una o más variables de segmentación que definan cómo deben agruparse los casos para crear datos agregados.
- ▶ Seleccione una o varias variables para incluir.
- ▶ Seleccione una función de agregación para cada variable agregada.

Si lo desea, puede crear una variable que contenga el número de casos en cada grupo de segmentación.

Nota: Si utiliza el muestreo aleatorio, la agregación no estará disponible.

Definición de variables

Nombres y etiquetas de variables. El nombre completo del campo (columna) de la base de datos se utiliza como etiqueta de la variable. A menos que modifique el nombre de la variable, el Asistente para bases de datos asignará nombres de variable a cada columna de la base de datos de una de las siguientes formas:

- Si el nombre del campo de la base de datos forma un nombre de variable válido y único, se usará como el nombre de la variable.
- Si el nombre del campo de la base de datos no es un nombre de variable válido y único, se generará automáticamente un nombre único.

Pulse en cualquier casilla para editar el nombre de la variable.

Conversión de cadenas en variables numéricas. Active la casilla Recodificar como numérica para convertir automáticamente una variable de cadena en una variable numérica. Los valores de cadena se convierten en valores enteros consecutivos en función del orden alfabético de los valores originales. Los valores originales se mantienen como etiquetas de valor para las nuevas variables.

Anchura para los campos de ancho variable. Controla la anchura de los valores de las cadenas de anchura variable. Por defecto, la anchura es de 255 bytes y sólo se leen los primeros 255 bytes (generalmente 255 caracteres en idiomas de un solo byte). El valor máximo que se puede asignar a este parámetro es de 32.767 bytes. Aunque es probable que desee truncar los valores de cadena, tampoco querrá especificar un valor innecesariamente alto, ya que provocaría una merma en el rendimiento del procesamiento de SPSS.

Figura 3-12
Cuadro de diálogo Definir las variables

Asistente para bases de datos [X]

Definir las variables

Editar nombres y propiedades de variable de SPSS.
Los nombres se generarán de manera automática si los deja en blanco.
Para convertir una variable alfanumérica en numérica, utilizando los códigos originales como etiquetas de los valores, seleccione Recodificar como numérica.

	Nombre de variable de resultado	Tipo de datos	Recodificar como numérica
demo: DIRECCN	DIRECCN	Numérica	
demo: INGRESOS	INGRESOS	Numérica	
demo: CATING	CATING	Numérica	
demo: COCHE	COCHE	Numérica	
demo: CATCOCHE	CATCOCHE	Numérica	
demo: EDUC	EDUC	Numérica	
demo: EMPLEO	EMPLEO	Numérica	
demo: RETIRO	RETIRO	Numérica	
demo: CATEMP	CATEMP	Numérica	
demo: GÉNERO	GÉNERO	Cadena	<input checked="" type="checkbox"/>
demo: REDISENT	REDISENT	Numérica	

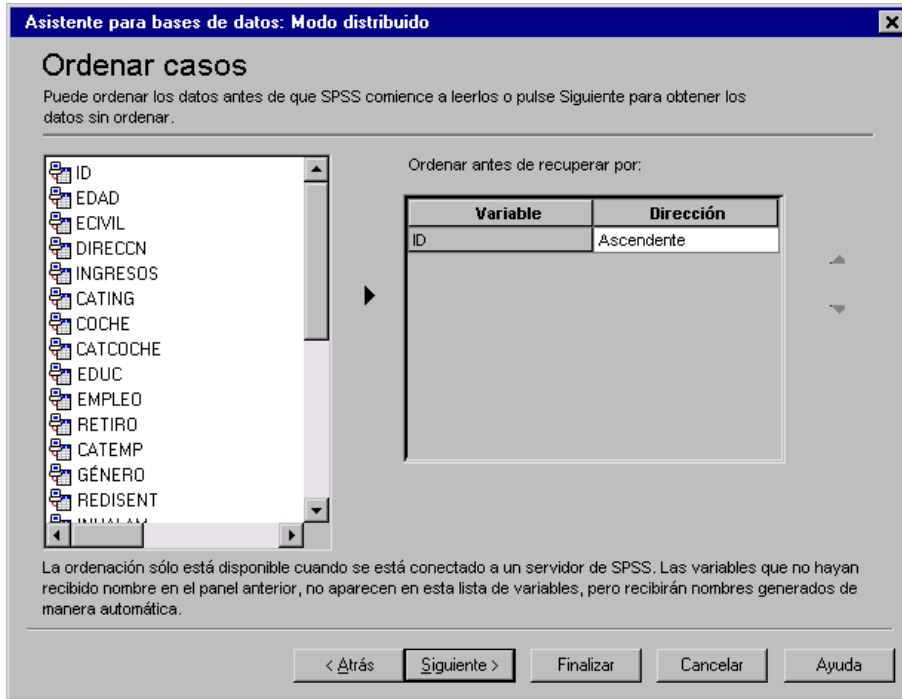
Anchura para los campos alfanuméricos de ancho variable:

< Atrás **Siguiente >** Finalizar Cancelar Ayuda

Ordenación de casos

Si se encuentra en modo distribuido, conectado a un servidor remoto (disponible con el servidor SPSS), podrá ordenar los datos antes de leerlos en SPSS.

Figura 3-13
Cuadro de diálogo Ordenar casos



También se pueden ordenar los datos después de leerlos en SPSS, pero si lo hace antes ahorrará tiempo en el caso de grandes fuentes de datos.

Resultados

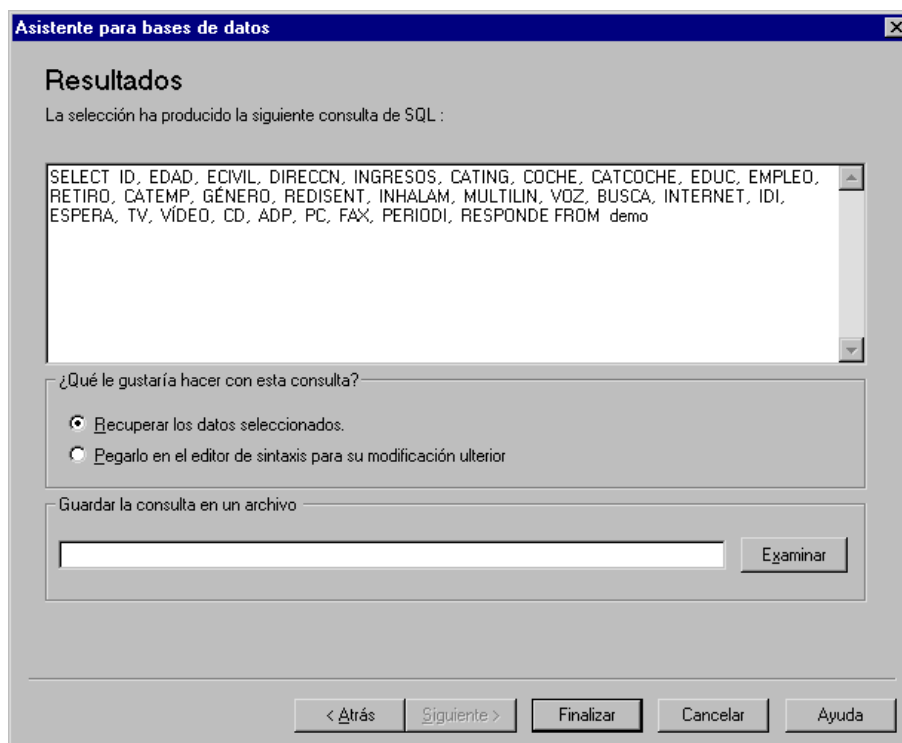
El cuadro de diálogo Resultados muestra la instrucción Select de SQL para la consulta.

- Se puede editar la instrucción Select de SQL antes de ejecutar la consulta, pero si pulsa el botón Anterior para introducir cambios en pasos anteriores, se perderán los cambios realizados en la instrucción Select.

- Se puede guardar la consulta para una futura utilización con Guardar la consulta en un archivo.
- Seleccione Pegarlo en el editor de sintaxis para su modificación ulterior para pegar la sintaxis de `GET DATA` completa en una ventana de sintaxis. Copiar y pegar la instrucción `Select` de la ventana Resultados no pegará la sintaxis de comandos necesaria.

Nota: La sintaxis pegada contiene un espacio en blanco delante de las comillas de cierre en cada línea de SQL generado por el asistente. Estos espacios no son superfluos. Cuando se procesa el comando, todas las líneas de la instrucción SQL se fusionan de un modo muy literal. Si esos espacios, los caracteres último y primero de cada línea se unirían.

Figura 3-14
Cuadro de diálogo Resultados



Asistente para la importación de texto

El Asistente para la importación de texto puede leer archivos de datos de texto de diversos formatos:

- Archivos delimitados por tabuladores
- Archivos delimitados por espacios
- Archivos delimitados por comas
- Archivos con formato de campos fijos

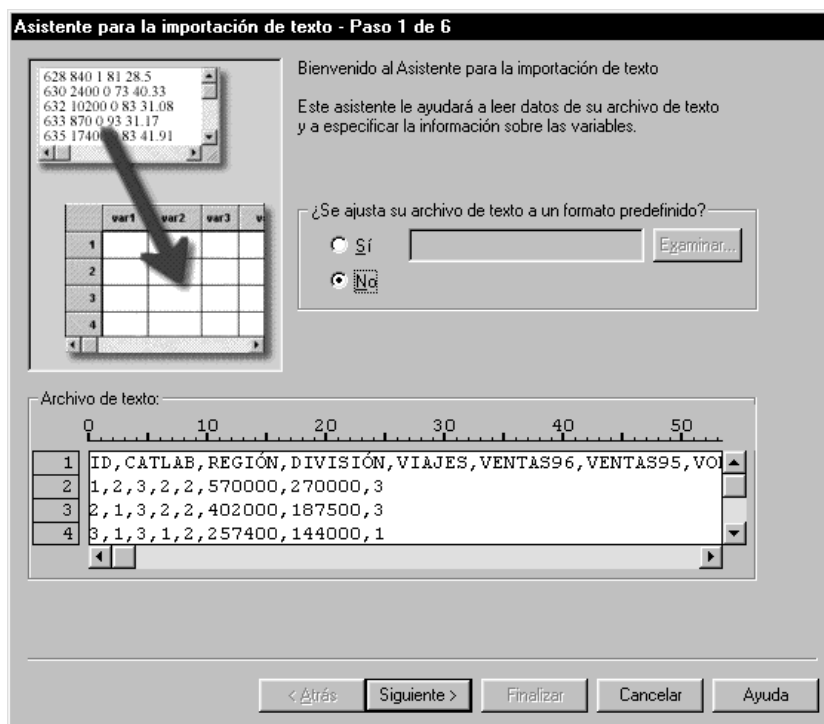
En los archivos delimitados, también se pueden especificar otros caracteres como delimitadores entre valores, o bien especificar varios delimitadores diferentes.

Para leer archivos de datos de texto

- ▶ Elija en los menús:
 - Archivo
 - Leer datos de texto
- ▶ Seleccione el archivo de texto en el cuadro de diálogo Abrir.
- ▶ Siga los pasos indicados en el Asistente para la importación de texto para definir cómo desea leer el archivo de datos de texto.

Asistente para la importación de texto: Paso 1

Figura 3-15
Asistente para la importación de texto: Paso 1



El archivo de texto se mostrará en una ventana de vista previa. Puede aplicar un formato predefinido (guardado con anterioridad desde el Asistente para la importación de texto) o seguir los pasos del asistente para especificar cómo desea que se lean los datos.

Asistente para la importación de texto: Paso 2

Figura 3-16
Asistente para la importación de texto: Paso 2

Asistente para la importación de texto - Paso 2 de 6

¿Cómo están organizadas sus variables?

Delimitadas - Las variables están delimitadas por un carácter concreto (coma, tabulador).

Ancho fijo - Las variables están alineadas en columnas de ancho fijo.

¿Están incluidos los nombres de las variables en la parte superior del archivo?

Sí

No

Archivo de texto:

	0	10	20	30	40	50
1	ID, CATLAB, REGIÓN, DIVISIÓN, VIAJES, VENTAS96, VENTAS95, VO					
2	1, 2, 3, 2, 2, 570000, 270000, 3					
3	2, 1, 3, 2, 2, 402000, 187500, 3					
4	3, 1, 3, 1, 2, 257400, 144000, 1					

< Atrás **Siguiente >** Finalizar Cancelar Ayuda

Este paso ofrece información sobre las variables. Una de las variables es similar a uno de los campos de la base de datos. Por ejemplo, cada elemento de un cuestionario es una variable.

¿Cómo están organizadas sus variables? Para leer los datos adecuadamente, el Asistente para la importación de texto necesita saber cómo determinar el lugar en el que terminan los valores de datos de una variable y comienzan los valores de datos de la variable siguiente. La organización de las variables define el método utilizado para diferenciar una variable de la siguiente.

- **Delimitado.** Se utilizan espacios, comas, tabulaciones u otros caracteres para separar variables. Las variables quedan registradas en el mismo orden para cada caso, pero no necesariamente conservando la misma ubicación para las columnas.
- **Ancho fijo.** Cada variable se registra en la misma posición de columna en el mismo registro (línea) para cada caso del archivo de datos. No se requiere delimitador entre variables. De hecho, en muchos archivos de datos de texto generados por programas de ordenador, podría parecer que los valores de los datos se suceden, sin espacios que los separen. La ubicación de la columna determina qué variable se está leyendo.

¿Están incluidos los nombres de las variables en la parte superior del archivo? Si la primera fila del archivo de datos contiene etiquetas descriptivas para cada variable, podrá utilizar dichas etiquetas como nombres de las variables. Los valores que no cumplan las normas de denominación de variables se convertirán en nombres de variables válidos.

Asistente para la importación de texto: Paso 3: Archivos delimitados

Figura 3-17

Asistente para la importación de texto: Paso 3 para archivos delimitados

Asistente para la importación de texto - Paso 3 de 6

¿En qué número de línea comienza el primer caso de los datos?

¿Cómo se encuentran representados sus casos?

Cada línea representa un caso

Un número concreto de variables representa un caso:

¿Cuántos casos desea importar?

Todos los casos

Los primeros casos.

Un porcentaje aleatorio de los casos (aproximado): %

Vista previa de datos

	0	10	20	30	40	50
1	ID, CATLAB, REGIÓN, DIVISIÓN, VIAJES, VENTAS96, VENTAS95,					
2	1, 2, 3, 2, 2, 570000, 270000, 3					
3	2, 1, 3, 2, 2, 402000, 187500, 3					

< Atrás Siguiente > Finalizar Cancelar Ayuda

Este paso ofrece información sobre los casos. Un caso es similar a un registro de una base de datos. Por ejemplo, cada persona que responde a un cuestionario es un caso.

¿En qué número de línea comienza el primer caso de datos? Indica la primera línea del archivo de datos que contiene valores de datos. Si la línea o líneas superiores del archivo de datos contienen etiquetas descriptivas o cualquier otro texto que no represente valores de datos, dicha línea o líneas *no* serán la línea 1.

¿Cómo se representan sus casos? Controla la manera en que el Asistente para la importación de texto determina dónde finaliza cada caso y comienza el siguiente.

- **Cada línea representa un caso.** Cada línea contiene un sólo caso. Es bastante común que cada línea (fila) contenga un sólo caso, aunque dicha línea puede ser muy larga para un archivo de datos con un gran número de variables. Si no todas

las líneas contienen el mismo número de valores de datos, el número de variables para cada caso quedará determinado por la línea que tenga el mayor número de valores de datos. A los casos con menos valores de datos se les asignarán valores perdidos para las variables adicionales.

- **Un número concreto de variables representa un caso.** El número de variables especificado para cada caso informa al Asistente para la importación de texto de dónde detener la lectura de un caso y comenzar la del siguiente. Una misma línea puede contener varios casos y los casos pueden empezar en medio de una línea y continuar en la línea siguiente. El Asistente para la importación de texto determina el final de cada caso basándose en el número de valores leídos, independientemente del número de líneas. Cada caso debe contener valores de datos (o valores perdidos indicados por delimitadores) para todas las variables; de otra forma, el archivo de datos no se leerá correctamente.

¿Cuántos casos desea importar? Puede importar todos los casos del archivo de datos, los primeros n casos (siendo n un número especificado por el usuario) o una muestra aleatoria a partir de un porcentaje especificado. Dado que esta rutina de muestreo aleatorio toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar al porcentaje especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.

Asistente para la importación de texto: Paso 3: Archivos de ancho fijo

Figura 3-18

Asistente para la importación de texto: Paso 3 para archivos de ancho fijo

Asistente para la importación de texto: Ancho fijo - Paso 3 de 6

¿En qué número de línea comienza el primer caso de los datos?

¿Cuántas líneas representan un caso?

¿Cuántos casos desea importar?

Todos los casos

Los primeros casos.

Un porcentaje de los casos: %

Vista previa de datos

	0	10	20	30	40	50
1	12	322	\$570,000	\$270,000	3	
2	21	322	\$402,000	\$187,500	3	
3	31	312	\$257,400	\$144,000	1	
4	41	312	\$262,800	\$158,400	1	
5	51	322	\$450,000	\$210,000	3	

< Atrás **Siguiente >** Finalizar Cancelar Ayuda

Este paso ofrece información sobre los casos. Un caso es similar a un registro de una base de datos. Por ejemplo, cada encuestado es un caso.

¿En qué número de línea comienza el primer caso de datos? Indica la primera línea del archivo de datos que contiene valores de datos. Si la línea o líneas superiores del archivo de datos contienen etiquetas descriptivas o cualquier otro texto que no represente valores de datos, dicha línea o líneas *no* serán la línea 1.

¿Cuántas líneas representan un caso? Controla la manera en que el Asistente para la importación de texto determina dónde finaliza cada caso y comienza el siguiente. Cada variable queda definida por su número de línea dentro del caso y por la ubicación de su columna. Para leer los datos correctamente, deberá especificar el número de líneas de cada caso.

¿Cuántos casos desea importar? Puede importar todos los casos del archivo de datos, los primeros n casos (siendo n un número especificado por el usuario) o una muestra aleatoria a partir de un porcentaje especificado. Dado que esta rutina de muestreo aleatorio toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar al porcentaje especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.

Asistente para la importación de texto: Paso 4: Archivos delimitados

Figura 3-19

Asistente para la importación de texto: Paso 4 para archivos delimitados

Asistente para la importación de texto - Paso 4 de 6

¿Qué delimitador se encuentra entre las variables?

Tabulador Espacio
 Coma Punto y coma
 Otro:

Vista previa de datos

ID	CATLAB	REGIÓN	DIVISIÓN	VIAJES	VENTAS96	VENTA
1	2	3	2	2	570000	27000
2	1	3	2	2	402000	18750
3	1	3	1	2	257400	14400
4	1	3	1	2	262800	15840

< Atrás Siguiente > Finalizar Cancelar Ayuda

Este paso muestra la mejor opción, según el Asistente para la importación de texto, para leer el archivo de datos y le permite modificar la manera en que el asistente leerá las variables del archivo de datos.

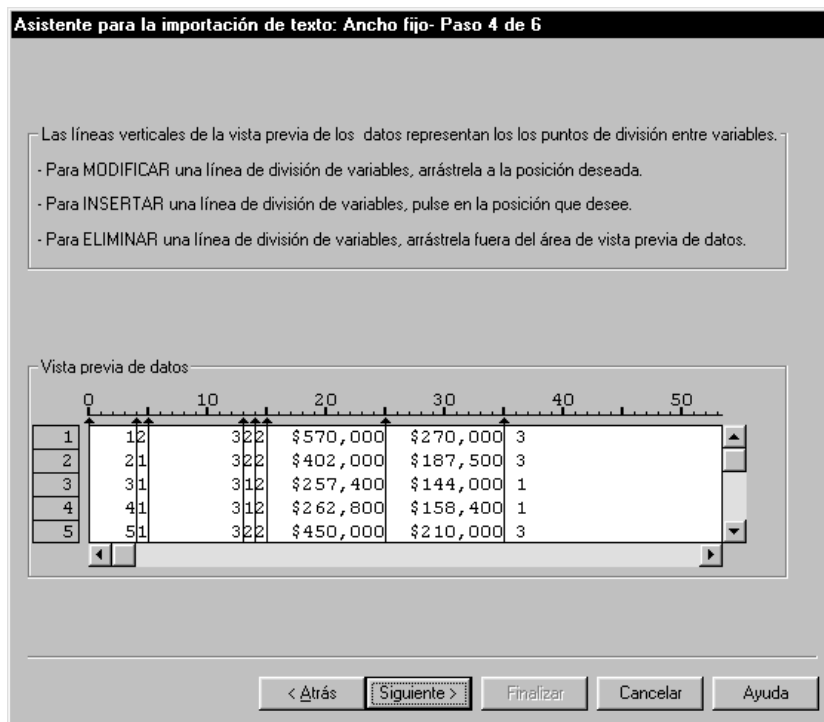
¿Qué delimitador desea para la separación entre variables? Indica los caracteres o símbolos que separan los valores de datos. Puede seleccionar cualquier combinación de espacios, comas, signos de punto y coma, tabulaciones o cualquier otro carácter. En caso de existir varios delimitadores consecutivos sin valores de datos, dichos delimitadores serán considerados valores perdidos.

¿Cuál es el calificador de texto? Caracteres utilizados para encerrar valores que contienen caracteres delimitadores. Por ejemplo, si una coma es el delimitador, los valores que contengan comas se leerán incorrectamente a menos que estos valores se encierre en un calificador de texto, impidiendo que las comas del valor se interpreten como delimitadores entre los valores. Los archivos de datos con formato CSV de Excel utilizan las comillas dobles (“”) como calificador de texto. El calificador de texto aparece tanto al comienzo como al final del valor, encerrándolo completamente.

Asistente para la importación de texto: Paso 4: Archivos de ancho fijo

Figura 3-20

Asistente para la importación de texto: Paso 4 para archivos de ancho fijo



Este paso muestra la mejor opción, según el Asistente para la importación de texto, para leer el archivo de datos y le permite modificar la manera en que el asistente leerá las variables del archivo de datos. Las líneas verticales de la ventana de vista previa indican el lugar en el que en ese momento el Asistente para la importación de texto piensa que cada variable comienza en el archivo.

Inserte, mueva y elimine líneas de ruptura de variable según convenga para separar variables. Si para cada caso se utilizan varias líneas, seleccione cada línea de la lista desplegable y modifique las líneas de ruptura de variable según convenga.

Nota: En archivos de datos generados por ordenador que producen un flujo continuo de valores de datos sin espacios ni otras características distintivas, puede resultar difícil determinar el lugar en el que comienza cada variable. Los archivos de datos

del tipo citado anteriormente suelen depender de un archivo de definición de datos u otro tipo de descripción escrita que especifique la ubicación por líneas y columnas de cada variable.

Asistente para la importación de texto: Paso 5

Figura 3-21

Asistente para la importación de texto: Paso 5

Asistente para la importación de texto - Paso 5 de 6

Especificaciones para la(s) variable(s) seleccionada(s) en la vista previa de datos:

Nombre de variable: Nombre original: ID

Formato de datos:

Vista previa de datos:

CATLAB	REGIÓN	DIVISIÓN	VIAJES	VENTAS96	VENTAS95	VOI ▲
2	3	2	2	570000	270000	3
1	3	2	2	402000	187500	3
1	3	1	2	257400	144000	1
1	3	1	2	262800	158400	1

< Atrás **Siguiente >** Finalizar Cancelar Ayuda

Este paso controla el nombre de la variable y el formato de datos que el Asistente para la importación de texto utilizará para leer cada variable, así como las que se incluirán en el archivo de datos definitivo.

Nombre de variable. Puede sobrescribir los nombres de variable por defecto y sustituirlos por otros diferentes. Si lee nombres de variable desde el archivo de datos, el Asistente para la importación de texto modificará de manera automática

los nombres de variable que no cumplan las normas de denominación de variables. Seleccione una variable en la ventana de vista previa e introduzca un nombre de variable.

Formato de datos. Seleccione una variable en la ventana de vista previa y, a continuación, seleccione un formato de la lista desplegable. Pulse el botón del ratón con la tecla MAYÚS presionada para seleccionar una serie de variables consecutivas, o bien con la tecla CTRL presionada para seleccionar una serie de variables no consecutivas.

Opciones para el formato del Asistente para la importación de texto

Entre las opciones de formato para la lectura de variables con el Asistente para la importación de texto se encuentran:

No importar. Omite la variable o variables seleccionadas del archivo de datos importado.

Numérico. Los valores válidos incluyen números, los signos más y menos iniciales y un indicador decimal.

Cadena. Son valores válidos prácticamente todos los caracteres del teclado y los espacios en blanco incrustados. En los archivos delimitados, puede especificar hasta un máximo de 32.767 de caracteres para el valor. El Asistente para la importación de texto fija como valor por defecto para el número de caracteres el valor de cadena más largo que se haya encontrado para la variable o variables seleccionadas. Para los archivos de ancho fijo, el número de caracteres en los valores de cadena queda definido por la ubicación de las líneas de ruptura de variable en el paso 4.

Fecha/hora. Entre los valores válidos se encuentran las fechas con formato general: *dd-mm-aaaa*, *mm/dd/aaaa*, *dd.mm.aaaa*, *aaaa/mm/dd*, *hh:mm:ss*, así como una amplia variedad de formatos de hora y fecha. Los meses se pueden representar en dígitos, números romanos, abreviaturas de tres letras o con el nombre completo. Seleccione un formato de fecha de la lista.

Dólar. Los valores válidos son números con un signo dólar inicial optativo y puntos separadores de millares también optativos.

Coma. Entre los valores válidos se encuentran los números que utilizan un punto para separar los decimales y una coma para separar los millares.

Punto. Entre los valores válidos se encuentran los números que utilizan una coma para separar los decimales y un punto para separar los millares.

Nota: Los valores que contengan caracteres no válidos para el formato seleccionado serán considerados valores perdidos. Los valores que contengan uno cualquiera de los delimitadores especificados serán considerados como valores múltiples.

Asistente para la importación de texto: Paso 6

Figura 3-22

Asistente para la importación de texto: Paso 6

Asistente para la importación de texto - Paso 6 de 6

Ha definido satisfactoriamente el formato del archivo de texto.

¿Desea guardar este formato de archivo para su uso en el futuro?

Sí No

¿Desea pegar la sintaxis?

Sí No

Pulse en el botón Finalizar para completar el Asistente para la importación de texto.

Vista previa de datos

ID	CATLAB	REGIÓN	DIVISIÓN	VIAJES	VENTAS96	VEI
1	2	3	2	2	570000	2700
2	1	3	2	2	402000	1875
3	1	3	1	2	257400	1440
4	1	3	1	2	262800	1584

< Atrás

Este es el paso final del Asistente para la importación de texto. Puede guardar sus propias especificaciones en un archivo para hacer uso de ellas cuando importe archivos de datos de texto similares. También puede pegar la sintaxis generada por el Asistente

para la importación de texto en una ventana de sintaxis. Así podrá personalizar y/o guardar dicha sintaxis para utilizarla en futuras sesiones o en trabajos de producción.

Caché local de los datos. Una caché de los datos es una copia completa del archivo de datos, almacenada en un espacio de disco temporal. La caché del archivo de datos puede mejorar el rendimiento.

Información sobre el archivo

Un archivo de datos contiene mucho más que datos sin formato. También contiene información sobre la definición de las variables, incluyendo:

- Los nombres de las variables.
- Los formatos de las variables.
- Las etiquetas descriptivas de variable y de valor.

Esta información se almacena en la parte del diccionario sobre el archivo de datos. El Editor de datos proporciona una forma de presentar la información sobre la definición de la variable. También se puede mostrar la información completa del diccionario para el archivo de datos de trabajo o para cualquier otro archivo de datos.

Obtención de información sobre un archivo de datos

- ▶ Seleccione en los menús de la ventana Editor de datos:
 - Archivo
 - Mostrar información del archivo de datos
 - ▶ Para el archivo de datos abierto actualmente, elija Archivo de trabajo.
 - ▶ Para otros archivos de datos, elija Archivo externo y seleccione el archivo de datos.
- La información sobre el archivo de datos se muestra en el Visor.

Almacenamiento de archivos de datos

Cualquier cambio que realice en un archivo de datos se mantendrá solamente durante la sesión actual, a menos que guarde los cambios de manera explícita.

Almacenamiento de archivos de datos modificados

- ▶ Active la ventana Editor de datos (pulse en cualquier punto de la ventana para activarla).
- ▶ Elija en los menús:
 - Archivo
 - Guardar

El archivo de datos modificado se guarda y sobrescribe la versión anterior del archivo.

Almacenamiento de archivos de datos en formato de Excel

Puede guardar los datos en uno de los tres formatos de archivo de Microsoft Excel. La opción de formato depende de la versión de Excel que se utilizar para abrir los datos. La aplicación Excel no puede abrir un archivo de Excel procedente de una versión más actual de la aplicación. Por ejemplo, Excel 5.0 no puede abrir un documento de Excel 2000. Sin embargo, Excel 2000 puede abrir fácilmente un documento de Excel 5.0.

Hay unas cuantas limitaciones acerca del formato de archivo de Excel que no existen en SPSS. Entre estas limitaciones se incluyen:

- La información acerca de las variables, como los valores perdidos y las etiquetas de variable, no se incluye en los archivos de Excel exportados.
- Al exportar a Excel 97 o versiones posteriores, se ofrece la posibilidad de incluir las etiquetas de valor en lugar de los valores.
- Al estar limitados los archivos de Excel a 256 columnas de datos, sólo se incluirán en el archivo exportado las primeras 256 variables.
- Los archivos de Excel 4.0 y Excel 5.0/95 están limitados a 16.384 registros o filas de datos. Los archivos de Excel 97–2000 permiten 65.536 registros. Si los datos superan estos límites, aparecer un mensaje de aviso y los datos se truncarán de acuerdo con el máximo tamaño permitido por Excel.

Tipos de variables

La siguiente tabla muestra la relación del tipo de las variables entre los datos originales de SPSS y los datos exportados a Excel.

Tipo de variable de SPSS	Formato de datos de Excel
Numérico	0.00; #,##0.00;...
Coma	0.00; #,##0.00;...
Dolar	\$#,##0_);...
Fecha	d-mmm-aaaa
Tiempo	hh:mm:ss
Cadena	General

Almacenamiento de archivos de datos en formato de SAS

Al guardar un archivo de SAS, se aplica un tratamiento especial a determinadas características de los datos. Entre estos casos se incluyen:

- Algunos caracteres que se permiten en los nombres de variables de SPSS no son válidos en SAS, como por ejemplo @, # y \$. Al exportar los datos, estos caracteres no válidos se reemplazan por un carácter de subrayado.
- Las etiquetas de variable de SPSS que contienen más de 40 caracteres se truncan al exportarlas a un archivo de SAS v6.
- Si existen, las etiquetas de variable de SPSS se asignan a etiquetas de variable de SAS. Si no hay ninguna etiqueta de variable en los datos de SPSS, el nombre de variable se asigna a la etiqueta de variable de SAS.
- SAS sólo permite que exista un valor perdido del sistema, mientras que SPSS permite que haya varios valores perdidos del sistema. Por tanto, todos los valores perdidos del sistema en SPSS se asignan a un único valor perdido del sistema en el archivo SAS.

Almacenamiento de etiquetas de valor

Existe la posibilidad de guardar los valores y las etiquetas de valor asociadas al archivo de datos en un archivo de sintaxis de SAS. Por ejemplo, al exportar las etiquetas de valor del archivo de datos *coches.sav*, el archivo de sintaxis generado contiene:

```
libname library 'd:\spss\' ;

proc format library = library ;
```

```
value ORIGIN /* País de origen */
  1 = 'EE.UU.'
  2 = 'Europa'
  3 = 'Japón' ;
value CYLINDER /* Número de cilindros */
  3 = '3 cilindros'
  4 = '4 cilindros'
  5 = '5 cilindros'
  6 = '6 cilindros'
  8 = '8 cilindros' ;
value FILTER__ /* cylrec = 1 | cylrec = 2 (FILTER) */
  0 = 'No seleccionado'
  1 = 'Seleccionado' ;

proc datasets library = library ;
modify cars;
  format    ORIGIN ORIGIN.;
  format    CYLINDER CYLINDER.;
  format    FILTER__ FILTER__.;
quit;
```

Esta función no se admite para el archivo de transporte de SAS.

Tipos de variables

La siguiente tabla muestra la relación del tipo de las variables entre los datos originales de SPSS y los datos exportados a SAS:

Tipo de variable de SPSS	Tipo de variable de SAS	Formato de datos de SAS
Numérico	Numérico	12
Coma	Numérico	12
Puntos	Numérico	12

Tipo de variable de SPSS	Tipo de variable de SAS	Formato de datos de SAS
Notación científica	Numérico	12
Fecha	Numérico	(Fecha) p.ej., MMDDAA10,...
Fecha (Hora)	Numérico	Hora18
Dolar	Numérico	12
Moneda personalizada	Numérico	12
Cadena	Carácter	\$8

Para guardar un archivo de datos nuevo o guardar los datos con un formato distinto

- ▶ Active la ventana Editor de datos (pulse en cualquier punto de la ventana para activarla).
- ▶ Elija en los menús:
 - Archivo
 - Guardar como...
- ▶ Seleccione un tipo de archivo de la lista desplegable.

- ▶ Introduzca un nombre de archivo para el nuevo archivo de datos.

Para escribir nombres de variable en la primera fila de una hoja de cálculo o de un archivo de datos delimitados por tabuladores:

- ▶ Pulse en Escribir nombres de variable en hoja de cálculo en el cuadro de diálogo Guardar datos como.

Para guardar las etiquetas de valor en lugar de los valores de los datos en formato de Excel 97:

- ▶ Pulse en Guardar etiquetas de valor donde se hayan definido en vez de valores de datos en el cuadro de diálogo Guardar datos como.

Para guardar etiquetas de valor en un archivo de sintaxis de SAS (esta opción sólo está activa si se ha seleccionado un tipo de archivo de SAS):

- ▶ Pulse en Guardar etiquetas de valor en un archivo .sas en el cuadro de diálogo Guardar datos como.

Almacenamiento de datos: Tipos de archivos de datos

Puede guardar datos en los siguientes formatos:

SPSS (*.sav). Formato SPSS.

- Los archivos de datos guardados con formato SPSS no se pueden leer en versiones anteriores a la 7.5.
- Al utilizar archivos de datos con nombres de variable con longitud superior a ocho bytes en SPSS 10.X u 11.X, se utilizan versiones únicas de ocho bytes de los nombres de variable, pero se mantienen los nombres originales de las variables para su utilización en la versión 12.0 o posterior. En versiones anteriores a SPSS 10, los nombres largos originales de las variables se pierden si se guarda el archivo de datos.
- Al utilizar archivos de datos con variables de cadena con más de 255 bytes en versiones de SPSS anteriores a la versión 13.0, dichas variables de cadena se fragmentan en variables de cadena de 255 bytes.

SPSS 7.0 (*.sav). Formato SPSS 7.0 para Windows. Los archivos de datos guardados con formato SPSS 7.0 se pueden leer en SPSS 7.0 y en versiones anteriores de SPSS para Windows, pero no incluyen los conjuntos de respuestas múltiples definidos ni la información sobre la introducción de datos para Windows.

SPSS/PC+ (*.sys). Formato SPSS/PC+. Si el archivo de datos contiene más de 500 variables, sólo se guardarán las 500 primeras. Para las variables con más de un valor perdido definido por el usuario, los valores perdidos definidos por el usuario adicionales se recodificarán en el primero de estos valores.

SPSS portátil (*.por). El formato portátil de SPSS puede leerse en otras versiones de SPSS y en versiones para otros sistemas operativos (por ejemplo, Macintosh y UNIX). Los nombres de variable se limitan a ocho bytes, y se convertirán a nombres únicos de ocho bytes si es preciso.

Delimitado por tabulaciones (*.dat). Archivos de texto ASCII con valores separados por tabuladores.

ASCII en formato fijo (*.dat). Archivos de texto ASCII con formato fijo, utilizando los formatos de escritura por defecto para todas las variables. No existen tabuladores ni espacios entre los campos de variable.

Excel 2.1(*.xls). Archivo de hoja de cálculo de Microsoft Excel 2,1. El número máximo de variables es de 256, y el número máximo de filas es de 16.384.

Excel 97 y posteriores (*.xls). Archivo de hoja de cálculo de Microsoft Excel 97/2000/XP. El número máximo de variables es de 256, y el número máximo de filas es de 65536.

1-2-3 versión 3.0 (*.wk3). Archivo de hoja de cálculo de Lotus 1-2-3, versión 3,0. El número máximo de variables que puede guardar es 256.

1-2-3 versión 2.0 (*.wk1). Archivo de hoja de cálculo de Lotus 1-2-3, versión 2,0. El número máximo de variables que puede guardar es 256.

1-2-3 versión 1.0 (*.wks). Archivo de hoja de cálculo de Lotus 1-2-3, versión 1A. El número máximo de variables que puede guardar es 256.

SYLK (*.slk). Formato de vínculo simbólico para archivos de hojas de cálculo de Microsoft Excel y de Multiplan. El número máximo de variables que puede guardar es 256.

dBASE IV (*.dbf). Formato dBASE IV.

dBASE III (*.dbf). Formato dBASE III.

dBASE II (*.dbf). Formato dBASE II.

SAS v6 para Windows (*.sd2). Formato de archivo de SAS v6 para Windows/OS2.

SAS v6 para UNIX (*.ssd01). Formato de archivo de SAS v6 para UNIX (Sun, HP, IBM).

SAS v6 para Alpha/OSF (*.ssd04). Formato de archivo de SAS v6 para Alpha/OSF (DEC UNIX).

SAS v7+ Windows extensión corta (*.sd7). Versión 7–8 de SAS para Windows con formato de nombre de archivo corto.

SAS v7+ Windows extensión larga (*.sas7bdat). Versión 7–8 de SAS para Windows con formato de nombre de archivo largo.

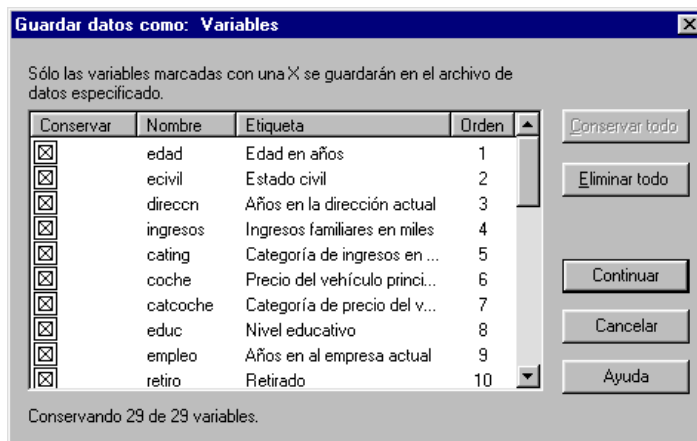
SAS v7+ para UNIX (*.ssd01). SAS v8 para UNIX.

Transporte de SAS (*.xpt). Archivo de transporte de SAS.

Almacenamiento de subconjuntos de variables

Figura 3-23

Cuadro de diálogo Guardar datos como: Variables



Para datos almacenados como un archivo de datos de SPSS, el cuadro de diálogo Guardar datos como: Variables permite seleccionar las variables que desea guardar en el nuevo archivo de datos. Por defecto, se almacenarán todas las variables. Anule la selección de las variables que no desea guardar o pulse en Eliminar todo y, a continuación, seleccione aquellas variables que desea guardar.

Para guardar un subconjunto de variables

- ▶ Active la ventana Editor de datos (pulse en cualquier punto de la ventana para activarla).
- ▶ Elija en los menús:
 - Archivo
 - Guardar como...
- ▶ Seleccione SPSS (*.sav) de la lista de tipos de archivos.
- ▶ Pulse en Variables.
- ▶ Seleccione las variables que desee almacenar.

Almacenamiento de archivos: Opciones

En los archivos de hoja de cálculo y los delimitados por tabuladores se pueden escribir nombres de variable en la primera fila del archivo.

Protección de datos originales

Para evitar la modificación o eliminación accidental de los datos originales, puede marcar el archivo como un archivo de sólo lectura.

- ▶ En los menús del Editor de datos, elija:
 - Archivo
 - Marcar archivo como de sólo lectura

Si hace modificaciones posteriores de los datos y a continuación, intenta guardar el archivo de datos, puede guardar los datos sólo con un nombre de archivo distinto; así, los datos originales no se verán afectados.

Puede restablecer los permisos de archivo a lectura/escritura seleccionando la opción Marcar archivo como de lectura/escritura en el menú Archivo.

Archivo activo virtual

El archivo activo virtual permite trabajar con grandes archivos de datos sin que sea necesaria una cantidad igual de grande (o mayor) de espacio temporal en disco. Para la mayoría de los procedimientos de análisis y gráficos, la fuente de datos original se vuelve a leer cada vez que se ejecuta un procedimiento diferente. Los procedimientos que modifican los datos necesitan una cierta cantidad de espacio temporal en disco para realizar un seguimiento de los cambios; además, algunas acciones necesitan disponer siempre de la cantidad suficiente de espacio en disco para, al menos, una copia completa del archivo de datos.

Figura 3-24
Requisitos de espacio temporal en disco

Acción	GET FILE = 'v1-5.sav'. FRECUENCIES...	COMPUTE v6 = ... RECODE v4... REGRESSION... /SAVE ZPRED...	SORT CASES BY... o CACHE																																																																																																																																					
Archivo virtual activo	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	11	12	13	14	15	21	22	23	24	25	31	32	33	34	35	41	42	43	44	45	51	52	53	54	55	61	62	63	64	65	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>1</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>2</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>3</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>4</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>5</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>6</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	v6	zpre	11	12	13	14	15	16	1	21	22	23	24	25	26	2	31	32	33	34	35	36	3	41	42	43	44	45	46	4	51	52	53	54	55	56	5	61	62	63	64	65	66	6	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>1</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>2</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>3</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>4</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>5</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>6</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	v6	zpre	11	12	13	14	15	16	1	21	22	23	24	25	26	2	31	32	33	34	35	36	3	41	42	43	44	45	46	4	51	52	53	54	55	56	5	61	62	63	64	65	66	6
v1	v2	v3	v4	v5																																																																																																																																				
11	12	13	14	15																																																																																																																																				
21	22	23	24	25																																																																																																																																				
31	32	33	34	35																																																																																																																																				
41	42	43	44	45																																																																																																																																				
51	52	53	54	55																																																																																																																																				
61	62	63	64	65																																																																																																																																				
v1	v2	v3	v4	v5	v6	zpre																																																																																																																																		
11	12	13	14	15	16	1																																																																																																																																		
21	22	23	24	25	26	2																																																																																																																																		
31	32	33	34	35	36	3																																																																																																																																		
41	42	43	44	45	46	4																																																																																																																																		
51	52	53	54	55	56	5																																																																																																																																		
61	62	63	64	65	66	6																																																																																																																																		
v1	v2	v3	v4	v5	v6	zpre																																																																																																																																		
11	12	13	14	15	16	1																																																																																																																																		
21	22	23	24	25	26	2																																																																																																																																		
31	32	33	34	35	36	3																																																																																																																																		
41	42	43	44	45	46	4																																																																																																																																		
51	52	53	54	55	56	5																																																																																																																																		
61	62	63	64	65	66	6																																																																																																																																		
Datos almacenados en espacio temporal del disco	Ninguno	<table border="1"> <thead> <tr><th>v4</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>14</td><td>16</td><td>1</td></tr> <tr><td>24</td><td>26</td><td>2</td></tr> <tr><td>34</td><td>36</td><td>3</td></tr> <tr><td>44</td><td>46</td><td>4</td></tr> <tr><td>54</td><td>56</td><td>5</td></tr> <tr><td>64</td><td>66</td><td>6</td></tr> </tbody> </table>	v4	v6	zpre	14	16	1	24	26	2	34	36	3	44	46	4	54	56	5	64	66	6	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>1</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>2</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>3</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>4</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>5</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>6</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	v6	zpre	11	12	13	14	15	16	1	21	22	23	24	25	26	2	31	32	33	34	35	36	3	41	42	43	44	45	46	4	51	52	53	54	55	56	5	61	62	63	64	65	66	6																																																															
v4	v6	zpre																																																																																																																																						
14	16	1																																																																																																																																						
24	26	2																																																																																																																																						
34	36	3																																																																																																																																						
44	46	4																																																																																																																																						
54	56	5																																																																																																																																						
64	66	6																																																																																																																																						
v1	v2	v3	v4	v5	v6	zpre																																																																																																																																		
11	12	13	14	15	16	1																																																																																																																																		
21	22	23	24	25	26	2																																																																																																																																		
31	32	33	34	35	36	3																																																																																																																																		
41	42	43	44	45	46	4																																																																																																																																		
51	52	53	54	55	56	5																																																																																																																																		
61	62	63	64	65	66	6																																																																																																																																		

Las acciones que no necesitan ningún espacio temporal en disco son:

- La lectura de archivos de datos de SPSS
- La fusión de dos o más archivos de datos de SPSS
- La lectura de tablas de bases de datos con el Asistente para bases de datos

- La fusión de un archivo de datos de SPSS con una tabla de base de datos
- La ejecución de procedimientos que leen datos (por ejemplo, Frecuencias, Tablas de contingencia, Explorar)

Las acciones que crean una o más columnas de datos en espacio temporal en disco son:

- El cálculo de nuevas variables
- La recodificación de variables existentes
- La ejecución de procedimientos que crean o modifican variables (por ejemplo, almacenamiento de valores pronosticados en Regresión lineal)

Las acciones que crean una copia completa del archivo de datos en espacio temporal en disco son:

- La lectura de archivos de Excel
- La ejecución de procedimientos que ordenan los datos (por ejemplo, Ordenar casos, Segmentar archivo)
- La lectura de datos con los comandos `GET TRANSLATE` o `DATA LIST`
- La utilización de la unidad Datos de caché o el comando `CACHE`
- La activación de otras aplicaciones de SPSS que leen el archivo de datos (por ejemplo, AnswerTree, DecisionTime)

Nota: El comando `GET DATA` proporciona una funcionalidad comparable a `DATA LIST`, sin crear una copia completa del archivo de datos en el espacio temporal del disco. El comando `SPLIT FILE` de la sintaxis de comandos no ordena el archivo de datos y por lo tanto no crea una copia del archivo de datos. Este comando, sin embargo, necesita tener los datos ordenados para un funcionamiento apropiado y la interfaz del cuadro de diálogo para este procedimiento ordenará de forma automática el archivo de datos, con la consiguiente copia completa de dicho archivo. En la versión para estudiantes no está disponible la sintaxis de comandos.

Acciones que crean una copia completa del archivo de datos por defecto:

- Lectura de bases de datos con el Asistente para bases de datos
- La lectura de archivos de texto con el Asistente para la importación de texto

El Asistente para la importación de texto proporciona un ajuste opcional para crear de forma automática una caché de los datos. Por defecto, se selecciona esta opción. Para desactivar esta opción, simplemente desmarque la casilla de verificación Caché local de los datos. En el Asistente para bases de datos puede pegar la sintaxis de comando generada y eliminar el comando `CACHE`.

Creación de una caché de datos

Aunque el archivo actual virtual puede reducir de forma drástica la cantidad de espacio temporal en disco necesario, la falta de una copia temporal del archivo “activo” significa que la fuente original de datos debe volver a leerse para cada procedimiento. Para archivos de datos grandes leídos desde una fuente externa, la creación de una copia temporal de los datos puede mejorar el rendimiento. Por ejemplo, para tablas de datos leídas desde una fuente de base de datos, la consulta SQL que lee la información de la base de datos debe volver a ejecutarse para cualquier comando o procedimiento que necesite leer los datos. Debido a que virtualmente todos los procedimientos de análisis estadísticos y procedimientos gráficos necesitan leer los datos, la ejecución de la consulta SQL se repite para cada procedimiento, lo que puede significar un importante incremento en el tiempo de procesamiento si se ejecuta un gran número de procedimientos.

Si se dispone de suficiente espacio en disco en el ordenador que realiza el análisis (el ordenador local o el servidor remoto), se pueden eliminar varias consultas SQL y mejorar el tiempo de procesamiento mediante la creación de una caché de datos del archivo activo. La caché de datos es una copia temporal de todos los datos.

Nota: Por defecto, el Asistente para bases de datos crea de forma automática una caché de datos, pero si se utiliza el comando `GET DATA` en la sintaxis de comandos para leer una base de datos, no se creará una caché de datos de forma automática. En la versión para estudiantes no está disponible la sintaxis de comandos.

Para crear una caché de datos

- ▶ Elija en los menús:
 - Archivo
 - Hacer caché de datos...

- ▶ Pulse en Aceptar o en Crear caché ahora.

Acceptar crea una caché de datos la siguiente vez que el programa lea los datos (por ejemplo, la próxima vez que se ejecute un procedimiento estadístico), que será lo que normalmente se quiera porque no necesita una lectura adicional de los datos. Crear caché ahora crea una caché de datos inmediatamente, lo cual no será necesario la mayoría de las veces. Crear caché ahora se utiliza principalmente por dos razones:

- Una fuente de datos está “bloqueada” y no se puede actualizar por nadie hasta que finalice la sesión actual, abra una fuente de datos diferente o haga una caché de los datos.
- Para grandes fuentes de datos, el desplazamiento por el contenido de la pestaña Vista de datos en el Editor de datos será mucho más rápido si se hace una caché de datos.

Para crear una caché de datos de forma automática

Se puede utilizar el comando `SET` para crear de forma automática una caché de datos después de un número especificado de cambios en el archivo de datos activo. Por defecto, se crea una caché del archivo de datos de forma automática cada 20 cambios realizados sobre el archivo.

- ▶ Elija en los menús:
 - Archivo
 - Nuevo
 - Sintaxis
- ▶ En la ventana de sintaxis, escriba `SET CACHE n`. (donde n representa el número de cambios realizados en el archivo de datos activo antes de crear una caché del archivo).
- ▶ En los menús de la ventana de sintaxis, elija:
 - Ejecutar
 - Todos

Nota: El ajuste de la caché no se almacena entre sesiones. Cada vez que se inicia una nueva sesión, se toma el valor por defecto de la opción que es 20.

Análisis en modo distribuido

El análisis en modo distribuido permite utilizar un ordenador que no es el local (o de escritorio) para realizar trabajos que requieren un gran consumo de memoria. Debido a que los servidores remotos utilizados para análisis distribuidos son normalmente más potentes y rápidos que los ordenadores locales, una utilización correcta del análisis en modo distribuido puede reducir significativamente el tiempo de procesamiento del ordenador. El análisis distribuido con un servidor remoto puede ser útil si el trabajo trata:

- Archivos de datos, en particular lecturas de datos de fuentes de bases de datos.
- Tareas que requieren un gran consumo de memoria. Cualquier tarea que tarde bastante tiempo en el análisis en modo local será una buena candidata para el análisis distribuido,

El análisis distribuido sólo afecta a las tareas relacionadas con los datos, como lectura de datos, transformación de datos, cálculo de nuevas variables y cálculo de estadísticos. No tiene ningún efecto sobre tareas relacionadas con la edición de resultados, como la manipulación de tablas pivote o la modificación de gráficos.

Nota: El análisis distribuido está disponible sólo si ambos equipos tienen una versión local y acceso a una versión de servidor con licencia del software instalado en un servidor remoto.

Análisis distribuido respecto al análisis local

A continuación, se enumeran algunas normas para elegir entre el análisis en modo distribuido o en modo local:

Acceso a bases de datos. Los trabajos que llevan a cabo consultas en la base de datos se ejecutarán con mayor rapidez en modo distribuido si el servidor dispone de un acceso superior a la base de datos o si el servidor se está ejecutando en la misma máquina

que el motor de base de datos. Si el software necesario para acceder a la base de datos sólo está disponible en el servidor o si el administrador de red no permite descargar grandes tablas de datos, sólo podrá acceder a la base de datos en modo distribuido.

Relación entre cálculos y resultados. Los comandos que realizan gran cantidad de cálculos y generan pocos resultados de salida (por ejemplo, tablas pivote pequeñas y pocas, breves resultados de texto, o pocos y sencillos gráficos) son los que más aprovechan la ejecución en modo distribuido. El grado de mejora depende sobre todo de la capacidad de cálculo del servidor remoto.

Trabajos pequeños. Los trabajos que se ejecutan con rapidez en modo local, siempre se ejecutarán más lentamente en modo distribuido debido a la carga general inherente al cliente /servidor.

Gráficos. Los gráficos orientados a los casos, como los diagramas de dispersión, los gráficos de residuos de regresión y los gráficos de secuencia, necesitan que los datos brutos se encuentren en el ordenador local. Para archivos de datos y tablas de bases de datos grandes, esto puede repercutir en un descenso del rendimiento en el modo distribuido debido a que los datos deben enviarse desde el servidor remoto al ordenador local. Otros gráficos se basan en datos resumidos o agregados y deben realizarse de forma adecuada ya que la agregación se lleva a cabo en el servidor.

Gráficos interactivos. Como es posible guardar los datos brutos con los gráficos interactivos (un ajuste opcional), esto puede producir que se transfieran grandes cantidades de datos desde el servidor remoto al ordenador local, incrementando de forma significativa el tiempo que se tarda en guardar los resultados.

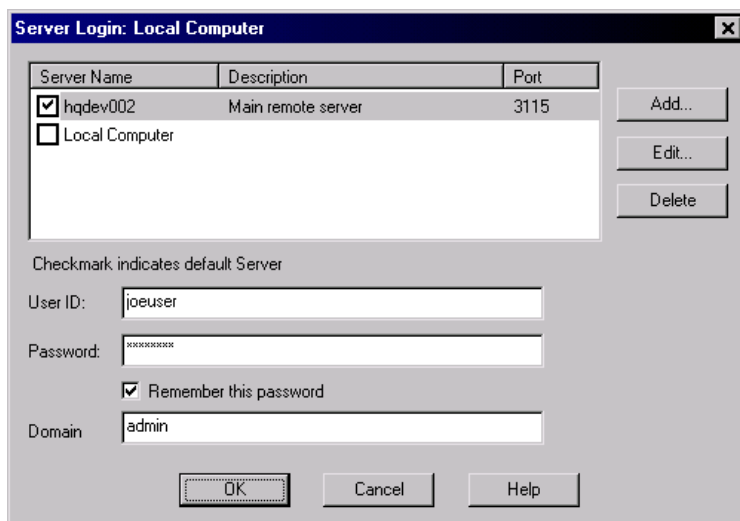
Tablas pivote. Las tablas pivote grandes pueden tardar más en crearse en el modo distribuido. Esto es particularmente cierto para el procedimiento Cubos OLAP y para las tablas que contienen datos sobre los casos individuales, como las disponibles en el procedimiento Resumir.

Resultados de texto. Cuanto más texto se produzca, más lento será en modo distribuido, ya que este texto se genera en el servidor remoto y se copia al ordenador local para ser visualizado. Sin embargo, los resultados de texto tienen una carga general pequeña y suelen transmitirse rápidamente.

Acceso al servidor

El cuadro de diálogo Acceso al servidor permite seleccionar el ordenador para procesar comandos y ejecutar procedimientos. Podrá ser el ordenador local o un servidor remoto.

Figura 4-1
Cuadro de diálogo Acceso al servidor



Se pueden añadir, modificar o eliminar servidores remotos de la lista. Los servidores remotos requieren normalmente un ID de usuario y una contraseña; también puede ser necesario un nombre de dominio. Póngase en contacto con el administrador del sistema para obtener información acerca de servidores, ID de usuario y contraseñas, nombres de dominio disponibles y demás información necesaria para la conexión.

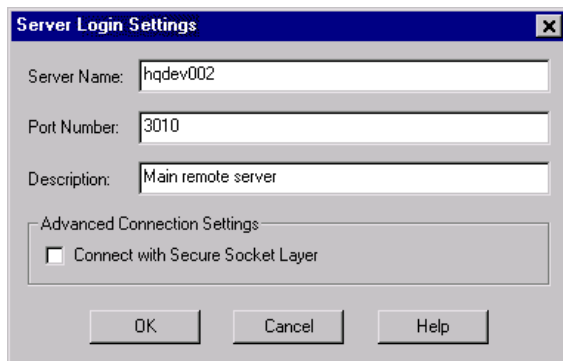
Puede seleccionar un servidor por defecto y guardar el ID de usuario, nombre de dominio y contraseña asociados a cualquier servidor. De esta manera, se conectará de forma automática al servidor por defecto en el momento de iniciar la sesión.

Adición y edición de la configuración de acceso al servidor

Utilice el cuadro de diálogo Configuración del acceso al servidor para añadir o editar la información de conexión para servidores remotos para utilizar en los análisis en modo distribuido.

Figura 4-2

Cuadro de diálogo Configuración del acceso al servidor



Para obtener una lista de servidores disponibles, los números de puerto para dichos servidores y toda la información adicional necesaria para la conexión, póngase en contacto con el administrador del sistema. No utilice el Nivel de socket seguro a menos que lo indique el administrador.

Nombre del servidor. Un “nombre” de servidor puede ser un nombre alfanumérico asignado a un ordenador (por ejemplo, hqdev001) o una dirección IP única asignada a un ordenador (por ejemplo, 202.123.456.78).

Número de puerto. El número de puerto es el puerto que el software del servidor utiliza para las comunicaciones.

Descripción. Introduzca una descripción opcional para que se visualice en la lista de servidores.

Conectar con Nivel de socket seguro. Las encriptaciones de Nivel de socket seguro (SSL) requieren el análisis distribuido cuando se envían al servidor remoto de SPSS. Antes de utilizar el SSL, consulte con el administrador. El SSL debe estar configurado en su equipo y el servidor de esta opción debe estar activado.

Para seleccionar, cambiar o añadir servidores

- ▶ Elija en los menús:
 - Archivo
 - Cambiar servidor...

Para seleccionar un servidor por defecto:

- ▶ En la lista de servidores, seleccione la casilla que se encuentra junto al servidor que desea utilizar.
- ▶ Introduzca el ID de usuario, nombre de dominio y contraseña suministrados por el administrador.

Nota: De esta manera, se conectará de forma automática al servidor por defecto en el momento de iniciar la sesión.

Para cambiar a otro servidor:

- ▶ Seleccione el servidor de la lista.
- ▶ Introduzca el ID de usuario, nombre de dominio y contraseña (si es necesario).

Nota: Al cambiar de servidor durante una sesión, se cierran todas las ventanas abiertas. Se solicitará guardar los cambios antes de que se cierren las ventanas.

Para añadir un servidor:

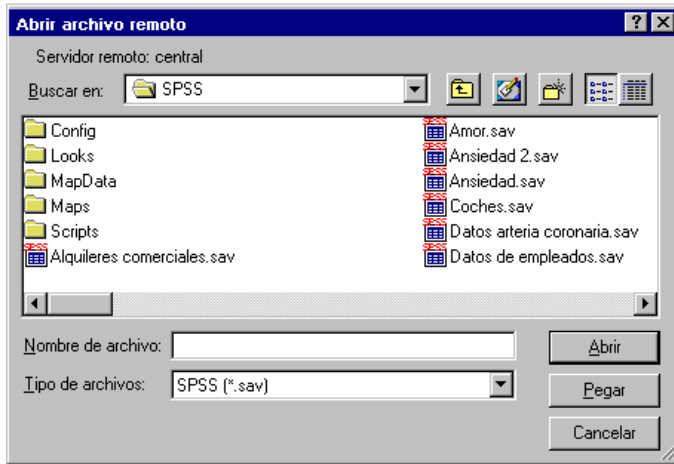
- ▶ Solicite al administrador la información de conexión del servidor.
- ▶ Pulse en Añadir para abrir el cuadro de diálogo Configuración del acceso al servidor.
- ▶ Introduzca la información de conexión y la configuración opcional y pulse en Aceptar.

Para editar un servidor:

- ▶ Solicite al administrador la información de conexión revisada.
- ▶ Pulse en Editar para abrir el cuadro de diálogo Configuración del acceso al servidor.
- ▶ Introduzca los cambios y pulse en Aceptar.

Apertura de archivos de datos desde un servidor remoto

Figura 4-3
Cuadro de diálogo Abrir archivo remoto



En el análisis en modo distribuido, el cuadro de diálogo Abrir archivo remoto sustituye al cuadro de diálogo estándar Abrir archivo.

- La lista de archivos, carpetas y unidades muestra lo que hay disponible en o desde el servidor remoto. En la parte superior del cuadro de diálogo se indica el nombre del servidor actual.
- En el análisis en modo distribuido, sólo tendrá acceso a los archivos de datos del equipo local si especifica la unidad como un dispositivo compartido y las carpetas que contienen los archivos de datos como carpetas compartidas.
- Si el servidor está ejecutando un sistema operativo diferente (por ejemplo, usted dispone de Windows y el servidor se ejecuta bajo UNIX), probablemente no dispondrá de acceso a los archivos de datos locales en el análisis en modo distribuido, aunque los archivos estén en carpetas compartidas.

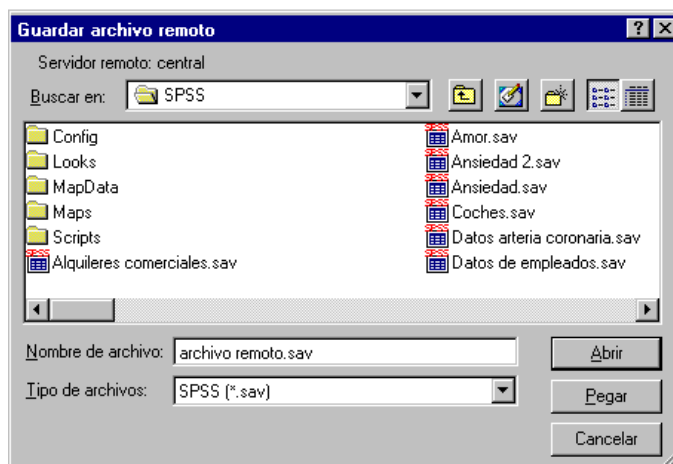
Sólo se puede abrir un archivo de datos cada vez. El archivo abierto se cierra automáticamente cuando se abre otro archivo de datos. Si desea mantener varios archivos de datos abiertos al mismo tiempo, puede iniciar varias sesiones.

Para abrir archivos de datos desde un servidor remoto

- ▶ Si aún no está conectado al servidor remoto, acceda al servidor remoto:
- ▶ Según el tipo de archivo de datos que quiera abrir, elija en los menús:
 - Archivo
 - Abrir
 - Datos...
 - o
 - Archivo
 - Abrir base de datos
 - o
 - Archivo
 - Leer datos de texto...

Almacenamiento de archivos de datos desde un servidor remoto

Figura 4-4
Cuadro de diálogo Guardar archivo remoto



En el análisis en modo distribuido, el cuadro de diálogo Guardar archivo remoto sustituye al cuadro de diálogo estándar Guardar archivo.

La lista de carpetas y unidades muestra lo que hay disponible en o desde el servidor remoto. En la parte superior del cuadro de diálogo se indica el nombre del servidor actual. Sólo tendrá acceso a las carpetas del equipo local si especifica la unidad como un dispositivo compartido y las carpetas como carpetas compartidas. Si el servidor está ejecutando un sistema operativo diferente (por ejemplo, usted dispone de Windows y el servidor se ejecuta bajo UNIX), probablemente no dispondrá de acceso a los archivos de datos locales en el análisis en modo distribuido, aunque los archivos estén en carpetas compartidas. Los derechos sobre las carpetas compartidas deben incluir el permiso de escritura en la carpeta si quiere guardar archivos de datos en una carpeta local.

Para guardar archivos de datos desde un servidor remoto

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Elija en los menús:
 - Archivo
 - Guardar (o Guardar como...)

Acceso a archivo de datos en análisis en modo local y Moda

La presentación de los archivos de datos, las carpetas (directorios) y las unidades para el ordenador local y la red está en función del ordenador que está utilizando para procesar comandos y ejecutar procedimientos, que no es necesariamente el ordenador que tiene delante.

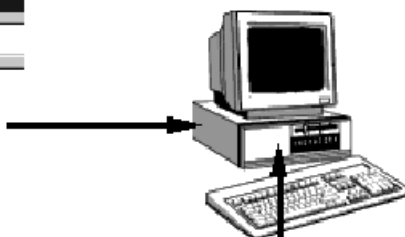
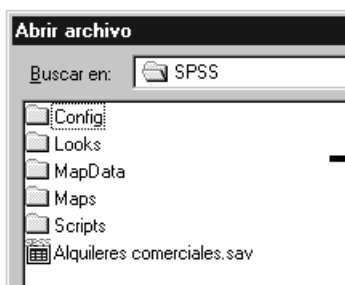
Análisis en modo local. Cuando utiliza el ordenador local como el “servidor”, la visualización de los archivos de datos, las carpetas y las unidades que ve en el cuadro de diálogo de acceso a los archivos (para la apertura de archivos de datos) es similar a lo que ve en otras aplicaciones o en el Explorador de Windows. Se pueden ver todos los archivos de datos y las carpetas en el ordenador y cualquier archivo y carpeta en las unidades de red que normalmente ve.

Análisis en modo distribuido. Cuando utiliza otro ordenador como “servidor remoto” para ejecutar comandos y procedimientos, la visualización de los archivos de datos y las unidades representa la vista desde la perspectiva del servidor remoto. Aunque vea nombres de carpetas que le son familiares, como *Archivos de programas* y

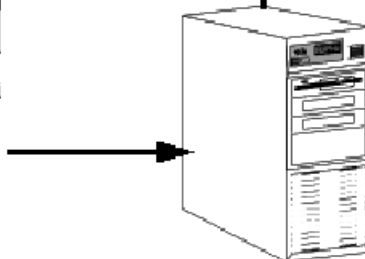
unidades como C, estas *no* son las carpetas y unidades del ordenador local, son las del servidor remoto.

Figura 4-5
Vistas de los equipos local y remoto

Vista del sistema local



Vista del sistema remoto



En el análisis en modo distribuido, sólo tendrá acceso a los archivos de datos del equipo local si especifica la unidad como un dispositivo compartido y las carpetas que contienen los archivos de datos como carpetas compartidas. Si el servidor está ejecutando un sistema operativo diferente (por ejemplo, si dispone de Windows y el servidor se ejecuta bajo UNIX), probablemente no dispondrá de acceso a los archivos de datos locales en el análisis en modo distribuido, aunque los archivos estén en carpetas compartidas.

El análisis en modo distribuido no es lo mismo que acceder a archivos de datos que se encuentran en otro ordenador de la red. Se puede acceder a archivos de datos en otros dispositivos de red tanto en análisis en modo local como en análisis en modo distribuido. En modo local, se accede a otros dispositivos desde el ordenador local. En el modo distribuido, se accede a otros dispositivos de red desde el servidor remoto.

Si no está seguro de si está utilizando el análisis en modo local o distribuido, mire la barra de título en el cuadro de diálogo para acceder a archivos de datos. Si el título del cuadro de diálogo contiene la palabra *remoto* (como en Abrir archivo remoto) o si el texto Servidor remoto: [nombre del servidor] aparece en la parte superior del cuadro de diálogo, significa que está en modo distribuido.

Nota: Esto afecta sólo a los cuadros de diálogo para acceder a archivos de datos (por ejemplo, Abrir datos, Guardar datos, Abrir base de datos y Aplicar diccionario de datos). Para todos los demás tipos de archivos (por ejemplo, archivos del Visor, archivos de sintaxis y archivos de procesos) se utiliza la visualización local.

Para establecer derechos de uso compartido para una unidad o carpeta

- ▶ En Mi PC, pulse en la carpeta (directorio) o unidad que quiera compartir.
- ▶ En el menú Archivo, pulse en Propiedades.
- ▶ Pulse en la pestaña Compartir y, a continuación, pulse en Compartido como.

Si desea obtener más información acerca de compartir unidades y carpetas, consulte la ayuda del sistema operativo.

Disponibilidad de procedimientos en análisis en modo Moda

En el análisis en modo distribuido, sólo están disponibles aquellos procedimientos instalados en la versión local y en la versión del servidor remoto. No se pueden utilizar procedimientos instalados en el servidor y que no lo están en la versión local, ni tampoco procedimientos instalados en la versión local y que no lo están en el servidor remoto.

Mientras esta última situación puede ser poco probable, es posible que se disponga de componentes opcionales instalados en el ordenador local que no están disponibles en el servidor remoto. Si este es el caso, cambiar del ordenador local a un servidor remoto provocará la eliminación de los procedimientos afectados de los menús y

la sintaxis de comandos relacionada generará errores. Todos los procedimientos afectados se restaurarán al cambiar de nuevo al modo local.

Para utilizar las especificaciones de ruta de acceso UNC

Con la versión de servidor de SPSS para Windows NT, las especificaciones de ruta de acceso relativas para los archivos de datos son relativas al servidor actual con análisis en modo distribuido y no lo son respecto al ordenador local. En términos prácticos, esto significa que una especificación de ruta de acceso como *c:\misdocs\misdatos.sav* no indica un directorio y archivo en la unidad local C; sino que indica un directorio y archivo en el disco duro del servidor remoto. Si el directorio y/o archivo no existen en el servidor remoto, se generará un error en la sintaxis del comando, como en:

```
GET FILE='c:\misdocs\misdatos.sav'.
```

Si utiliza la versión de servidor de SPSS para Windows NT, puede usar las especificaciones de la UNC (convención de denominación universal) al acceder a los archivos de datos mediante la sintaxis de comandos. El formato general de una especificación UNC es:

```
\\nombre_servidor\nombre_compartido\ruta\nombre_archivo
```

- *Nombre_servidor* es el nombre del ordenador que contiene el archivo de datos.
- *Nombre_compartido* es la carpeta (directorio) en el ordenador que aparece designada como una carpeta compartida.
- *Ruta* es cualquier ruta de acceso de carpetas (subdirectorios) por debajo de la carpeta compartida.
- *Nombre_archivo* es el nombre del archivo de datos.

Por ejemplo:

```
GET FILE = '\\hqdev001\public\julio\ventas.sav'.
```

Si el ordenador no tiene un nombre asignado, puede utilizar su dirección IP, como en:

```
GET FILE = '\\204.125.125.53\public\julio\ventas.sav'.
```

Incluso con especificaciones de ruta de acceso UNC, sólo se pueden acceder a archivos de datos que estén en carpetas y dispositivos compartidos. Cuando se utiliza análisis en modo distribuido, esto incluye archivos de datos del ordenador local.

Servidores UNIX. En las plataformas UNIX, no hay un equivalente de las rutas UNC y todas las rutas de acceso de los directorios deben ser rutas absolutas que comienzan en la raíz del servidor; las rutas relativas no están permitidas. Por ejemplo, si el archivo de datos está ubicado en */bin/spss/datos* y el directorio actual también es */bin/spss/datos*, la sintaxis `GET FILE='ventas.sav'` no es válida; se debe especificar la ruta completa, como en:

```
GET FILE='/bin/datos/spss/ventas.sav'.
```

Editor de datos

El Editor de datos proporciona un método práctico (al estilo de las hojas de cálculo) para la creación y edición de archivos de datos. La ventana Editor de datos se abre automáticamente cuando se inicia una sesión.

El Editor de datos proporciona dos vistas de los datos.

- **Vista de datos.** Muestra los valores de datos reales o las etiquetas de valor definidas.
- **Vista de variables.** Muestra la información de definición de variable, que incluye las etiquetas de la variable definida y de valor, tipo de dato (por ejemplo, de cadena, fecha y numérico), nivel de medida (nominal, ordinal o de escala) y los valores perdidos definidos por el usuario.

En ambas vistas, se puede añadir, modificar y eliminar la información contenida en el archivo de datos.

Vista de datos

Figura 5-1
Vista de datos

	id	sexo	fechnac	educ	catlab	sa
1	1	h	03.02.52	15	3	\$
2	2	h	23.05.58	16	1	\$
3	3	m	26.07.29	12	1	\$
4	4	m	15.04.47	8	1	\$
5	5	h	09.02.55	15	1	\$
6	6	h	22.08.58	15	1	\$
7	7	h	26.04.56	15	1	\$
8	8	m	06.05.66	12	1	\$
9	9	m	23.01.46	15	1	\$

Muchas de las funciones de la Vista de datos son similares a las que se encuentran en aplicaciones de hojas de cálculo. Sin embargo, existen varias diferencias importantes:

- Las filas son casos. Cada fila representa un caso o una observación. Por ejemplo, cada individuo que responde a un cuestionario es un caso.
- Las columnas son variables. Cada columna representa una variable o una característica que se mide. Por ejemplo, cada elemento en un cuestionario es una variable.
- Las casillas contienen valores. Cada casilla contiene un valor único de una variable para cada caso. La casilla es la intersección del caso y la variable. Las casillas sólo contienen valores de datos. A diferencia de los programas de hoja de cálculo, las casillas del Editor de datos no pueden contener fórmulas.
- El archivo de datos es rectangular. Las dimensiones del archivo de datos vienen determinadas por el número de casos y de variables. Se pueden introducir datos en cualquier casilla. Si introduce datos en una casilla fuera de los límites del archivo de datos definido, el rectángulo de datos se ampliará para incluir todas las filas y columnas situadas entre esa casilla y los límites del archivo. No hay casillas “vacías” en los límites del archivo de datos. Para variables numéricas, las casillas vacías se convierten en el valor perdido del sistema. Para variables de cadena, un espacio en blanco se considera un valor válido.

Vista de variables

Figura 5-2
Vista de variables

	Nombre	Tipo	Anchura	Decimales	Etiqueta
1	id	Numérico	4	0	Código de empleado
2	sexo	Cadena	1	0	Sexo
3	fechnac	Fecha	8	0	Fecha de nacimiento
4	educ	Numérico	2	0	Nivel educativo
5	catlab	Numérico	1	0	Categoría laboral
6	salario	Dólar	8	0	Salario actual
7	salini	Dólar	8	0	Salario inicial
8	tiempemp	Numérico	2	0	Meses desde el contr
9	expPrev	Numérico	6	0	Experiencia previa (m
10	minoría	Numérico	1	0	Clasificación étnica

La Vista de variables contiene descripciones de los atributos de cada variable del archivo de datos. En la Vista de variables:

- Las filas son variables.
- Las columnas son atributos de las variables.

Se pueden añadir o eliminar variables, y modificar los atributos de las variables, incluyendo:

- Nombre de variable
- Tipo de datos
- Número de dígitos o caracteres
- Número de decimales
- Las etiquetas descriptivas de variable y de valor.
- Valores perdidos definidos por el usuario
- Ancho de columna.
- Nivel de medida

Todos estos atributos se guardan al guardar el archivo de datos.

Además de la definición de propiedades de variables en la vista Variable, hay dos otros métodos para definir las propiedades de variables:

- El Asistente para la copia de propiedades de datos ofrece la posibilidad de utilizar un archivo de datos de SPSS como plantilla para definir las propiedades de las variables y del archivo del archivo de datos de trabajo. También puede utilizar las variables del archivo de datos de trabajo como plantillas para el resto de las variables del archivo de datos de trabajo. La opción Copiar propiedades de datos está disponible en el menú Datos en la ventana Editor de datos.
- La opción Definir propiedades de variables (también disponible en el menú Datos de la ventana Editor de datos) explora los datos y muestra una lista con todos los valores de datos únicos para las variables seleccionadas, indica los valores sin etiquetas y ofrece una función de etiquetas automáticas. Esta opción es especialmente útil para variables categóricas que utilizan códigos numéricos para representar las categorías (por ejemplo, 0 = *hombre*, 1 = *mujer*).

Para visualizar o definir los atributos de las variables

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña Vista de variables.
- ▶ Para definir variables nuevas, introduzca un nombre de variable en cualquier fila vacía.
- ▶ Seleccione los atributos que desea definir o modificar.

Nombres de variable

Para los nombres de variable se aplican las siguientes normas:

- El nombre debe comenzar por una letra. Los demás caracteres pueden ser letras, dígitos, puntos o los símbolos @, #, _ o \$.
- Los nombres de variable no pueden terminar en punto.
- Se deben evitar los nombres de variable que terminan con subrayado (para evitar conflictos con las variables creadas automáticamente por algunos procedimientos).

- La longitud del nombre no debe exceder los 64 bytes. Sesenta y cuatro bytes suelen equivaler a 64 caracteres en idiomas de un solo byte (por ejemplo, inglés, francés, alemán, español, italiano, hebreo, ruso, griego, árabe, tailandés) y 32 caracteres en los idiomas de dos bytes (por ejemplo, japonés, chino, coreano).
- No se pueden utilizar espacios en blanco ni caracteres especiales (por ejemplo, !, ?, ' y *).
- Cada nombre de variable debe ser único; no se permiten duplicados.
- Las palabras reservadas no se pueden utilizar como nombres de variable. Las palabras reservadas son: ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH.
- Los nombres de variable se pueden definir combinando de cualquier manera caracteres en mayúsculas y en minúsculas, esta distinción entre mayúsculas y minúsculas se conserva en lo que se refiere a la presentación.
- Cuando es necesario dividir los nombres largos de variable en varias líneas en los resultados, SPSS intenta dividir las líneas aprovechando los subrayados, los puntos y los cambios de minúsculas a mayúsculas.

Nivel de medida de variable

Puede especificar el nivel de medida como Escala (datos numéricos de una escala de intervalo o de razón), Ordinal o Nominal. Los datos nominales y ordinales pueden ser de cadena (alfanuméricos) o numéricos. La especificación de medida sólo es relevante para:

- procedimientos de los gráficos o de tablas personalizadas que identifican las variables como de escala o categóricas. Las variables nominales y ordinales se tratan como categóricas. (Las tablas personalizadas sólo están disponibles en el componente adicional Tablas.)
- Los archivos de datos con formato SPSS utilizados con AnswerTree.

Se puede seleccionar uno de los tres niveles de medida:

Escala. Los valores de los datos son numéricos en una escala de intervalo o de razón (por ejemplo, *edad* o *ingresos*). Las variables de escala deben ser numéricas.

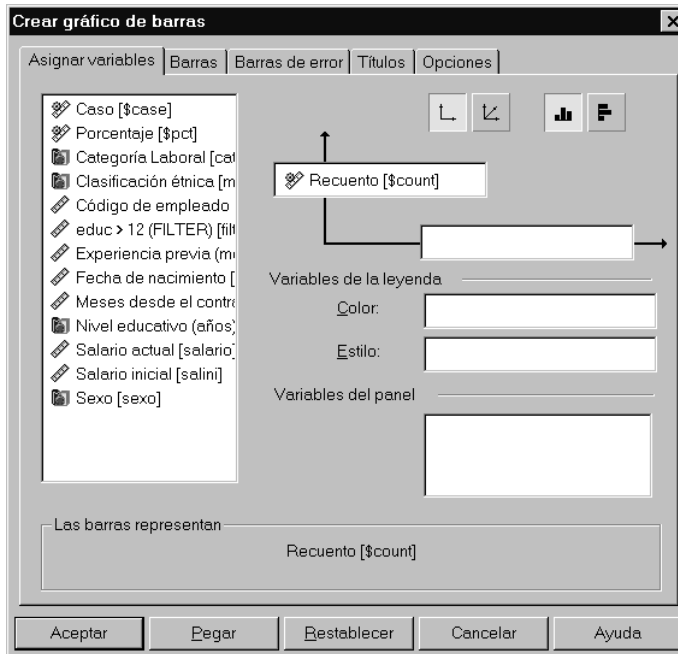
Ordinal. Los valores de datos representan categorías con un cierto orden intrínseco (por ejemplo, *bajo, medio, alto; totalmente de acuerdo, de acuerdo, en desacuerdo, totalmente en desacuerdo*). Las variables ordinales pueden ser valores de cadena (alfanuméricos) o numéricos que representen diferentes categorías (por ejemplo, 1 = *bajo*, 2 = *medio*, 3 = *alto*).

Nota: Para variables de cadena ordinales, se asume que el orden alfabético de los valores de cadena indica el orden correcto de las categorías. Por ejemplo, en una variable de cadena cuyos valores sean *bajo, medio, alto*, se interpreta el orden de las categorías como *alto, bajo, medio* (orden que no es el correcto). Por norma general, se puede indicar que es más fiable utilizar códigos numéricos para representar datos ordinales.

Nominal. Los valores de datos representan categorías sin un orden intrínseco (por ejemplo, categoría laboral o división de la compañía). Las variables nominales pueden ser valores de cadena (alfanuméricos) o numéricos que representen diferentes categorías (por ejemplo, 1 = *hombre*, 2 = *mujer*).

Figura 5-3

Las variables categóricas y de escala en procedimientos de gráficos



Para los archivos de datos con formato SPSS creados en versiones anteriores de productos de SPSS son aplicables las reglas siguientes:

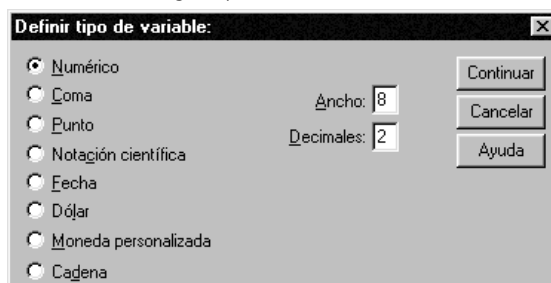
- Las variables de cadena (alfanuméricas) se establecen en nominales.
- Las variables de cadena y numéricas con etiquetas de valor definidas se establecen en ordinales.
- Las variables numéricas sin etiquetas de valor definidas que no superen un número especificado de valores únicos se establecen en ordinales.
- Las variables numéricas sin etiquetas de valor definidas que superen un número especificado de valores únicos se establecen en variables de escala.

Por defecto, el número de valores únicos es de 24. Para cambiar el valor especificado, modifique las opciones de los gráficos interactivos: en el menú Edición, seleccione Opciones y pulse en la pestaña Interactivo.

Tipo de variable

Tipo de variable especifica los tipos de datos de cada variable. Por defecto se asume que todas las variables nuevas son numéricas. Se puede utilizar Tipo de variable para cambiar el tipo de datos. El contenido del cuadro de diálogo Tipo de variable depende del tipo de datos seleccionado. Para algunos tipos de datos, hay cuadros de texto para el ancho y el número de decimales; para otros, simplemente puede seleccionar un formato de una lista desplegable de ejemplos.

Figura 5-4
Cuadro de diálogo Tipo de variable



Los tipos de datos disponibles son los siguientes:

Numérico. Una variable cuyos valores son números. Los valores se muestran en formato numérico estándar. El Editor de datos acepta valores numéricos en formato estándar o en notación científica.

Coma. Una variable numérica cuyos valores se muestran con comas que delimitan cada tres posiciones y con el punto como delimitador decimal. El Editor de datos acepta valores numéricos para este tipo de variables con o sin comas, o bien en notación científica. Los valores no pueden contener comas a la derecha del indicador decimal.

Punto. Una variable numérica cuyos valores se muestran con puntos que delimitan cada tres posiciones y con la coma como delimitador decimal. El Editor de datos acepta valores numéricos para este tipo de variables con o sin puntos, o bien en notación científica. Los valores no pueden contener puntos a la derecha del indicador decimal.

Notación científica. Una variable numérica cuyos valores se muestran con una E intercalada y un exponente con signo que representa una potencia de base diez. El Editor de datos acepta para estas variables valores numéricos con o sin el exponente. El exponente puede aparecer precedido por una E o una D con un signo opcional, o bien sólo por el signo (por ejemplo, 123, 1,23E2, 1,23D2, 1,23E+2 e incluso 1,23+2).

Fecha. Una variable numérica cuyos valores se muestran en uno de los diferentes formatos de fecha-calendario u hora-reloj. Seleccione un formato de la lista. Puede introducir las fechas utilizando como delimitadores: barras, guiones, puntos, comas o espacios. El rango de siglo para los valores de año de dos dígitos está determinado por la configuración de las Opciones (en el menú Edición, seleccione Opciones y pulse en la pestaña Datos).

Moneda personalizada. Una variable numérica cuyos valores se muestran en uno de los formatos de moneda personalizados que se hayan definido previamente en la pestaña Moneda del cuadro de diálogo Opciones. Los caracteres definidos en la moneda personalizada no se pueden emplear en la introducción de datos pero sí se mostrarán en el Editor de datos.

Cadena. Variable cuyos valores no son numéricos y, por lo tanto, no se utilizan en los cálculos. Pueden contener cualquier carácter siempre que no se exceda la longitud definida. Las mayúsculas y las minúsculas se consideran diferentes. También se conoce como variable alfanumérica.

Para definir el tipo de variable

- ▶ Pulse en el botón de la casilla *Tipo* para la variable que se quiere definir.
- ▶ Seleccione el tipo de datos en el cuadro de diálogo Tipo de variable.

Formatos de entrada frente a formatos de presentación

Dependiendo del formato, la presentación de valores en la Vista de datos puede ser diferente del valor real que se ha introducido y almacenado internamente. A continuación, se proporcionan algunas normas generales:

- Para formatos numéricos, de coma y de punto, se pueden introducir valores con cualquier número de dígitos decimales (hasta 16) y el valor completo se almacena internamente. La Vista de datos muestra sólo el número definido de dígitos decimales y redondea los valores con más decimales. Sin embargo, el valor completo se utiliza en todos los cálculos.
- Para las variables de cadena, todos los valores se rellenan por la derecha hasta el ancho máximo. Para una variable de cadena con un ancho de tres, un valor de *No* se almacena internamente como 'No ' y no es equivalente a ' No '.
- Para formatos de fecha, se pueden utilizar guiones, barras, espacios, comas o puntos como separadores entre valores de día, mes y año; se pueden introducir números, abreviaciones de tres letras o nombres completos para el valor de mes. Las fechas del formato general dd-mmm-aa aparecen separadas por guiones y con abreviaciones de tres letras para el mes. Las fechas del formato general dd/mm/aa y mm/dd/aa se muestran con barras como separadores y números para el mes. Internamente, las fechas se almacenan como el número de segundos transcurridos desde el 14 de octubre de 1582. El rango de siglo para los valores de año de dos dígitos está determinado por la configuración de las Opciones (en el menú Edición, seleccione Opciones y pulse en la pestaña Datos).
- Para formatos de hora, se pueden utilizar dos puntos, puntos o espacios como separadores entre horas, minutos y segundos. Las horas se muestran separadas por dos puntos. Internamente, las horas se almacenan como el número de segundos transcurridos desde el 14 de octubre de 1582.

Etiquetas de variable

Puede asignar etiquetas de variable descriptivas de hasta 256 caracteres de longitud (128 caracteres en los idiomas de doble byte). Las etiquetas de variable pueden contener espacios y caracteres reservados que no se admiten en los nombres de variable.

Para especificar etiquetas de variable

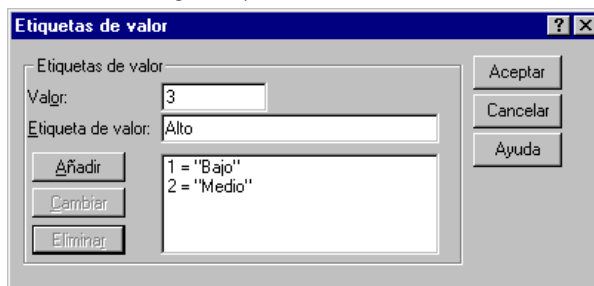
- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña Vista de variables.
- ▶ Escriba la etiqueta de variable descriptiva en la casilla *Etiqueta* para la variable.

Etiquetas de valor

Puede asignar etiquetas de valor descriptivas a cada valor de una variable. Esta posibilidad es especialmente útil si los archivos de datos utilizan códigos numéricos para representar categorías que no son numéricas (por ejemplo, códigos 1 y 2 para *hombre* y *mujer*).

- Las etiquetas de valor pueden tener hasta 60 caracteres.
- Las etiquetas de valor no están disponibles para las variables de cadena larga (variables de cadena de más de 8 caracteres).

Figura 5-5
Cuadro de diálogo *Etiquetas de valor*



Para especificar etiquetas de valor

- ▶ Pulse en el botón de la casilla *Valores* para la variable que se quiere definir.
- ▶ Para cada valor, escriba el valor y una etiqueta.
- ▶ Pulse en Añadir para introducir la etiqueta de valor.

Inserción de saltos de línea en etiquetas

Las etiquetas de valor y de variable se dividen automáticamente en varias líneas en los gráficos y en las tablas pivote si el ancho de casilla o el área no es suficiente para mostrar la etiqueta entera en una línea. Se pueden editar los resultados para insertar saltos de línea manuales si se quiere dividir la etiqueta en un punto diferente. También puede crear etiquetas de variable y de valor que *siempre* se dividan en puntos especificados y se muestren en varias líneas:

- ▶ Para etiquetas de variable, seleccione la casilla Etiqueta de la variable en la Vista de variables del Editor de datos.
- ▶ Para etiquetas de valor, seleccione la casilla Valores correspondiente a la variable en la Vista de variables del Editor de datos, pulse en el botón que aparece en la casilla y, a continuación, seleccione la etiqueta que desea modificar en el cuadro de diálogo Etiquetas de valor.
- ▶ En el punto de la etiqueta en el que desea dividir la etiqueta, escriba \n.

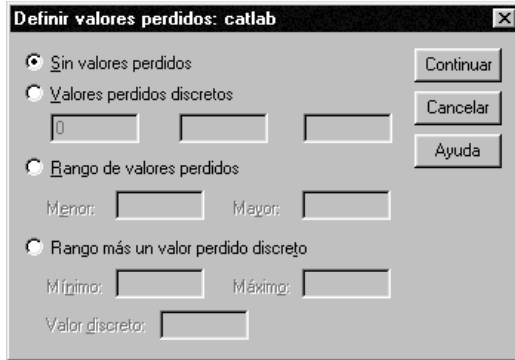
El \n no aparece en las tablas pivote ni en los gráficos; se interpreta como un carácter de salto de línea.

Valores perdidos

Valores perdidos define los valores de los datos definidos como **perdidos por el usuario**. A menudo es útil para saber por qué se pierde información. Por ejemplo, puede desear distinguir los datos perdidos porque un encuestado se niegue a responder de los datos perdidos porque la pregunta no afecta a dicho encuestado. Los valores

de datos especificados como perdidos por el usuario aparecen marcados para un tratamiento especial y se excluyen de la mayoría de los cálculos.

Figura 5-6
Cuadro de diálogo Valores perdidos



- Se pueden introducir hasta tres valores perdidos (individuales) de tipo discreto, un rango de valores perdidos o un rango más un valor de tipo discreto.
- Sólo pueden especificarse rangos para las variables numéricas.
- No se pueden definir los valores perdidos para variables de cadena larga (variables de cadena de más de ocho caracteres).

Valores perdidos para las variables de cadena. Se considera que son válidos todos los valores de cadena, incluidos los valores vacíos o nulos, a no ser que se definan explícitamente como perdidos. Para definir como perdidos los valores nulos o vacíos de una variable de cadena, escriba un espacio en blanco en uno de los campos debajo de la selección Valores perdidos discretos.

Para definir los valores perdidos

- ▶ Pulse en el botón de la casilla *Perdido* para la variable que se quiere definir.
- ▶ Introduzca los valores o el rango de valores que representen los datos perdidos.

Se considera que son válidos todos los valores de cadena, incluidos los valores vacíos o nulos, a no ser que se definan explícitamente como perdidos. Para definir como perdidos los valores nulos o vacíos de una variable de cadena, escriba un espacio en blanco en uno de los campos debajo de la selección Valores perdidos discretos.

Ancho de columna

Se puede especificar un número de caracteres para el ancho de la columna. Los anchos de columna también se pueden cambiar en la Vista de datos pulsando y arrastrando los bordes de las columnas.

Los formatos de columna afectan sólo a la presentación de valores en el Editor de datos. Al cambiar el ancho de columna no se cambia el ancho definido de una variable. Si el ancho real y definido de un valor es más ancho que la columna, aparecerán asteriscos (*) en la ventana Vista de datos.

Alineación de la variable

La alineación controla la presentación de los valores de los datos y/o de las etiquetas de valor en la Vista de datos. La alineación por defecto es a la derecha para las variables numéricas y a la izquierda para las variables de cadena. Este ajuste sólo afecta a la presentación en la Vista de datos.

Aplicación de atributos de definición de variables a varias variables

Una vez que se han definido los atributos de definición de variables correspondientes a una variable, puede copiar uno o más atributos y aplicarlos a una o más variables.

Se utilizan las operaciones básicas de copiar y pegar para aplicar atributos de definición de variables. Tiene la posibilidad de:

- Copiar un único atributo (por ejemplo, etiquetas de valor) y pegarlo en la misma casilla de atributo para una o más variables.
- Copiar todos los atributos de una variable y pegarlos en una o más variables.
- Crear varias variables nuevas con todos los atributos de una variable copiada.

Aplicación de atributos de definición de variables a otras variables

Para aplicar atributos individuales de una variable definida:

- ▶ En la Vista de variables, seleccione la casilla de atributos que quiere aplicar a otras variables.

- ▶ Elija en los menús:
 - Edición
 - Copiar
 - ▶ Seleccione la casilla de atributos a la que quiere aplicar el atributo. Puede seleccionar varias variables de destino.
 - ▶ Elija en los menús:
 - Edición
 - Pegar
- Si pega el atributo en filas vacías, se crean nuevas variables con atributos por defecto excepto para el atributo seleccionado.

Para aplicar todos los atributos de una variable definida:

- ▶ En la Vista de variables, seleccione el número de fila para la variable con los atributos que quiere utilizar. Se resaltará la fila entera.
- ▶ Elija en los menús:
 - Edición
 - Copiar
- ▶ Seleccione los números de fila de las variables a la que desea aplicar los atributos. Puede seleccionar varias variables de destino.
- ▶ Elija en los menús:
 - Edición
 - Pegar

Generación de varias variables nuevas con los mismos atributos

- ▶ En la Vista de variables, seleccione el número de fila para la variable con los atributos que quiere utilizar para la nueva variable. Se resaltará la fila entera.
- ▶ Elija en los menús:
 - Edición
 - Copiar

- ▶ Pulse en el número de la fila vacía situada bajo la última variable definida en el archivo de datos.
- ▶ Elija en los menús:
 - Edición
 - Pegar variables...
- ▶ Indique el número de variables que quiere crear.
- ▶ Introduzca un prefijo y un número inicial para las nuevas variables.

Los nombres de las nuevas variables se compondrán del prefijo especificado, más un número secuencial que comienza por el número indicado.

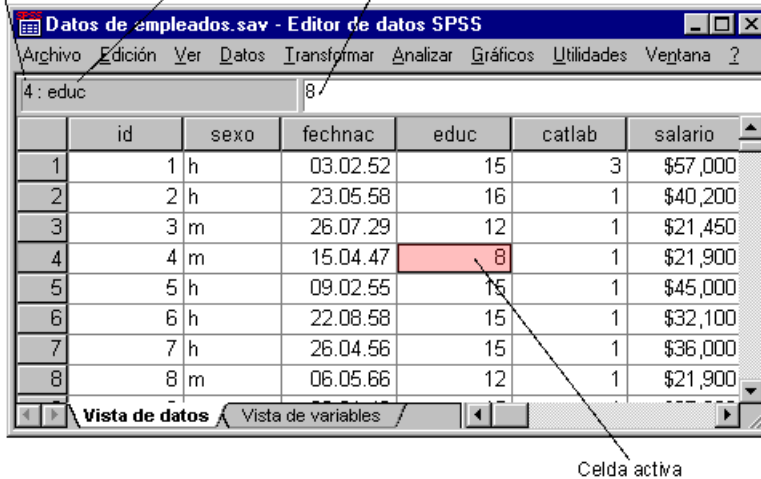
Introducción de datos

Se pueden introducir datos directamente en el Editor de datos de la Vista de datos. Se puede introducir datos en cualquier orden. Asimismo, se pueden introducir datos por caso o por variable, para áreas seleccionadas o para casillas individuales.

- Se resaltarán la casilla activa.
- El nombre de la variable y el número de fila de la casilla activa aparecen en la esquina superior izquierda del Editor de datos.
- Cuando seleccione una casilla e introduzca un valor de datos, el valor se muestra en el editor de casillas situado en la parte superior del Editor de datos.
- Los valores de datos no se registran hasta que se pulsa Intro o se selecciona otra casilla.
- Para introducir datos distintos de los numéricos, en primer lugar, se debe definir el tipo de variable.

Si introduce un valor en una columna vacía, el Editor de datos creará automáticamente una nueva variable y asignará un nombre de variable.

Figura 5-7
Archivo de datos de trabajo en la Vista de datos
 Numero de fila Nombre de variable Editor de casillas



Para introducir datos numéricos

- ▶ Seleccione una casilla en la Vista de datos.
- ▶ Introduzca el valor de los datos. El valor se muestra en el editor de casillas situado en la parte superior del Editor de datos.
- ▶ Pulse Intro o seleccione otra casilla para registrar el valor.

Para introducir datos no numéricos

- ▶ Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña Vista de variables.
- ▶ Pulse en el botón de la casilla *Tipo* de la variable.
- ▶ Seleccione el tipo de datos en el cuadro de diálogo Tipo de variable.
- ▶ Pulse en Aceptar.

- ▶ Pulse dos veces en el número de fila o pulse en la pestaña Vista de datos.
- ▶ Introduzca en la columna los datos de la variable que se va a definir.

Para utilizar etiquetas de valor en la introducción de datos

- ▶ Si las etiquetas de valor no aparecen en la Vista de datos, seleccione en los menús:
Ver
Etiquetas de valor
- ▶ Pulse en la casilla en la que quiere introducir el valor.
- ▶ Seleccione una etiqueta de valor en la lista desplegable.

De este modo se introducirá el valor y la etiqueta de valor se mostrará en la casilla.

Nota: Este método sólo funciona si ha definido etiquetas de valor para la variable.

Restricciones de los valores de datos en el Editor de datos

El ancho y el tipo de variable definidos determinan el tipo de valor que se puede introducir en la casilla en la Vista de datos.

- Si escribe un carácter no permitido por el tipo de variable definido, el Editor de datos emitirá una señal acústica y el carácter no se introducirá.
- Para variables de cadena, no se permiten los caracteres que sobrepasen el ancho definido.
- Para variables numéricas, se pueden introducir valores enteros que excedan el ancho definido, pero el Editor de datos mostrará la notación científica o asteriscos en la casilla para indicar que el valor es más ancho que el ancho definido. Para mostrar el valor de la casilla, cambie el ancho definido de la variable. (*Nota:* Cambiar el ancho de la columna no afecta al ancho de la variable.)

Edición de datos

Con el Editor de datos es posible modificar los valores de datos en la Vista de datos de muchas maneras. Tiene la posibilidad de:

- Cambiar los valores de datos.
- Cortar, copiar y pegar valores de datos.
- Añadir y eliminar casos.
- Añadir y eliminar variables.
- Cambiar el orden de las variables.

Para reemplazar o modificar un valor de datos

Para eliminar el valor anterior e introducir un valor nuevo:

- ▶ En la Vista de datos, pulse dos veces en la casilla. Su valor aparecerá en el editor de casillas.
- ▶ Edite el valor directamente en la casilla o en el editor de casillas.
- ▶ Pulse Intro (o desplácese a otra casilla) para registrar el valor nuevo.

Cortar, copiar y pegar valores de datos

Puede cortar, copiar y pegar valores de casillas individuales o grupos de valores en el Editor de datos. Tiene la posibilidad de:

- Mover o copiar un único valor de casilla a otra casilla.
- Mover o copiar un único valor de casilla en un grupo de casillas.
- Mover o copiar los valores de un único caso (fila) en varios casos.
- Mover o copiar los valores de una única variable (columna) en varias variables.
- Mover o copiar un grupo de valores de casillas en otro grupo de casillas.

Conversión de datos para valores pegados en el Editor de datos

Si los tipos de variable definidos de las casillas de origen y de destino no son iguales, el Editor de datos intentará convertir el valor. Si no es posible realizar la conversión, el valor perdido del sistema se insertará en la casilla de destino.

Numérico o fecha a cadena. Los formatos numéricos (por ejemplo, numérico, dólar, de punto o de coma) y de fechas se convierten en cadenas si se pegan en una casilla de variable de cadena. El valor de cadena es el valor numérico tal como se muestra en la casilla. Por ejemplo, para la variable con formato de dólar, el signo dólar que se muestra se convierte en parte del valor de cadena. Los valores que sobrepasan el ancho de la variable de cadena definida quedan cortados.

Cadena a numérico o fecha. Los valores de cadena que contienen caracteres admisibles por el formato numérico o de fecha de la casilla de destino se convierten al valor numérico o de fecha equivalente. Por ejemplo, un valor de cadena de 25/12/91 se convierte a una fecha válida si el tipo de formato de la casilla de destino es uno de los formatos día-mes-año, pero se convierte en perdido por el sistema si el tipo de formato de la casilla de destino es uno de los formatos mes-día-año.

Fecha a numérico. Los valores de fecha y hora se convierten a un número de segundos si la casilla de destino es uno de los formatos numéricos (por ejemplo, numérico, dólar, de punto o de coma). Como las fechas se almacenan internamente como el número de segundos transcurridos desde el 14 de octubre de 1582, la conversión de fechas a valores numéricos puede generar números extremadamente grandes. Por ejemplo, la fecha 10/29/91 se convierte al valor numérico 12.908.073.600.

Numérico a fecha u hora. Los valores numéricos se convierten a fechas u horas si el valor representa un número de segundos que puede producir una fecha u hora válidos. Para las fechas, los valores numéricos menores de 86.400 se convierten al valor perdido del sistema.

Inserción de nuevos casos

Al introducir datos en una casilla de una fila vacía, se crea automáticamente un nuevo caso. El Editor de datos inserta el valor perdido del sistema para el resto de las variables de dicho caso. Si hay alguna fila vacía entre el nuevo caso y los casos existentes, las filas en blanco también se convierten en casos nuevos con el valor

perdido del sistema para todas las variables. También puede insertar nuevos casos entre casos existentes.

Para insertar nuevos casos entre los casos existentes

- ▶ En la Vista de datos, seleccione cualquier casilla del caso (fila) debajo de la posición donde desea insertar el nuevo caso.

- ▶ Elija en los menús:

Datos
Insertar casos

Se inserta una fila nueva para el caso y todas las variables reciben el valor perdido del sistema.

Inserción de nuevas variables

La introducción de datos en una columna vacía en la Vista de datos o en una fila vacía en la Vista de variables crea de forma automática una variable nueva con un nombre de variable por defecto (el prefijo *var* y un número secuencial) y un tipo de formato de datos por defecto (numérico). El Editor de datos inserta el valor perdido del sistema en todos los casos de la nueva variable. Si hay columnas vacías en la Vista de datos o filas vacías en Vista de variables entre la nueva variable y las variables existentes, estas columnas también se convierten en nuevas variables con el valor perdido del sistema para todos los casos. También se pueden insertar variables nuevas entre las variables existentes.

Para insertar nuevas variables entre variables existentes

- ▶ Seleccione cualquier casilla de la variable a la derecha (Vista de datos) o debajo (Vista de variables) de la posición donde desea insertar la nueva variable.

- ▶ Elija en los menús:

Datos
Insertar variable

Se insertará una nueva variable con el valor perdido del sistema para todos los casos.

Para mover variables

- ▶ Pulse en el nombre de variable en la Vista de datos o en el número de fila de la variable en la Vista de variables para seleccionarla.
- ▶ Arrastre y suelte la variable en la nueva ubicación.
- ▶ Si desea situar la variable entre dos variables existentes, en la Vista de datos coloque la variable en la columna de variables a la derecha del lugar donde desea situar la variable. En la Vista de variables, colóquela en la fila de variables por debajo del lugar donde desea situarla.

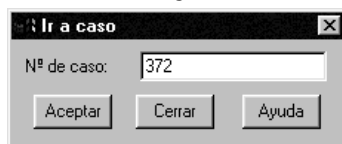
Para cambiar el tipo de datos

Puede cambiar el tipo de datos de una variable en cualquier momento utilizando el cuadro de diálogo Tipo de variable en la Vista de variables y el Editor de datos intentará convertir los valores existentes al nuevo tipo. Si no se puede realizar esta conversión, se asignará el valor perdido del sistema. Las reglas de conversión son las mismas que las del pegado de valores de datos en una variable con distinto tipo de formato. Si el cambio del formato de los datos puede generar la pérdida de las especificaciones de valores perdidos o de las etiquetas de valor, el Editor de datos mostrará un cuadro de alerta solicitando confirmación para proseguir o cancelar la operación.

Ir a caso

Ir a caso va al número de caso (fila) especificado en el Editor de datos.

Figura 5-8
Cuadro de diálogo Ir a caso



Para buscar un caso en el Editor de datos

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Elija en los menús:
Datos
Ir a caso...
- ▶ Introduzca el número de fila del Editor de datos para el caso.

Estado de selección de casos en el Editor de datos

Si ha seleccionado un subconjunto de casos pero no ha descartado los casos no seleccionados, éstos se marcarán en el Editor de datos con una línea diagonal (barra diagonal) atravesando el número de fila.

Figura 5-9
Casos filtrados en el Editor de datos

	id	sexo	fechnac	educ	catlab	salario
1	1	h	03.02.52	15	3	\$57,000
2	2	h	23.05.58	16	1	\$40,200
3	3	m	26.07.29	12	1	\$21,450
4	4	m	15.04.47	8	1	\$21,900
5	5	h	09.02.55	15	1	\$45,000
6	6	h	22.08.58	15	1	\$32,100
7	7	h	26.04.56	15	1	\$36,000
8	8	m	06.05.66	12	1	\$21,900
9	9	m	23.01.46	15	1	\$27,900

Editor de datos: Opciones de presentación

El menú Ver proporciona varias opciones de presentación para el Editor de datos:

Fuentes. Controla las características de fuentes de la presentación de datos.

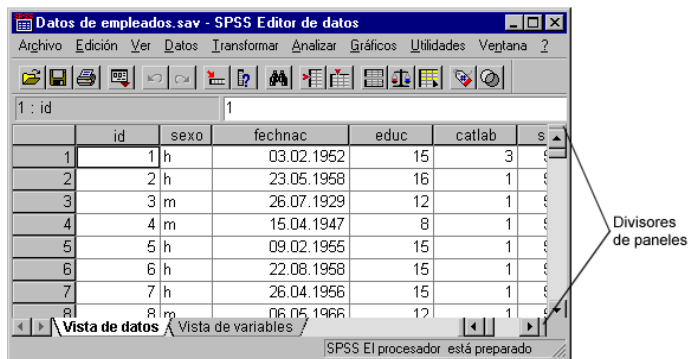
Líneas de cuadrícula. Activa y desactiva la presentación de las líneas de cuadrícula.

Etiquetas de valor. Activa y desactiva la presentación de los valores reales de los datos y las etiquetas de valor descriptivas definidas por el usuario. Esta opción sólo está disponible en la Vista de datos.

Uso de varias vistas

En la Vista de datos, puede crear varias vistas (paneles) mediante los divisores situados sobre la barra de desplazamiento horizontal y a la derecha de la barra de desplazamiento vertical.

Figura 5-10
Divisores de paneles de Vista de datos



También puede utilizar el menú Ventana para insertar y eliminar divisores de paneles. Para insertar divisores:

- ▶ En la Vista de datos, elija en los menús:
Ventana
Dividir

Los divisores se insertan sobre y a la izquierda de la casilla seleccionada.

- Si se ha seleccionado la casilla superior izquierda, los divisores se insertan para dividir la vista actual aproximadamente por la mitad horizontal y verticalmente.

- Si se selecciona una casilla distinta de la casilla superior de la primera columna, se inserta un divisor de paneles horizontales sobre la casilla seleccionada.
- Si se selecciona una casilla distinta de la primera casilla de fila superior, se inserta un divisor de paneles verticales a la izquierda de la casilla seleccionada.

Impresión en el Editor de datos

Los archivos de datos se imprimen tal y como aparece en la pantalla.

- Se imprime la información que está en la vista actualmente mostrada. En la Vista de datos, se imprimen los datos. En la Vista de variables, se imprime la información de definición de los datos.
- Las líneas de cuadrícula se imprimen si aparecen actualmente en la vista seleccionada.
- Las etiquetas de valor se imprimen si aparecen actualmente en la Vista de datos. En caso contrario, se imprimirán los valores de datos reales.

Utilice el menú Ver en la ventana Editor de datos para mostrar u ocultar las líneas de cuadrícula y para que se muestren o no los valores de los datos y las etiquetas de valor.

Impresión del contenido del Editor de datos

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Seleccione la pestaña de la vista que desea imprimir.
- ▶ Elija en los menús:
 - Archivo
 - Imprimir...

Preparación de datos

Cuando abra un archivo de datos o introduzca datos en el Editor de datos, podrá empezar a crear informes, gráficos y análisis sin ningún trabajo preliminar adicional. Sin embargo, hay algunas funciones de preparación adicional de los datos que pueden resultarle útiles, entre las que se incluyen:

- Asignar propiedades de las variables que describan los datos y determinen cómo se deben tratar ciertos valores.
- Identificar los casos que pueden contener información duplicada y excluir dichos casos de los análisis o eliminarlos del archivo de datos.
- Crear nuevas variables con algunas categorías distintas que representen rangos de valores de variables que tengan un mayor número de valores posibles.

Propiedades de variables

Los datos introducidos simplemente en la Vista de datos del Editor de datos o leídos en SPSS desde un formato de archivos externo (por ejemplo, una hoja de cálculo de Excel o un archivo de datos de texto) carecen de ciertas propiedades de variables que pueden resultar muy útiles, como:

- Definición de etiquetas de valor descriptivas para códigos numéricos (por ejemplo, 0 = *Hombre* y 1 = *Mujer*).
- Identificación de códigos de valores perdidos (por ejemplo, 99 = *No procede*).
- Asignación del nivel de medida (nominal, ordinal o de escala).

Todas estas propiedades de variables (y otras) se pueden asignar en la Vista de variables del Editor de datos. También hay algunas utilidades que le pueden ofrecer asistencia en este proceso:

- **Definir propiedades de variables** puede ayudarle a definir etiquetas de valor descriptivas y valores perdidos. Esto es especialmente útil para datos categóricos con códigos numéricos utilizados para valores de categorías.
- **Copiar propiedades de datos** ofrece la posibilidad de utilizar un archivo de datos con formato SPSS como plantilla para definir las propiedades de variables y archivos en el archivo de datos actual. Esto es particularmente útil si utiliza frecuentemente archivos de datos con un formato externo que tenga un contenido similar, como puedan ser informes mensuales en formato Excel.

Definición de propiedades de variables

Definir propiedades de variables está diseñado para ayudarle en el proceso de creación de etiquetas de valor descriptivas para variables categóricas (nominales u ordinales).

Definir propiedades de variables:

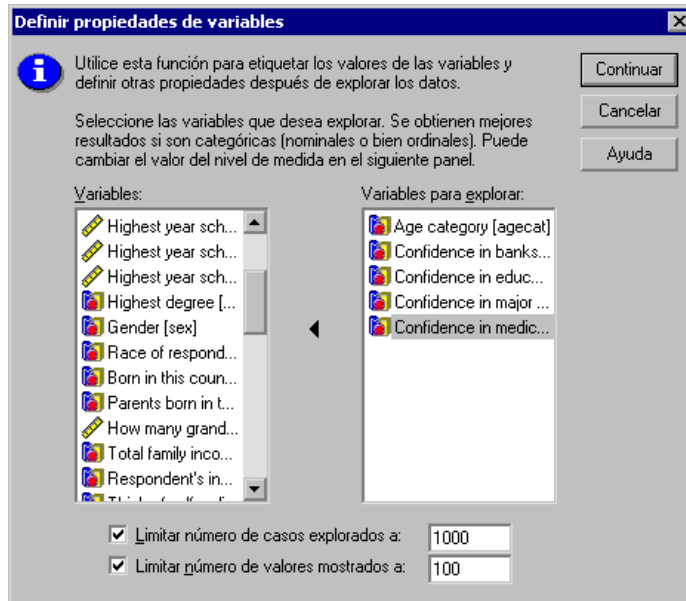
- Explora los datos reales y enumera todos valores de datos únicos para cada variable seleccionada.
- Identifica valores sin etiquetas y ofrece una función de “etiquetas automáticas”.
- Permite copiar etiquetas de valor definidas de otra variable en la variable seleccionada o de la variable seleccionada a varias variables adicionales.

Nota: Para utilizar Definir propiedades de variables sin explorar primero los casos, introduzca 0 para el número de casos que se van a explorar.

Para definir propiedades de variables

- ▶ Elija en los menús:
 - Datos
 - Definir propiedades de variables...

Figura 6-1
Cuadro de diálogo inicial para seleccionar las variables que se van a definir



- Seleccione las variables numéricas o de cadena corta para las que desea crear etiquetas de valor o definir o cambiar otras propiedades de las variables, como los valores perdidos o las etiquetas de variable descriptivas.

Nota: Las variables de cadena largas (variables de cadena con una anchura definida de más de ocho caracteres) no se visualizan en la lista de variables. Las variables de cadena largas no pueden tener etiquetas de valor definidas ni categorías de valores perdidos.

- Especificar el número de casos que se van a explorar para generar la lista de valores únicos. resulta especialmente útil para los archivos de datos con mayor número de casos, para el cual una exploración del archivo de datos completo podría tardar una gran cantidad de tiempo.
- Especifique un límite superior para el número de valores únicos que se va a visualizar. Esto es especialmente útil para evitar que se generen listas de cientos, miles o incluso millones de valores para las variables de escala (intervalo continuo, razón).

- ▶ Pulse en Continuar para abrir el cuadro de diálogo Definir propiedades de variables principal.
- ▶ Seleccione una variable para la que desee crear etiquetas de valor o definir o cambiar otras propiedades de las variables.
- ▶ Introduzca el texto de etiqueta para los valores sin etiquetas que se visualicen en Rejilla etiqueta valores.
- ▶ Si hay valores para los que desea crear etiquetas de valor, pero no se visualizan dichos valores, puede introducirlos en la columna *Valores* por debajo del último valor explorado.
- ▶ Repita este proceso para cada variable de la lista para la que desee crear etiquetas de valor.
- ▶ Pulse en Aceptar para aplicar las etiquetas de valor y otras propiedades de las variables.

Definición de etiquetas de valor y otras propiedades de las variables

Figura 6-2

Cuadro de diálogo principal Definir propiedades de variables

Definir propiedades de las variables

Lista de variables exploradas: agecat, confinan, coneduc, conbus, conmedic

Variable actual: agecat Etiqueta: Age category

Nivel de medida: Ordinal Sugerir

Valores sin etiqueta: 0 Tipo: Numérico Ancho: 8 Decimales: 2

Cuadrícula eti. va Añada etiquetas a la cuadrícula o editelas. Puede añadir valores abajo.

	Cambiado	Perdido	Recuento	Valor	Etiqueta
1	<input type="checkbox"/>	<input type="checkbox"/>	67	1,00	Less than 25
2	<input type="checkbox"/>	<input type="checkbox"/>	221	2,00	25 to 34
3	<input type="checkbox"/>	<input type="checkbox"/>	233	3,00	35 to 44
4	<input type="checkbox"/>	<input type="checkbox"/>	171	4,00	45 to 54
5	<input type="checkbox"/>	<input type="checkbox"/>	131	5,00	55 to 64
6	<input type="checkbox"/>	<input type="checkbox"/>	174	6,00	65 or older
7	<input type="checkbox"/>	<input checked="" type="checkbox"/>	0	98,00	DK
8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	0	99,00	NA
9	<input type="checkbox"/>	<input type="checkbox"/>			

Copiar propiedades: De otra variable... A otras variables... Valores sin etiquetas: Etiquetas automáticas

Casos explor.: 1000 Lím. lista valores: 100

Aceptar Pegar Restabl. Cancelar Ayuda

El cuadro de diálogo principal Definir propiedades de variables proporciona la siguiente información para las variables exploradas:

Lista de variables exploradas. Para cada variable explorada, aparecerá una marca de verificación en la columna *Sin etiqueta* indicando que la variable contiene valores sin etiquetas de valor asignadas.

Para ordenar la lista de variables para que aparezcan todas las variables con valores sin etiquetas en la parte superior de la lista:

- ▶ Pulse en el encabezado de columna *Sin etiqueta* debajo de la Lista de variables exploradas.

También puede ordenarla por nombre de variable o nivel de medida pulsando en el encabezado de columna correspondiente debajo de la Lista de variables exploradas.

Rejilla etiqueta valores

- **Etiqueta.** Muestra las etiquetas de valor que ya se han definido. Puede añadir o cambiar las etiquetas de esta columna.
- **Valor.** Valores únicos para cada variable seleccionada. Esta lista de valores únicos se basa en el número de casos explorados. Por ejemplo, si sólo ha explorado los primeros 100 casos del archivo de datos, la lista reflejará sólo los valores únicos presentes en esos casos. Si el archivo de datos ya se ha ordenado por la variable para la que desea asignar etiquetas de valor, la lista puede mostrar muchos menos valores únicos de los que hay realmente presentes en los datos.
- **Recuento.** Número de veces que aparece cada valor en los casos explorados.
- **Perdidos.** Valores definidos para representar valores perdidos. Puede cambiar la designación de la categoría de los valores perdidos pulsando en la casilla de verificación. Una marca indica que la categoría se ha definido como categoría perdida por el usuario. Si una variable ya tiene un rango de valores definidos como perdidos por el usuario (por ejemplo 90 - 99), no podrá añadir ni eliminar categorías de valores perdidos para esa variable con Definir propiedades de variables. Puede utilizar la Vista de variables del Editor de datos para modificar las categorías de valores perdidos para la variable con rangos de valores perdidos. Si desea obtener más información, consulte “Valores perdidos” en Capítulo 5 en p. 93.
- **Cambiado.** Indica que ha añadido o cambiado una etiqueta de valor.

Nota: Si ha especificado 0 para el número de casos que se van a explorar en el cuadro de diálogo inicial, la rejilla etiqueta valores estará en blanco al principio, a excepción de algunas etiquetas de valor ya existentes y/o categorías de valores perdidos definidas para la variable seleccionada. Además, se desactivará el botón Sugerir para el nivel de medida.

Nivel de medida. Las etiquetas de valor son especialmente útiles para las variables categóricas (nominales u ordinales), y algunos procedimientos tratan a las variables categóricas y de escala de manera diferente, por lo que a veces es importante asignar el nivel de medida correcto. Sin embargo, por defecto, todas las nuevas variables numéricas se asignan al nivel de medida de escala. Por tanto, puede que muchas variables que son de hecho categóricas, aparezcan inicialmente como variables de escala.

Si no está seguro de qué nivel de medida debe asignar a una variable, pulse en Sugerir.

Copiar propiedades. Puede copiar las etiquetas de valor y otras propiedades de las variables de otra variable a la variable seleccionada en ese momento o desde la variable seleccionada en ese momento a una o varias otras variables.

Valores sin etiquetas. Para crear automáticamente etiquetas para valores sin etiquetas, pulse en Etiquetas automáticas.

Etiqueta de variable y formato de presentación

Puede cambiar de la etiqueta de variable descriptiva y el formato de presentación.

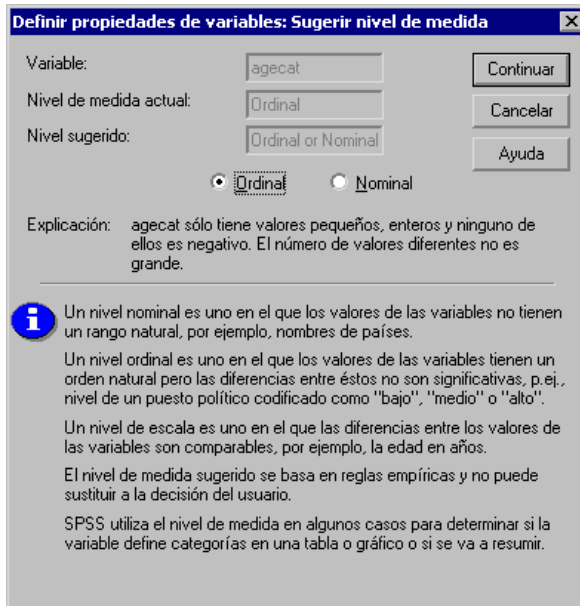
- No puede cambiar el tipo fundamental de la variable (numérica o de cadena).
- Para las variables de cadena, sólo puede cambiar la etiqueta de variable, no el formato de presentación.
- Para las variables numéricas, puede cambiar el tipo numérico (como numérico, fecha, dólar o moneda personalizada), el ancho (número máximo de dígitos, incluyendo los indicadores decimales y/o de agrupación) y el número de posiciones decimales.
- Para el formato de fecha numérica, puede seleccionar un formato de fecha específico (como dd-mm-aaaa, mm/dd/aa, aaaaddd, etc.)
- Para formato numérico personalizado, puede seleccionar uno de los cinco formatos de moneda personalizados (de CCA a CCE). Si desea obtener más información, consulte “Moneda: Opciones” en Capítulo 43 en p. 631.

- Aparece un asterisco en la columna *Valor* si el ancho especificado es inferior al ancho de los valores explorados o los valores mostrados para etiquetas de valor definidas ya existentes o categorías de valores perdidos.
- Aparece un período (.) si los valores explorados o los valores mostrados para etiquetas de valor definidas ya existentes o categorías de valores perdidos no son válidos para el tipo de formato de presentación seleccionado. Por ejemplo, un valor numérico interno inferior a 86.400 no es válido para una variable de formato de fecha.

Asignación del nivel de medida

Cuando pulse en Sugerir para seleccionar un nivel de medida en el cuadro de diálogo principal Definir propiedades de variables, la variable actual se evalúa en función de los casos explorados y las etiquetas de valor definidas y se sugiere un nivel de medida en el cuadro de diálogo Sugerir nivel de medida que se abre. El área Explicación ofrece una breve descripción de los criterios utilizados para proporcionar el nivel de medida sugerido.

Figura 6-3
Cuadro de diálogo Sugerir nivel de medida



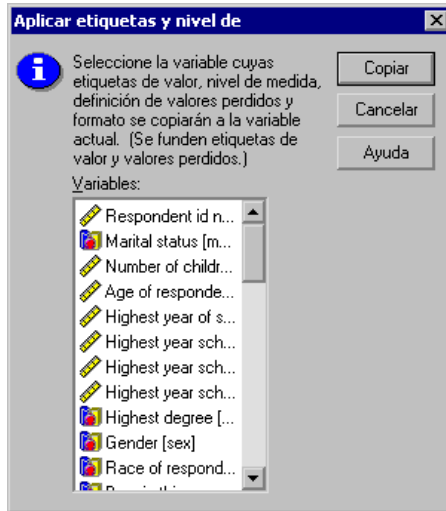
Nota: Los valores definidos para representar valores perdidos no se incluyen en la evaluación para el nivel de medida. Por ejemplo, la explicación del nivel de medida sugerido puede indicar que la sugerencia se basa, en parte, en el hecho de que la variable no contiene valores negativos, mientras que, de hecho, puede contener valores negativos, pero dichos valores ya se han definido como valores perdidos.

- Pulse en Continuar para aceptar el nivel de medida sugerido o en Cancelar para mantener el mismo.

Copia de propiedades de variables

El cuadro de diálogo Aplicar etiquetas y nivel a aparece al pulsar en De otra variable o A otras variables en el cuadro de diálogo principal Definir propiedades de variables. Muestra todas las variables exploradas que coinciden con el tipo de variable actual (de cadena o numérico). Para las variables de cadena, también debe coincidir la anchura definida.

Figura 6-4
Cuadro de diálogo *Aplicar etiquetas y medida a*



- ▶ Seleccione una única variable desde la que va a copiar las etiquetas de valor y otras propiedades de las variables (excepto la etiqueta de la variable).
 - o
- ▶ Seleccione una o más variables a las que va a copiar las etiquetas de valor y otras propiedades de las variables.
- ▶ Pulse en Copiar para copiar las etiquetas de valor y el nivel de medida.
 - Las etiquetas de valor existentes y categorías de valores perdidos para las variables de destino no se sustituyen.
 - Las etiquetas de valor y las categorías de valores perdidos para los valores que no se han definido aún para las variables de destino se añaden al conjunto de etiquetas de valor y categorías de valores perdidos para las variables de destino.
 - El nivel de medida para las variables de destino siempre se sustituye.
 - Si la variable de origen o de destino tiene un rango definido de valores perdidos, no se copian las definiciones de los valores perdidos.

Copia de propiedades de datos

El Asistente para la copia de propiedades de datos ofrece la posibilidad de utilizar un archivo de datos de SPSS como plantilla para definir las propiedades de las variables y del archivo del archivo de datos de trabajo. También puede utilizar las variables del archivo de datos de trabajo como plantillas para el resto de las variables del archivo de datos de trabajo. Tiene la posibilidad de:

- Copiar las propiedades de archivo seleccionadas de un archivo de datos externo en el archivo de datos de trabajo. Las propiedades de archivo incluyen documentos, etiquetas de archivos, conjuntos de respuestas múltiples, conjuntos de variables y ponderación.
- Copiar las propiedades de variable seleccionadas de un archivo de datos externo en las variables coincidentes del archivo de datos de trabajo. Las propiedades de variable incluyen etiquetas de valor, valores perdidos, nivel de medida, etiquetas de variable, formatos de impresión y escritura, alineación y ancho de columna (en el Editor de datos).
- Copiar las propiedades de variable seleccionadas de una variable, ya sea del archivo de datos externo o del archivo de datos de trabajo, en diversas variables del archivo de datos de trabajo.
- Crear nuevas variables en el archivo de datos de trabajo basándose en las variables seleccionadas del archivo de datos externo.

Al copiar las propiedades de datos, se aplicarán las reglas siguientes:

- Para copiar un archivo de datos externo como archivo de datos de origen, deberá tratarse de un archivo de datos con formato SPSS.
- Para utilizar el archivo de datos de trabajo como archivo de datos de origen, deberá contener al menos una variable. No podrá utilizar un archivo de datos de trabajo que esté completamente en blanco como archivo de datos de origen.
- Las propiedades no definidas (vacías) del archivo de datos de origen no sobrescribirán las propiedades definidas en el archivo de datos de trabajo.
- Las propiedades de variable se copian desde la variable de origen únicamente a las variables de destino de un tipo coincidente: de cadena (alfanuméricas) o numérico (incluidas numéricas, fecha y moneda).

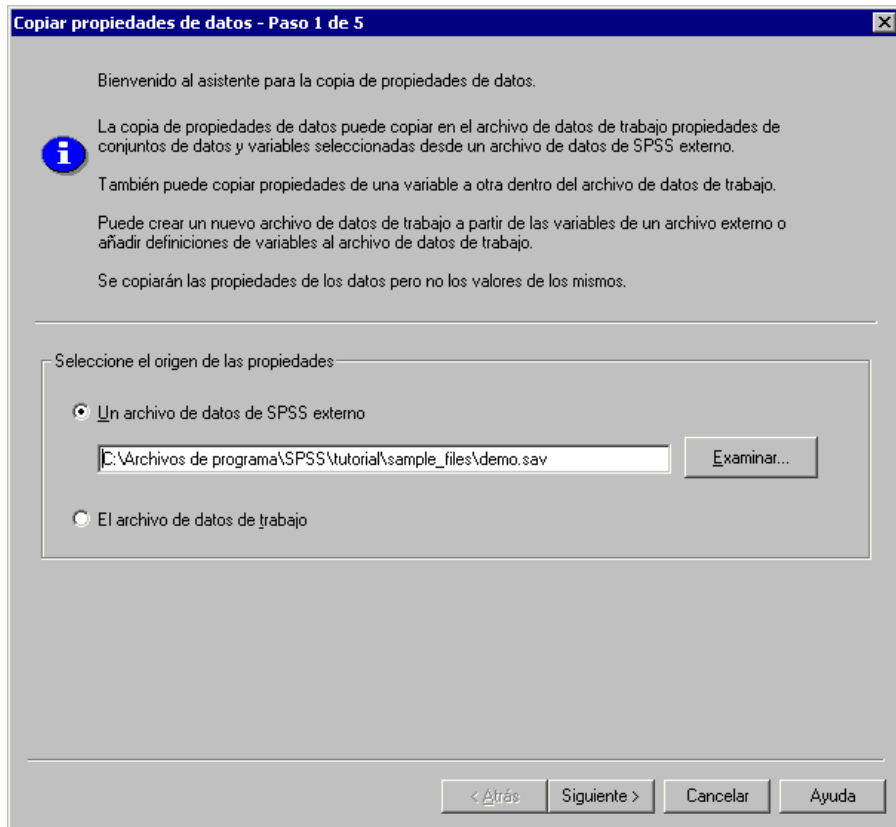
Nota: En el menú Archivo, Copiar propiedades de datos sustituirá a Aplicar diccionario de datos, disponible anteriormente.

Para copiar propiedades de datos

- ▶ Seleccione en los menús de la ventana Editor de datos:
 - Datos
 - Copiar propiedades de datos...

Figura 6-5

Asistente para la copia de propiedades de datos: Paso 1



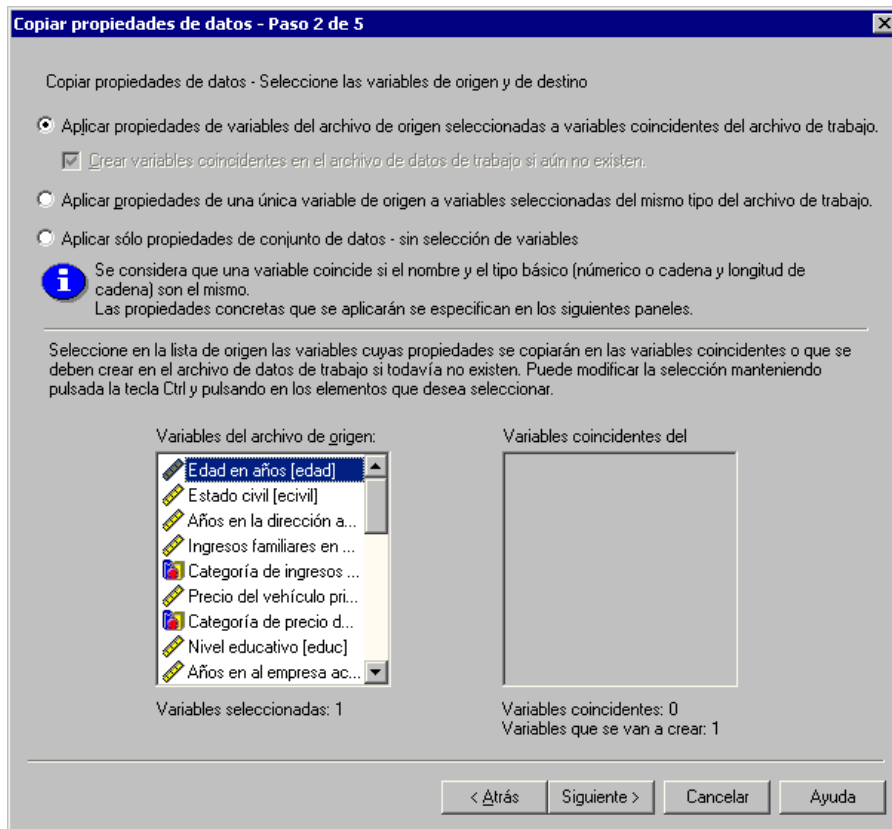
- ▶ Seleccione el archivo de datos que contenga las propiedades de archivo y/o variable que desee copiar. Se puede tratar de un archivo de datos externo con formato SPSS o del archivo de datos de trabajo.
- ▶ Siga las instrucciones detalladas del Asistente para la copia de propiedades de datos.

Selección de las variables de origen y de destino

En este paso, puede especificar tanto las variables de origen que contienen las propiedades de variable que desea copiar como las variables de destino en las que se copiarán estas propiedades de variable.

Figura 6-6

Asistente para la copia de propiedades de datos: Paso 2



Aplicar propiedades de variables del archivo de origen seleccionadas a variables coincidentes del archivo de trabajo. Las propiedades de variable se copian desde una o varias variables de origen seleccionadas en las variables coincidentes del archivo de datos de trabajo. Las variables “coinciden” si el nombre y el tipo de variable (de cadena o numérico) son los mismos. En el caso de las variables de cadena, la

longitud también debe ser la misma. Por defecto, sólo se muestran en las dos listas de variables las variables coincidentes.

■ **Crear variables coincidentes en el archivo de datos de trabajo si aún no existen.**

Actualiza la lista de origen para que muestre todas las variables del archivo de datos de origen. Si se seleccionan variables de origen que no existen en el archivo de datos de trabajo (basándose en el nombre de variable), se crearán nuevas variables en este archivo con los nombres y las propiedades de variable del archivo de datos de origen.

Si el archivo de datos de trabajo no contiene variables (un nuevo archivo de datos en blanco), se mostrarán todas las variables del archivo de datos de origen y se crearán automáticamente en el archivo de datos de trabajo nuevas variables basadas en las variables de origen seleccionadas.

Aplicar propiedades de una única variable de origen a variables seleccionadas del mismo tipo del archivo de trabajo. Las propiedades de variable de una única variable seleccionada se pueden aplicar a una o varias variables seleccionadas de la lista de archivos de trabajo. En esta lista sólo se mostrarán las variables que sean del mismo tipo (numérico o de cadena) que la variable seleccionada en la lista de origen. Si se trata de variables de cadena, sólo se mostrarán las cadenas con la misma longitud definida que la variable de origen. Esta opción no está disponible si el archivo de datos de trabajo no contiene variables.

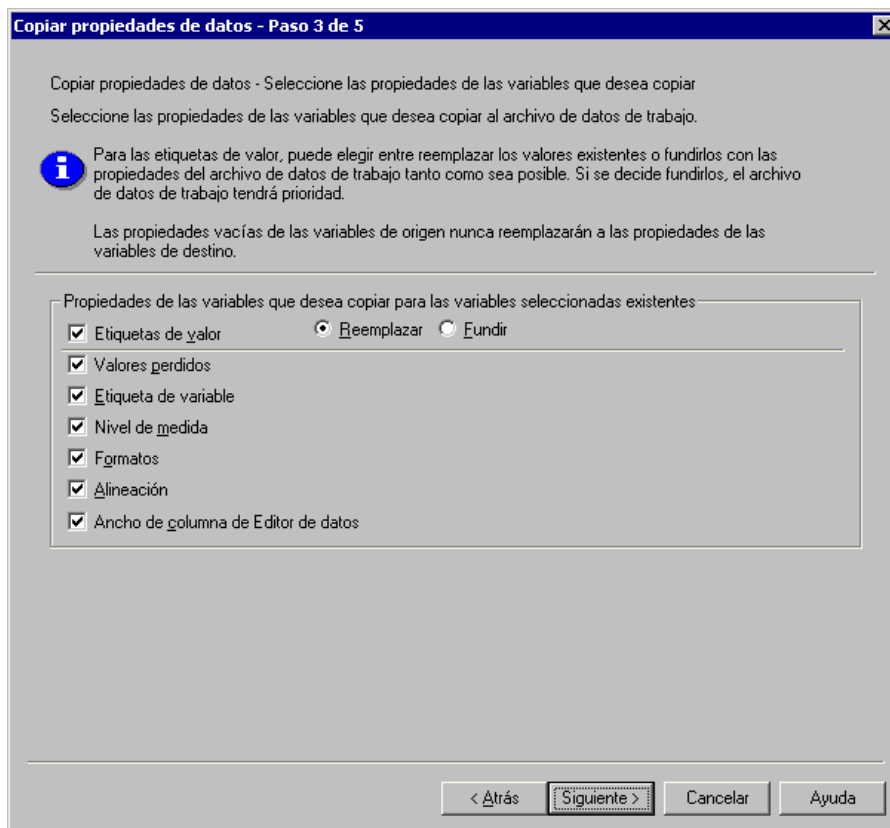
Nota: Mediante esta opción no puede crear nuevas variables en el archivo de datos de trabajo.

Aplicar sólo propiedades de conjunto de datos (sin selección de variables). Sólo se pueden aplicar al archivo de datos de trabajo las propiedades de archivo (por ejemplo, documentos, etiquetas de archivo, grosor). No se podrá aplicar ninguna propiedad de variable. Esta opción no está disponible si el archivo de datos de trabajo es al mismo tiempo el archivo de datos de origen.

Selección de propiedades de variable para copiar

Desde las variables de origen, las propiedades de variable seleccionadas no se pueden copiar en las variables de destino. Las propiedades no definidas (vacías) de las variables de origen no sobrescriben las propiedades definidas en las variables de destino.

Figura 6-7
Asistente para la copia de propiedades de datos: Paso 3



Etiquetas de valor. Las etiquetas de valor son etiquetas descriptivas asociadas a valores de datos. Se suelen utilizar cuando se seleccionan valores de datos numéricos para representar categorías no numéricas (por ejemplo, códigos 1 y 2 para *Hombre* y *Mujer*). Puede reemplazar o fundir las etiquetas de valor en las variables de destino.

- Reemplazar elimina todas las etiquetas de valor definidas para la variable de destino y las reemplaza por las etiquetas de valor definidas en la variable de origen.
- Fundir funde las etiquetas de valor definidas en la variable de origen con cualquier etiqueta de valor definida existente en la variable de destino. Si existe una etiqueta de valor definida con el mismo valor tanto en la variable de origen como en la de destino, la etiqueta de valor de la variable de destino permanecerá inalterada.

Valores perdidos. Los valores perdidos son valores identificados como representantes de datos perdidos (por ejemplo, 98 para *No se conoce* y 99 para *No procede*). Por lo general, estos valores tienen también etiquetas de valor definidas que describen el significado de códigos de valores perdidos. Todos los valores perdidos existentes definidos para la variable de destino se eliminarán y se reemplazarán por los valores perdidos de la variable de origen.

Etiqueta de variable. Las etiquetas de variable descriptivas pueden contener espacios y caracteres reservados que no se permiten en los nombres de las variables. Si desea utilizar esta opción para copiar propiedades de variable desde una variable de origen en varias variables de destino, reflexione antes de hacerlo.

Nivel de medida. El nivel de medida puede ser nominal, ordinal o de escala. En los procedimientos que diferencian entre los distintos niveles de medida, tanto el nivel nominal como el ordinal se consideran **categoricos**.

Formatos. Controla el tipo numérico (como numérico, fecha o moneda), el ancho (número total de caracteres que se muestran, incluidos los caracteres iniciales y finales y el indicador decimal) y el número de decimales que se van a mostrar para las variables numéricas. Esta opción no se tendrá en cuenta para las variables de cadena.

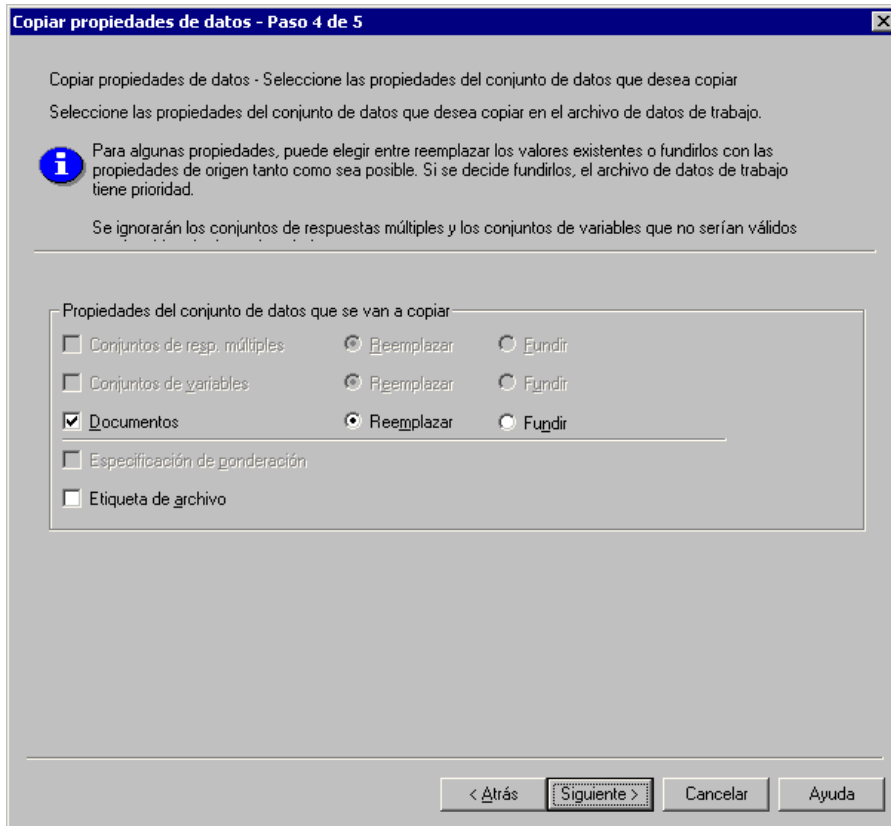
Alineación. Afecta únicamente a la alineación (izquierda, derecha, central) del Editor de datos de la Vista de datos.

Ancho de columna del Editor de datos. Afecta únicamente al ancho de columna de la Vista de datos del Editor de datos.

Copia de propiedades (de archivo) de conjunto de datos

Las propiedades de conjunto de datos globales seleccionadas del archivo de datos de origen se pueden aplicar al archivo de datos de trabajo. (Esta opción no está disponible si el archivo de datos de trabajo es el archivo de datos de origen.)

Figura 6-8
Asistente para la copia de propiedades de datos: Paso 4



Conjuntos resp. múltiples. Aplica múltiples definiciones del conjunto de respuestas del archivo de datos de origen al archivo de datos de trabajo. (*Nota:* En este momento, únicamente el componente adicional de tablas utiliza los conjuntos de respuestas múltiples.)

- Se ignorarán los conjuntos de respuestas múltiples del archivo de datos de origen que contengan variables no existentes en el archivo de datos de trabajo, a menos que se creen estas variables basándose en las especificaciones del paso 2 (Selección de las variables de origen y de destino) del Asistente para la copia de propiedades de datos.

- Reemplazar elimina todos los conjuntos de respuestas múltiples del archivo de datos de trabajo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir añade conjuntos de respuestas múltiples del archivo de datos de origen a la colección de este tipo de conjuntos incluida en el archivo de datos de trabajo. En caso de que exista un conjunto con el mismo nombre en ambos archivos, el conjunto existente en el archivo de datos de trabajo permanecerá inalterado.

Conjuntos de variables. Los conjuntos de variables se utilizan para controlar la lista de variables que se muestra en los cuadros de diálogo. Para definir conjuntos de variables, seleccione Definir conjuntos en el menú Utilidades.

- Se ignorarán los conjuntos del archivo de datos de origen que contengan variables no existentes en el archivo de datos de trabajo, a menos que se creen estas variables basándose en las especificaciones del paso 2 (Selección de las variables de origen y de destino) del Asistente para la copia de propiedades de datos.
- Reemplazar elimina todos los conjuntos de variables existentes en el archivo de datos de trabajo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir añade conjuntos de variables del archivo de datos de origen a la colección de este tipo de conjuntos incluida en el archivo de datos de trabajo. En caso de que exista un conjunto con el mismo nombre en ambos archivos, el conjunto existente en el archivo de datos de trabajo permanecerá inalterado.

Documentos. Notas añadidas al archivo de datos a través del comando `DOCUMENT`.

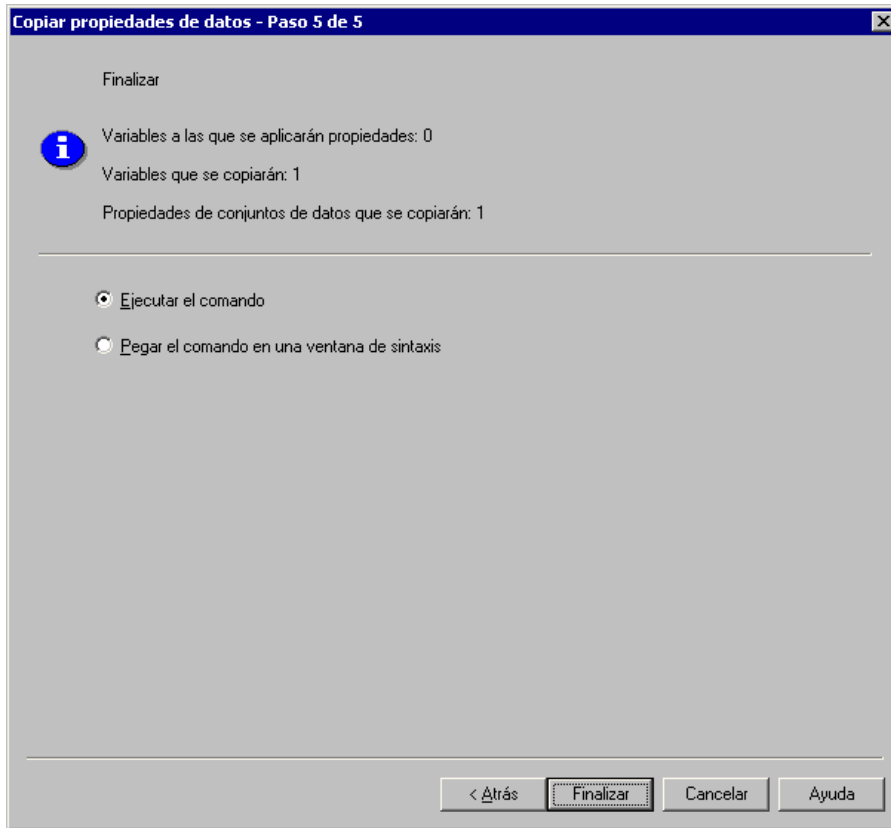
- Reemplazar elimina todos los documentos existentes en el archivo de datos de trabajo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir combina documentos incluidos en los archivos de datos de origen y de trabajo. Los documentos exclusivos del archivo de origen que no existan en el archivo de datos de trabajo se añadirán al archivo de datos de trabajo. A continuación, todos los documentos se ordenarán por fecha.

Especificación de ponderación. Pondera los casos por la variable de ponderación actual del archivo de datos de origen, siempre que exista una variable coincidente en el archivo de datos de trabajo. Sobrescribe cualquier ponderación activada actualmente en el archivo de datos de trabajo.

Etiqueta de archivo. Etiqueta descriptiva que se aplica a un archivo de datos mediante el comando `FILE LABEL`.

Resultados

Figura 6-9
Asistente para la copia de propiedades de datos: Paso 5



El último paso del Asistente para la copia de propiedades de datos proporciona información sobre el número de variables para las que se van a copiar las propiedades de variable del archivo de datos de origen, el número de nuevas variables que se van a crear y el número de propiedades (de archivo) de conjunto de datos que se van a copiar.

También puede pegar la sintaxis de comandos generada en una ventana de sintaxis y guardarla para su posterior uso.

Identificación de casos duplicados

Puede haber distintos motivos por los que haya casos “duplicados” en los datos, entre ellos:

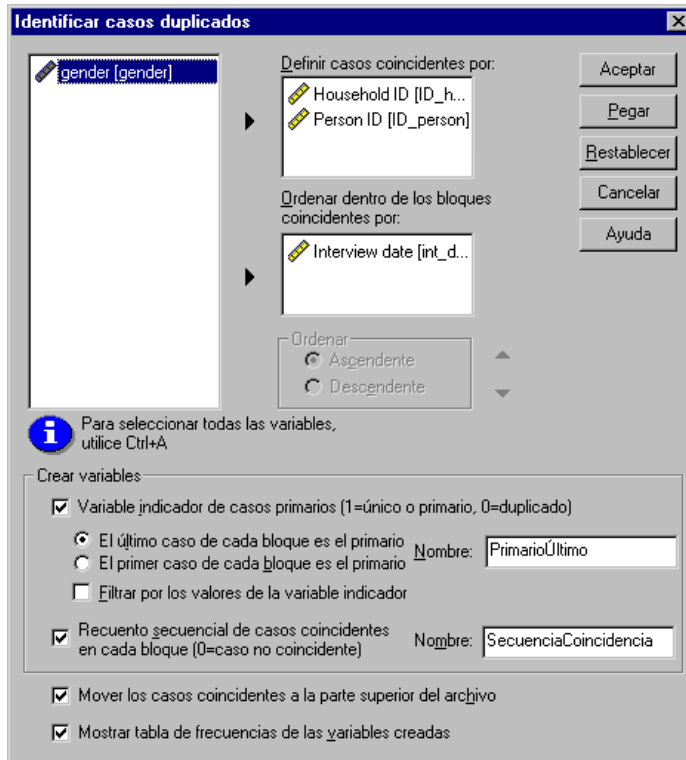
- Errores en la entrada de datos si por accidente se introduce el mismo caso más de una vez.
- Casos múltiples que comparten un valor de identificador primario común pero tienen valores diferentes de un identificador secundario, como los miembros de una familia que viven en el mismo domicilio.
- Casos múltiples que representan el mismo caso pero con valores diferentes para variables que no sean las que identifican el caso, como en el caso de varias compras realizadas por la misma persona o empresa de diferentes productos o en diferentes momentos.

La identificación de los casos duplicados le permite definir prácticamente como quiera lo que se considera *duplicado* y le proporciona cierto control sobre la determinación automática de los casos primarios frente a los duplicados.

Para identificar y marcar los casos duplicados:

- ▶ Elija en los menús:
 - Datos
 - Identificar casos duplicados...
- ▶ Seleccione una o varias variables que identifiquen los casos coincidentes.
- ▶ Seleccione una o varias de las opciones del grupo Crear variables.
Si lo desea, tiene la posibilidad de:
 - ▶ Seleccionar una o varias variables para ordenar los casos dentro de los bloques definidos por las variables seleccionadas de casos coincidentes. El orden definido por estas variables determina el “primer” y el “último” caso de cada bloque. En caso contrario, se utilizará el orden del archivo original.
 - ▶ Filtrar automáticamente los casos duplicados de manera que no se incluyan en los informes, los gráficos o el cálculo de estadísticos.

Figura 6-10
Cuadro de diálogo Identificar casos duplicados



Definir casos coincidentes por. Los casos se consideran duplicados si sus valores coinciden para *todas* las variables seleccionadas. Si desea identificar únicamente aquellos casos que coincidan al 100% en todos los aspectos, seleccione todas las variables.

Ordenar dentro de los bloques coincidentes por. Los casos se ordenan automáticamente por las variables que definen los casos coincidentes. Puede seleccionar otras variables de ordenación que determinarán el orden secuencial de los casos en cada bloque de coincidencia.

- Para cada variable de ordenación, el orden puede ser ascendente o descendente.

- Si selecciona más de una variable de ordenación, los casos se ordenarán por cada variable dentro de las categorías de la variable anterior de la lista. Por ejemplo, si selecciona *fecha* como la primera variable de ordenación y *cantidad* como la segunda, los casos se ordenarán por cantidad dentro de cada fecha.
- Utilice los botones de flecha hacia arriba y hacia abajo que hay a la derecha de la lista para cambiar el orden de las variables.
- El orden determina el “primer” y el “último” caso de cada bloque de coincidencia, que determina el valor de la variable indicador del caso primario opcional. Por ejemplo, si desea descartar todos los casos salvo el más reciente de cada bloque de coincidencia, puede ordenar los casos del bloque en orden ascendente por una variable de fecha, lo cual haría que la fecha más reciente fuese la última fecha del bloque.

Variable indicador de casos primarios. Crea una variable con un valor de 1 para todos los casos únicos y para el caso identificado como caso primario en cada bloque de casos coincidentes y un valor de 0 para los duplicados no primarios de cada bloque.

- El caso primario puede ser el primer o el último caso de cada bloque de coincidencia, según determine el orden del bloque de coincidencia. Si no especifica ninguna variable de ordenación, el orden del archivo original determina el orden de los casos dentro de cada bloque.
- Puede utilizar la variable indicador como una **variable de filtro** para excluir los duplicados que no sean primarios de los informes y los análisis sin eliminar dichos casos del archivo de datos.

Recuento secuencial de casos coincidentes en cada bloque. Crea una variable con un valor secuencial de 1 a n para los casos de cada bloque de coincidencia. La secuencia se basa en el orden actual de los casos de cada bloque, que puede ser el orden del archivo original o el orden determinado por las variables de ordenación especificadas.

Mover los casos coincidentes a la parte superior del archivo. Ordena el archivo de datos de manera que todos los bloques de casos coincidentes estén en la parte superior del archivo de datos, facilitando la inspección visual de los casos coincidentes en el Editor de datos.

Mostrar tabla de frecuencias de las variables creadas. Las tablas de frecuencias contienen los recuentos de cada valor de las variables creadas. Por ejemplo, para la variable de indicador de caso primario, la tabla mostraría tanto el número de casos

con un valor de 0 en esa variable, que indica el número de duplicados, como el número de casos con un valor de 1 para esa variable, que indica el número de casos únicos y primarios.

Valores perdidos. En el caso de variables numéricas, los valores perdidos del sistema se tratan como cualquier otro valor: los casos que tengan el valor perdido del sistema para una variable de identificación se tratarán como si tuviesen valores coincidentes para dicha variable. En el caso de variables de cadena, los casos que no tengan ningún valor para una variable de identificación se tratarán como si tuviesen valores coincidentes para dicha variable.

Categorizador visual

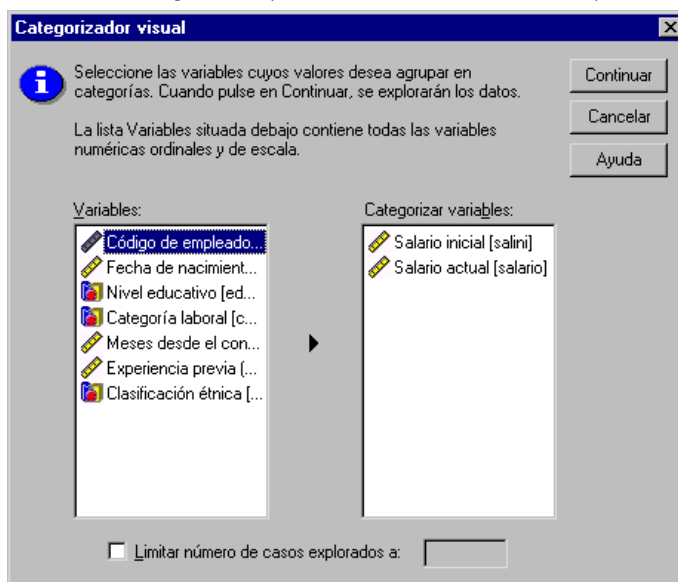
El Categorizador visual está concebido para ayudarle en el proceso de creación de variables nuevas, basado en la agrupación de los valores contiguos de las variables existentes para dar lugar a un número limitado de categorías diferentes. Puede utilizar el Categorizador visual para:

- Crear variables categóricas a partir de variables de escala continuas. Por ejemplo, puede utilizar una variable de escala con los ingresos para crear una variable categórica nueva que contenga intervalos de ingresos.
- Colapsar un número elevado de categorías ordinales en un conjunto menor de categorías. Por ejemplo, es posible colapsar una escala de evaluación de nueve categorías en tres categorías que representen: bajo, medio y alto.

En el primer paso del Categorizador visual, puede:

- ▶ Seleccione las variables numéricas de escala u ordinales para las que desee crear nuevas variables categóricas (en intervalos).

Figura 6-11
Cuadro de diálogo inicial para seleccionar las variables que se van a categorizar



Como alternativa, puede limitar la cantidad de casos que se van a explorar. Con los archivos de datos que contengan un gran número de casos, la limitación del número de casos que se va a explorar puede ahorrar tiempo, pero debe evitarse este procedimiento en lo posible, ya que influirá en la distribución de los valores que serán utilizados en los cálculos posteriores en el Categorizador visual.

Nota: Las variables de cadena y las variables numéricas nominales no se muestran en la lista de variables origen. El Categorizador visual requiere que las variables sean numéricas, medidas bien a nivel ordinal o de escala, puesto que asume que los valores de los datos representan algún tipo de orden lógico que puede ser utilizado para agrupar los valores con sentido. Puede cambiar el nivel de medida de una variable en la Vista de variables del Editor de datos. Si desea obtener más información, consulte “Nivel de medida de variable” en Capítulo 5 en p. 87.

Para categorizar variables

- ▶ Seleccione en los menús de la ventana Editor de datos:
 - Transformar
 - Categorizador visual...
- ▶ Seleccione las variables numéricas de escala u ordinales para las que desee crear nuevas variables categóricas (en intervalos).
- ▶ Seleccione una variable de la Lista de variables exploradas.
- ▶ Introduzca un nombre para la nueva variable categorizada. Los nombres de variable no pueden repetirse, y deben seguir las normas de denominación de variables de SPSS. Si desea obtener más información, consulte “Nombres de variable” en Capítulo 5 en p. 86.
- ▶ Defina los criterios de categorización para la nueva variable. Si desea obtener más información, consulte “Categorización de variables” en p. 131.
- ▶ Pulse en Aceptar.

Categorización de variables

Figura 6-12
Categorizador visual, cuadro de diálogo principal

Categorizador visual

Lista de variables exploradas

N...	Variable
✓	Salario inicial [salini]
✓	Salario actual [salario]
✓	Fecha de nacimiento [fech]
✓	Nivel educativo [educ]
✓	Código de empleado [id]
✓	Meses desde el contrato [t]
✓	Experiencia previa (meses)
✓	Categoría laboral [catlab]
✓	Clasificación étnica [minor]

Nombre: Variable actual: salario Etiqueta: Salario actual

Variable categorizada: catsal Etiqueta: Salario actual (Bandas)

Mínimo: \$15,750 Valores no perdidos Máximo: \$135,000

Introduzca puntos de corte de los intervalos o pulse Crear puntos de corte para generar intervalos automáticamente. Un valor de 10, por ejemplo, define un intervalo que comienza encima del intervalo anterior y finaliza en el 10.

Cuadrícula:

	Valor	Etiqueta
1	\$25.000	<= \$25.000
2	\$50.000	\$25.001 - \$50.000
3	\$75.000	\$50.001 - \$75.000
4	\$100.000	\$75.001 - \$100.000
5	\$125.000	\$100.001 - \$125.000
6	ALTO	\$125.000+
7		

Límites superiores:
 Incluidos (<=)
 Excluidos (<)

Crear puntos de corte...
 Crear etiquetas
 Invertir escala

Aceptar Pegar Restablecer Cancelar Ayuda

El cuadro de diálogo principal del Categorizador visual proporciona la siguiente información de las variables exploradas:

Lista de variables exploradas. Muestra las variables que fueron seleccionadas en el cuadro de diálogo inicial. Puede ordenar la lista por el nivel de medida (de escala u ordinal) o por la etiqueta o el nombre de variable, pulsando en los encabezados de las columnas.

Casos explorados. Indica el número de casos explorados. Todos los casos explorados sin valores perdidos definidos por el usuario o del sistema para la variable seleccionada, se usan en la generación de la distribución de valores que emplea el Categorizador visual, incluyendo el histograma que se visualiza en el cuadro

de diálogo principal y los puntos de corte basados en percentiles o unidades de desviación típica.

Valores perdidos. Indica el número de casos explorados con valores perdidos definidos por el usuario y perdidos del sistema. Los valores perdidos no se incluyen en ninguno de los intervalos de categorización. Si desea obtener más información, consulte “Valores perdidos definidos por el usuario en el Categorizador visual” en p. 139.

Variable actual. El nombre y etiqueta de variable (si existe) de la variable actualmente seleccionada y que se usará como base para la nueva variable categorizada.

Variable categorizada. Nombre y etiqueta de variable alternativa para la nueva variable categorizada.

- **Nombre.** Debe introducir un nombre para la nueva variable. Los nombres de variable no pueden repetirse, y deben seguir las normas de denominación de variables de SPSS. Si desea obtener más información, consulte “Nombres de variable” en Capítulo 5 en p. 86.
- **Etiqueta.** Puede introducir una etiqueta de variable descriptiva con una longitud de hasta 255 caracteres. La etiqueta de variable por defecto será la etiqueta de variable (si existe) o el nombre de variable de la variable origen con el texto (*Categorizada*) añadido al final de la etiqueta.

Mínimo y Máximo. Valores mínimo y máximo para la variable seleccionada actualmente, basados en los casos explorados y excluyendo los valores definidos como perdidos por el usuario.

Valores no perdidos. El histograma muestra la distribución de valores no perdidos correspondiente a la variable seleccionada actualmente, basándose en los casos explorados.

- Después de haber definido las categorías para la nueva variable, se mostrarán líneas verticales en el histograma para indicar los puntos de corte que definen los intervalos.
- Puede pulsar y arrastrar las líneas de los puntos de corte a distintos puntos del histograma, modificando así la amplitud de los intervalos.
- Puede eliminar categorías arrastrando las líneas de los puntos de corte fuera del histograma.

Nota: El histograma (que muestra valores no perdidos), el mínimo y el máximo se basan en los casos explorados. Si no incluye todos los casos en la exploración, es posible que no se refleje con precisión la distribución real, sobre todo si el archivo de datos se ordenó según la variable seleccionada. Si no explora ningún caso, no encontrará disponible información sobre la distribución de valores.

Cuadrícula. Muestra los valores que definen los puntos de corte superiores de cada intervalo, así como las etiquetas de valor opcionales para cada intervalo.

- **Valor.** Valores que definen los puntos de corte superiores en cada intervalo. Puede introducir los valores o utilizar **Crear puntos de corte** para crear automáticamente los intervalos basándose en los criterios seleccionados. Por defecto, se incluye automáticamente un punto de corte con el valor *SUPERIOR*. Este intervalo contendrá cualesquiera valores no perdidos por encima de los restantes puntos de corte. El intervalo definido por el punto de corte inferior incluirá todos los valores no perdidos que sean menores que o iguales a dicho valor (o, sencillamente, inferiores a ese valor, dependiendo de la forma en que haya definido los puntos de corte superiores).
- **Etiquetas.** Etiquetas opcionales y descriptivas de los valores de la nueva variable categorizada. Puesto que los valores de la nueva variable sólo serán números enteros en secuencia, del 1 *an*, las etiquetas que describan lo que representan los valores pueden resultar muy útiles. Puede introducir las etiquetas o usar **Crear etiquetas** para crear las etiquetas de valor de forma automática.

Para eliminar un intervalo de la rejilla:

- ▶ Pulse con el botón derecho las casillas *Valor* o *Etiqueta* del intervalo.
- ▶ En el menú contextual emergente, seleccione **Eliminar fila**.

Nota: Si elimina la categoría *SUPERIOR*, los casos con valores superiores al valor del último punto de corte especificado recibirán el valor perdido del sistema en la nueva variable.

Para eliminar todas las etiquetas o todas las categorías definidas:

- ▶ Pulse en cualquier parte de la rejilla con el botón derecho del ratón.
- ▶ En el menú contextual emergente, seleccione **Eliminar todas las etiquetas** o **Eliminar todos los puntos de corte**.

Límites superiores. Controla el tratamiento de los valores de los límites superiores introducidos en la columna *Valor* de la rejilla.

- **Incluidos (\leq).** Los casos con el valor especificado en la casilla *Valor* se incluyen en el intervalo. Por ejemplo, si especifica los valores 25, 50 y 75, los casos con el valor exacto 25 se incluirán en el primer intervalo, de esta forma se incluirán todos los casos con valor menor o igual a 25.
- **Excluido ($<$).** Los casos con el valor especificado en la casilla *Valor* no se incluyen en el intervalo. Por contra, se incluyen en la siguiente categoría. Por ejemplo, si especifica los valores 25, 50 y 75, los casos con el valor exacto 25 se incluirán en el segundo intervalo en vez de en el primero, puesto que el primero sólo contendrá casos con valores inferiores a 25.

Crear puntos de corte. Genera categorías automáticamente para crear intervalos de igual amplitud, intervalos con el mismo número de casos o intervalos basados en un número de desviaciones típicas. Esta posibilidad no está disponible si no se ha explorado ningún caso. Si desea obtener más información, consulte “Generación automática de categorías” en p. 134.

Crear etiquetas. Genera etiquetas descriptivas para los valores enteros consecutivos contenidos en la nueva variable categorizada, en función de los valores de la rejilla y el tratamiento especificado para los límites superiores (incluidos o excluidos).

Invertir la escala. Por defecto, los valores de la nueva variable categorizada serán números enteros consecutivos, del 1 a n . La inversión de la escala convierte los valores en números enteros consecutivos, de n a 1.

Copiar categorías. Puede copiar las especificaciones de categorización de otra variable a la variable seleccionada en ese momento, o desde la variable seleccionada en ese momento a otras varias variables. Si desea obtener más información, consulte “Copia de categorías” en p. 137.

Generación automática de categorías

El cuadro de diálogo Crear puntos de corte permite la creación automática de categorías en función de los criterios seleccionados.

Para utilizar el cuadro de diálogo Crear puntos de corte:

- ▶ Seleccione (pulse) una variable de la Lista de variables exploradas.
- ▶ Pulse Crear puntos de corte.
- ▶ Seleccione los criterios de generación de los puntos de corte que definirán los intervalos.
- ▶ Pulse en Aplicar.

Figura 6-13
Cuadro de diálogo Crear puntos de corte

Crear puntos de corte

Intervalos de igual amplitud

Intervalos: rellene al menos dos campos:

Posición del primer punto de corte: \$25,000

Número de puntos de corte: 5

Amplitud: 25000

Posición del último punto de corte: \$125,000

Percentiles iguales basados en los casos explorados

Intervalos - rellene cualquiera de los dos campos:

Número de puntos de corte:


% de casos:

Puntos de corte en media y desviaciones típicas seleccionadas, basadas en casos explorados

+/- 1 Desv. típica

+/- 2 Desv. típicas

+/- 3 Desv. típicas

 Aplicar reemplazará las definiciones de los puntos de corte actuales con esta especificación. Un intervalo final incluirá todos los valores restantes: N puntos de corte generan N+1 intervalos.

Aplicar

Cancelar

Ayuda

Nota: El cuadro de diálogo Crear puntos de corte no está disponible si no se ha explorado ningún caso.

Intervalos de igual amplitud. Genera categorías de igual amplitud (por ejemplo, 1–10, 11–20, 21–30, etc.), basándose en dos (cualesquiera) de los tres criterios siguientes:

- **Posición del primer punto de corte.** Valor que define el límite superior de la categoría inferior (por ejemplo, el valor 10 indica un intervalo que incluya todos los valores hasta 10).
- **Número de puntos de corte.** El número de intervalos es el número de puntos de corte *más uno*. Por ejemplo, 9 puntos de corte generan 10 intervalos.
- **Amplitud.** La amplitud de cada intervalo. Por ejemplo, el valor 10 categorizará la variable *Edad en años* en intervalos de 10 años.

Percentiles iguales basados en los casos explorados. Genera intervalos con un número igual de casos en cada categoría (utilizando el algoritmo "empírico" para el cálculo de percentiles), en función de uno de los criterios siguientes:

- **Número de puntos de corte.** El número de intervalos es el número de puntos de corte *más uno*. Por ejemplo, tres puntos de corte generan cuatro categorías percentiles (cuartiles), conteniendo cada una el 25% de los casos.
- **% de casos.** Amplitud de cada intervalo, expresado en forma de porcentaje sobre el número total de casos. Por ejemplo, el valor 33,3 generaría tres categorías (dos puntos de corte), conteniendo cada una el 33,3% de los casos.

Si la variable origen contiene un número relativamente pequeño de valores distintos, o un gran número de casos con el mismo valor, es posible que obtenga menos categorías que las solicitadas. En caso de haber varios valores idénticos en un punto de corte, todos se incluyen en el mismo intervalo; por consiguiente, los porcentajes reales pueden no ser siempre iguales.

Puntos de corte en media y desviaciones típicas seleccionadas, basadas en casos explorados. Genera categorías basándose en los valores de la media y la desviación típica de la distribución de la variable.

- Si no selecciona ninguno de los intervalos de desviación típica, se crearán dos intervalos, siendo la media el punto de corte que divide las categorías.
- Puede seleccionar cualquier combinación de los intervalos de desviación típica, basándose en una, dos o tres desviaciones típicas. Por ejemplo, al seleccionar las tres opciones se obtendrán ocho intervalos: seis intervalos distanciados en una desviación típica de amplitud y dos intervalos para los casos que se encuentren a más de tres desviaciones típicas por encima y por debajo de la media.

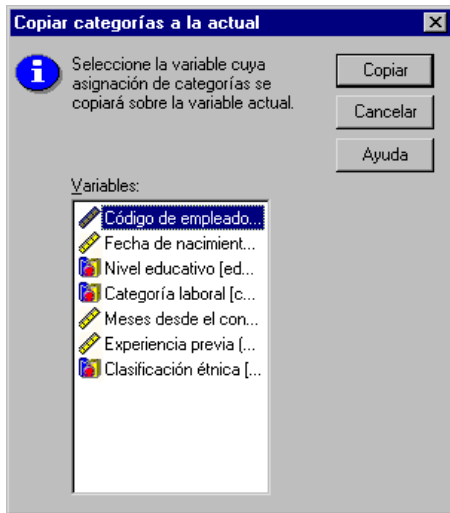
En una distribución normal, el 68% de los casos se encuentra dentro de una distancia de una desviación típica respecto a la media, el 95% entre dos desviaciones típicas y el 99% dentro de tres desviaciones típicas. La creación de categorías basadas en desviaciones típicas puede ocasionar que algunas categorías queden definidas fuera del rango real de los datos, e incluso fuera del rango de valores posibles de los datos (por ejemplo, un rango de salarios negativos).

Nota: Los cálculos de los percentiles y las desviaciones típicas se basan en los casos explorados. Si limita el número de casos explorados, puede que las categorías resultantes no incluyan la proporción de casos deseada en dichas categorías, sobre todo si el archivo de datos se ordenó según la variable origen. Por ejemplo, si limita la exploración a los primeros 100 casos de un archivo de datos con 1000 casos y el archivo de datos está ordenado en orden descendente por edad del encuestado, en lugar de cuatro categorías percentiles de la edad, cada una con el 25% de los casos, podría encontrarse con que las tres primeras categorías contuvieran cada una sólo en torno al 3,3% de los casos, mientras que la última categoría albergaría el 90% de los casos.

Copia de categorías

Al crear categorías para una o más variables, puede copiar las especificaciones de categorización de otra variable a la seleccionada en ese momento, o desde la variable seleccionada en ese momento a varias otras variables.

Figura 6-14
Copia de categorías en o desde la variable actual



Para copiar especificaciones de categorización:

- ▶ Defina los intervalos para una variable como mínimo; pero *no* pulse en Aceptar ni en Pegar.
 - ▶ Seleccione (pulse) una variable de la Lista de variables exploradas para la cual haya definido intervalos.
 - ▶ Pulse A otras variables.
 - ▶ Seleccione las variables para las que desea crear nuevas variables con los mismos intervalos.
 - ▶ Pulse Copiar.
- o*
- ▶ Seleccione (pulse) una variable de la Lista de variables exploradas sobre la cual desea copiar intervalos ya definidos.
 - ▶ Pulse De otra variable.

- ▶ Seleccione la variable que contiene los intervalos definidos que desea copiar.
- ▶ Pulse Copiar.

También se copiarán las etiquetas de valor si se especificaron en la variable cuyas especificaciones de categorización se van a copiar.

Nota: Una vez que haya pulsado en Aceptar en el cuadro de diálogo principal del Categorizador visual, para crear nuevas variables categorizadas (o cerrado el cuadro de diálogo de alguna otra forma), no podrá usar de nuevo el Categorizador visual para copiar dichas categorías en otras variables.

Valores perdidos definidos por el usuario en el Categorizador visual

Los valores perdidos definidos por el usuario (valores identificados como los códigos para los datos perdidos) para la variable origen no se incluyen en los intervalos de la nueva variable. Los valores perdidos definidos por el usuario de las variables se copian como valores perdidos definidos por el usuario en la nueva variable, copiándose también cualquier otra etiqueta de valor definida para los códigos de los valores perdidos.

Si un código de valor perdido entra en conflicto con alguno de los valores de categorización de la nueva variable, el código de valor perdido en la nueva variable se recodificará a un valor no conflictivo, sumando 100 al valor de categoría superior. Por ejemplo, si el usuario define el valor 1 como valor perdido para la variable origen, y la nueva variable va a contar con seis categorías, cualquier caso con el valor 1 en la variable origen tendrá el valor 106 en la nueva variable, y 106 será definido como un valor perdido del usuario. Si el valor perdido definido por el usuario en la variable de origen tenía definida una etiqueta de valor dicha etiqueta se mantendrá como etiqueta de valor para el valor recodificado de la nueva variable.

Nota: Si la variable de origen tiene definido un rango de valores perdidos de usuario con la forma *MENOR-n*, donde *n* es un número positivo, los valores perdidos definidos por el usuario correspondientes a la nueva variable, serán números negativos.

Transformaciones de los datos

En una situación ideal, los datos brutos son perfectamente apropiados para el tipo de análisis que se desea realizar y cualquier relación existente entre las variables o es adecuadamente lineal o es claramente ortogonal. Desafortunadamente, esto ocurre pocas veces. El análisis preliminar puede revelar esquemas de codificación poco prácticos o errores de codificación, o bien pueden requerirse transformaciones de los datos para exponer la verdadera relación existente entre las variables.

Puede realizar transformaciones de los datos de todo tipo, desde tareas sencillas, como la agrupación de categorías para su análisis posterior, hasta otras más avanzadas, como la creación de nuevas variables basadas en ecuaciones complejas e instrucciones condicionales.

Cálculo de variables

Utilice el cuadro de diálogo Calcular para calcular los valores de una variable basándose en transformaciones numéricas de otras variables.

- Puede calcular valores para las variables numéricas o de cadena (alfanuméricas).
- Puede crear nuevas variables o bien reemplazar los valores de las variables existentes. Para las nuevas variables, también se puede especificar el tipo y la etiqueta de variable.
- Puede calcular valores de forma selectiva para subconjuntos de datos basándose en condiciones lógicas.
- Puede utilizar más de 70 funciones preincorporadas, incluyendo funciones aritméticas, funciones estadísticas, funciones de distribución y funciones de cadena.

Figura 7-1
Cuadro de diálogo *Calcular variable*



Para calcular variables

- ▶ Elija en los menús:
Transformar
Calcular...
- ▶ Escriba el nombre de una sola variable de destino. Puede ser una variable existente o una nueva que se vaya a añadir al archivo de datos de trabajo.
- ▶ Para crear una expresión, puede pegar los componentes en el campo Expresión o escribir directamente en dicho campo.
 - Puede pegar las funciones o las variables de sistema utilizadas habitualmente seleccionando un grupo de la lista Grupo de funciones y pulsando dos veces en la función o variable de las listas de funciones y variables especiales (o seleccione la función o variable y pulse en la flecha que se encuentra sobre la lista Grupo de

funciones). Rellene los parámetros indicados mediante interrogaciones (aplicable sólo a las funciones). El grupo de funciones con la etiqueta Todo contiene una lista de todas las funciones y variables de sistema disponibles. En un área reservada del cuadro de diálogo se muestra una breve descripción de la función o variable actualmente seleccionada.

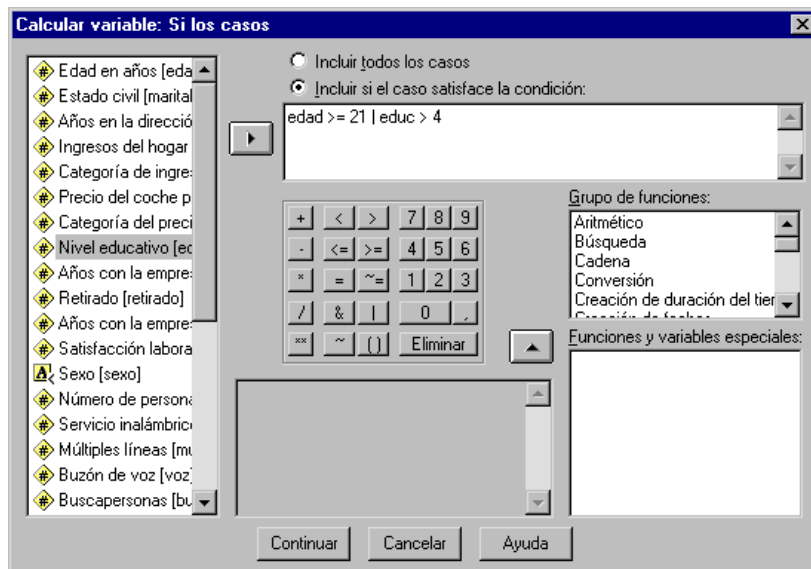
- Las constantes de cadena deben ir entre comillas o apóstrofos.
- Si los valores contienen decimales, debe utilizarse una coma(,) como indicador decimal.
- Para las nuevas variables de cadena, también deberán seleccionar Tipo y etiqueta para especificar el tipo de datos.

Calcular variable: Si los casos

El cuadro de diálogo Si los casos permite aplicar transformaciones de los datos para subconjuntos de casos seleccionados utilizando expresiones condicionales. Una expresión condicional devuelve un valor *verdadero*, *falso* o *perdido* para cada caso.

Figura 7-2

Cuadro de diálogo Calcular variable: Si los casos



- Si el resultado de una expresión condicional es *verdadero*, se incluirá el caso en el subconjunto seleccionado.
- Si el resultado de una expresión condicional es *falso* o *perdido*, no se incluirá el caso en el subconjunto seleccionado.
- La mayoría de las expresiones condicionales utilizan al menos uno de los seis operadores de relación (<, >, <=, >=, =, and ~=) de la calculadora.
- Las expresiones condicionales pueden incluir nombres de variable, constantes, operadores aritméticos, funciones numéricas y de otros tipos, variables lógicas y operadores de relación.

Calcular variable: Tipo y etiqueta

Por defecto, las nuevas variables calculadas son numéricas. Para calcular una nueva variable de cadena, deberá especificar el tipo de los datos y su ancho.

Etiqueta. Etiqueta de variable descriptiva y opcional de hasta 120 caracteres. Puede escribir el texto de una etiqueta o bien utilizar los primeros 110 caracteres de la expresión de cálculo.

Tipo. Las variables calculadas pueden ser numéricas o de cadena (alfanuméricas). Las variables de cadena no se pueden utilizar en cálculos aritméticos.

Figura 7-3
Cuadro de diálogo Tipo y etiqueta



Funciones

Se dispone de muchos tipos de funciones, entre ellos:

- Funciones aritméticas

- Funciones estadísticas
- Funciones de cadena
- Funciones de fecha y hora
- Funciones de distribución
- Funciones de variables aleatorias
- Funciones de valores perdidos
- Funciones de puntuación (sólo servidor de SPSS)

Si desea obtener una lista completa de las funciones, busque funciones en el índice del sistema de Ayuda en pantalla.

Valores perdidos en funciones

Las funciones y las expresiones aritméticas sencillas tratan los valores perdidos de diferentes formas. En la expresión:

```
(var1+var2+var3)/3
```

El resultado es el valor perdido si un caso tiene un valor perdido para cualquiera de las tres variables.

En la expresión:

```
MEAN(var1, var2, var3)
```

El resultado es el valor perdido sólo si el caso tiene valores perdidos para las tres variables.

En las funciones estadísticas se puede especificar el número mínimo de argumentos que deben tener valores no perdidos. Para ello, escriba un punto y el número mínimo de argumentos después del nombre de la función, como en:

```
MEAN.2(var1, var2, var3)
```

Generadores de números aleatorios

El cuadro de diálogo Generadores de números aleatorios le permite seleccionar el generador de números aleatorios y establecer el valor de secuencia de inicio de modo que pueda reproducir una secuencia de números aleatorios.

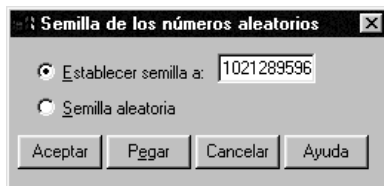
Generador activo. Hay dos generadores de números aleatorios disponibles:

- **Compatible con SPSS 12.** El generador de números aleatorios utilizado en SPSS 12 y versiones anteriores. Utilice este generador de números aleatorios si necesita reproducir los resultados aleatorizados generados por versiones previas basadas en una semilla de aleatorización especificada.
- **Tornado de Mersenne.** Un generador de números aleatorios nuevo que es más fiable en los procesos de simulación. Utilice este generador de números aleatorios si no es necesario reproducir resultados aleatorizados correspondientes a SPSS 12 o anteriores.

Inicialización del generador activo. La semilla de aleatorización cambia cada vez que se genera un número aleatorio para utilizarlo en las transformaciones (como las funciones de distribución aleatorias), el muestreo aleatorio o la ponderación de los casos. Para replicar una secuencia de números aleatorios, establezca el valor de inicialización del punto de inicio antes de cada análisis que utilice los números aleatorios. El valor debe ser un entero positivo.

Figura 7-4

Cuadro de diálogo Semilla de aleatorización



Para seleccionar el generador de números aleatorios y establecer el valor de inicialización:

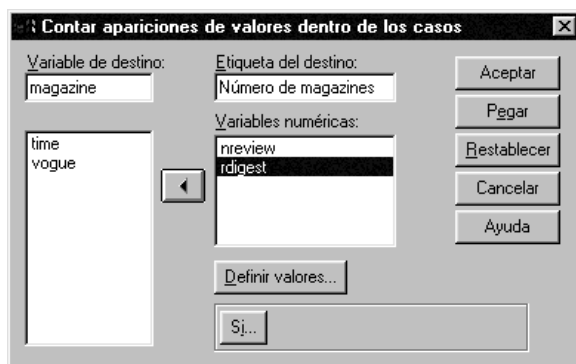
- ▶ Elija en los menús:
Transformar
Generadores de números aleatorios

Contar apariciones de valores dentro de los casos

Este cuadro de diálogo crea una variable que, para cada caso, cuenta las apariciones del mismo valor, o valores, en una lista de variables. Por ejemplo, un estudio podrá contener una lista de revistas con las casillas de verificación *sí/no* para indicar qué revistas lee cada encuestado. Se podría contar el número de respuestas *sí* de cada encuestado para crear una nueva variable que contenga el número total de revistas leídas.

Figura 7-5

Cuadro de diálogo Contar apariciones de valores dentro de los casos



Para contar apariciones de valores dentro de los casos

- ▶ Elija en los menús:
Transformar
Contar apariciones...
- ▶ Introduzca el nombre de la variable de destino.
- ▶ Seleccione dos o más variables del mismo tipo (numéricas o de cadena).
- ▶ Pulse en Definir valores y especifique los valores que se deben contar.

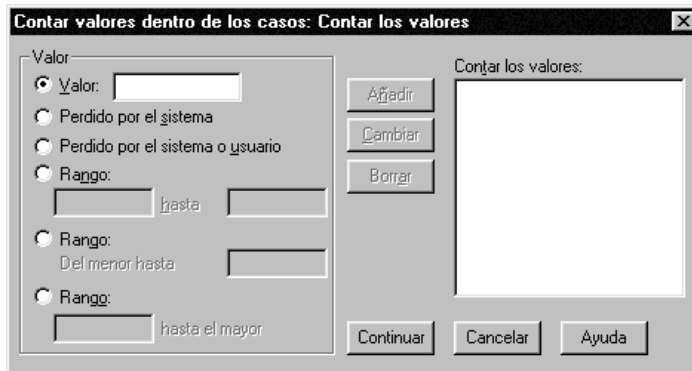
Si lo desea, puede definir un subconjunto de casos en los que contar las apariciones de valores.

Contar valores dentro de los casos: Contar los valores

El valor de la variable de destino (en el cuadro de diálogo principal) se incrementa en 1 cada vez que una de las variables seleccionadas coincide con una especificación de la lista Contar los valores. Si un caso coincide con varias de las especificaciones en cualquiera de las variables, la variable de destino se incrementa varias veces para esa variable.

Las especificaciones de valores pueden incluir valores individuales, valores perdidos o valores perdidos del sistema y rangos de valores. Los rangos incluyen sus puntos finales y los valores definidos como perdidos por el usuario que estén dentro del rango.

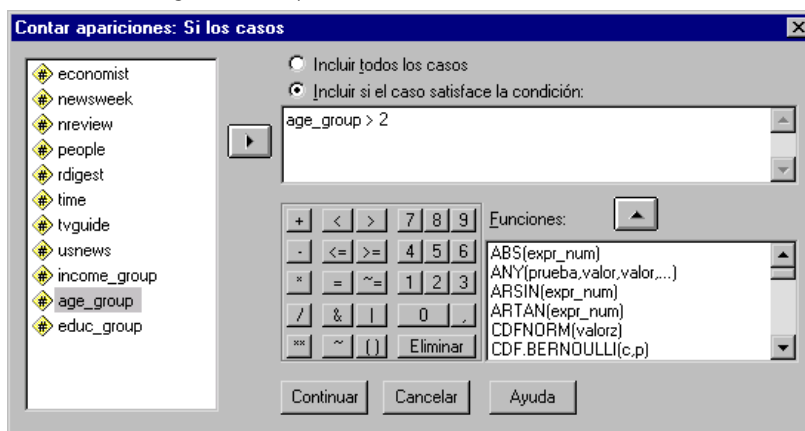
Figura 7-6
Cuadro de diálogo Contar los valores



Contar apariciones: Si los casos

El cuadro de diálogo Si los casos permite contar apariciones de valores para un subconjunto de casos seleccionado utilizando expresiones condicionales. Una expresión condicional devuelve un valor *verdadero*, *falso* o *perdido* para cada caso.

Figura 7-7
Cuadro de diálogo Contar apariciones: Si los casos



Si desea conocer las consideraciones generales sobre el uso del cuadro de diálogo Si los casos, consulte “Calcular variable: Si los casos” en p. 143.

Recodificación de valores

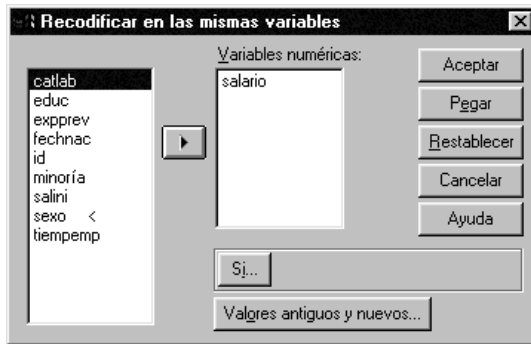
Los valores de datos se pueden modificar mediante la recodificación. Esto es particularmente útil para agrupar o combinar categorías. Puede recodificar los valores dentro de las variables existentes o crear variables nuevas que se basen en los valores recodificados de las variables existentes.

Recodificar en las mismas variables

El cuadro de diálogo Recodificar en las mismas variables le permite reasignar los valores de las variables existentes o agrupar rangos de valores existentes en nuevos valores. Por ejemplo, podría agrupar los salarios en categorías que sean rangos de salarios.

Puede recodificar las variables numéricas y de cadena. Si selecciona múltiples variables, todas deben ser del mismo tipo. No se pueden recodificar juntas las variables numéricas y de cadena.

Figura 7-8
Cuadro de diálogo *Recodif. en mismas var.*



Para recodificar los valores de una variable

- ▶ Elija en los menús:
Transformar
 Recodificar
 En las mismas variables...
- ▶ Seleccione las variables que desee recodificar. Si selecciona múltiples variables, todas deberán ser del mismo tipo (numéricas o de cadena).
- ▶ Pulse en Valores antiguos y nuevos y especifique cómo deben recodificarse los valores.

Si lo desea, puede definir un subconjunto de los casos para su recodificación. El cuadro de diálogo Si los casos para esto es igual al que se describe para Contar apariciones.

Recodificar en las mismas variables: Valores antiguos y nuevos

Este cuadro de diálogo permite definir los valores que se van a recodificar. Todas las especificaciones de valores deben pertenecer al mismo tipo de datos (numéricos o de cadena) que las variables seleccionadas en el cuadro de diálogo principal.

Valor antiguo. Determina el valor o los valores que se van a recodificar. Puede recodificar valores individuales, rangos de valores y valores perdidos. Los rangos y los valores perdidos del sistema no se pueden seleccionar para las variables de

cadena, ya que ninguno de los conceptos es aplicable a estas variables. Los rangos incluyen sus puntos finales y los valores definidos como perdidos por el usuario que estén dentro del rango.

- **Valor (Recodificar).** Valor antiguo individual que se a recodificar en un valor nuevo. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos de las variables desee recodificar.
- **Perdido por el sistema.** Valores asignados por SPSS cuando los valores de sus datos no están definidos de acuerdo al tipo de formato que haya especificado, cuando un campo numérico está vacío, o cuando no está definido un valor como resultado de un comando de transformación. Los valores numéricos perdidos del sistema se muestran como puntos. Las variables de cadena no pueden tener valores perdidos del sistema, ya que es lícito cualquier carácter en las variables de cadena.
- **Perdido por el sistema o por el usuario.** Observaciones que tienen valores que, o bien han sido declarados perdidos por el usuario o bien son desconocidos y han sido asignados el valor perdido por el sistema, lo cual se indica mediante un punto (.).
- **Rango (Rango inclusivo de valores).** Un rango inclusivo de valores. No está disponible para variables de cadena. Se incluirán cualesquiera valores definidos por el usuario como perdidos que se encuentren dentro del rango.
- **Todos los demás valores.** Cualquier valor no incluido en una de las especificaciones de la lista Antiguo->Nuevo. Aparece en la lista Antiguo->Nuevo como ELSE (otros).

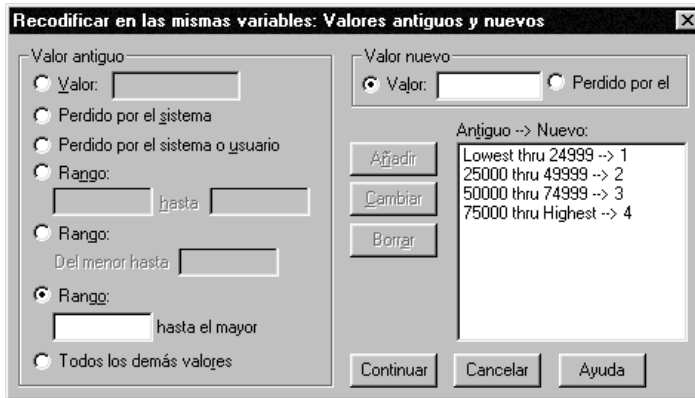
Valor nuevo. Es el valor individual en el que se recodifica cada valor o rango de valores antiguo. Puede introducir un valor o asignar el valor perdido del sistema.

- **Valor.** Valor en el que se va a recodificar uno o más valores antiguos. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos del valor antiguo.
- **Perdido por el sistema.** Recodifica el valor antiguo especificado como valor perdido por el sistema. El valor perdido por el sistema no se utiliza en los cálculos. Además, los casos con valor perdido por el sistema se excluyen de muchos procedimientos. No se encuentra disponible para las variables de cadena.

Antiguo->Nuevo. Contiene la lista de especificaciones que se va a utilizar para recodificar la variable o las variables. Puede añadir, cambiar y borrar las especificaciones que desee. La lista se ordena automáticamente basándose en la especificación del valor antiguo y siguiendo este orden: valores únicos, valores

perdidos, rangos y todos los demás valores. Si cambia una especificación de recodificación en la lista, el procedimiento volverá a ordenar la lista automáticamente, si fuera necesario, para mantener este orden.

Figura 7-9
Cuadro de diálogo Valores antiguos y nuevos

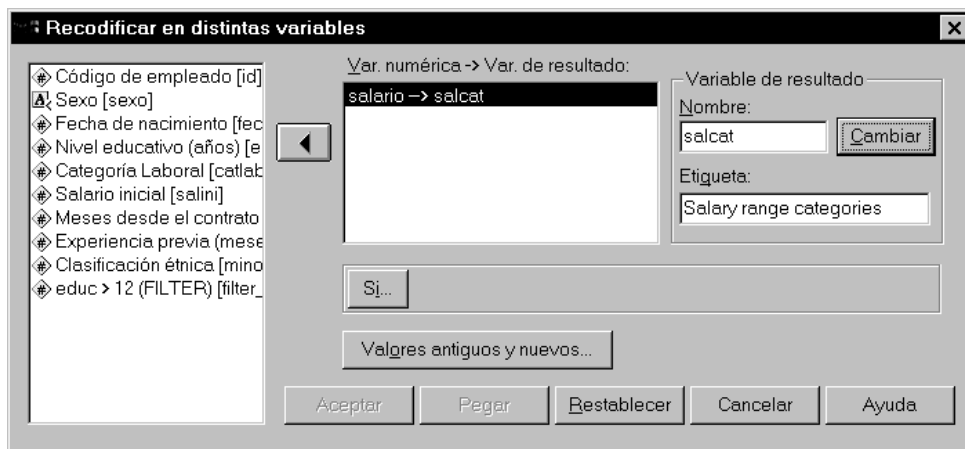


Recodificar en distintas variables

El cuadro de diálogo Recodificar en distintas variables le permite reasignar los valores de las variables existentes o agrupar rangos de valores existentes en nuevos valores para una variable nueva. Por ejemplo, podría agrupar los salarios en una nueva variable que contenga categorías de rangos de salarios.

- Puede recodificar las variables numéricas y de cadena.
- Puede recodificar variables numéricas en variables de cadena y viceversa.
- Si selecciona múltiples variables, todas deben ser del mismo tipo. No se pueden recodificar juntas las variables numéricas y de cadena.

Figura 7-10
Cuadro de diálogo Recodif. en distintas var. (Recodificar en distintas variables)



Para recodificar los valores de una variable en una nueva variable

- ▶ Elija en los menús:
 - Transformar
 - Recodificar
 - En distintas variables...
- ▶ Seleccione las variables que desee recodificar. Si selecciona múltiples variables, todas deberán ser del mismo tipo (numéricas o de cadena).
- ▶ Introduzca el nombre de la nueva variable de resultado para cada nueva variable y pulse en Cambiar.
- ▶ Pulse en Valores antiguos y nuevos y especifique cómo deben recodificarse los valores.

Si lo desea, puede definir un subconjunto de los casos para su recodificación. El cuadro de diálogo Si los casos para esto es igual al que se describe para Contar apariciones.

Recodificar en distintas variables: Valores antiguos y nuevos

Este cuadro de diálogo permite definir los valores que se van a recodificar.

Valor antiguo. Determina el valor o los valores que se van a recodificar. Puede recodificar valores individuales, rangos de valores y valores perdidos. Los rangos y los valores perdidos del sistema no se pueden seleccionar para las variables de cadena, ya que ninguno de los conceptos es aplicable a estas variables. Los valores antiguos deben ser del mismo tipo de datos (numéricos o de cadena) que la variable original. Los rangos incluyen sus puntos finales y los valores definidos como perdidos por el usuario que estén dentro del rango.

- **Valor (Recodificar).** Valor antiguo individual que se a recodificar en un valor nuevo. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos de las variables desee recodificar.
- **Perdido por el sistema.** Valores asignados por SPSS cuando los valores de sus datos no están definidos de acuerdo al tipo de formato que haya especificado, cuando un campo numérico está vacío, o cuando no está definido un valor como resultado de un comando de transformación. Los valores numéricos perdidos del sistema se muestran como puntos. Las variables de cadena no pueden tener valores perdidos del sistema, ya que es lícito cualquier carácter en las variables de cadena.
- **Perdido por el sistema o por el usuario.** Observaciones que tienen valores que, o bien han sido declarados perdidos por el usuario o bien son desconocidos y han sido asignados el valor perdido por el sistema, lo cual se indica mediante un punto (.).
- **Rango (Rango inclusivo de valores).** Un rango inclusivo de valores. No está disponible para variables de cadena. Se incluirán cualesquiera valores definidos por el usuario como perdidos que se encuentren dentro del rango.
- **Todos los demás valores.** Cualquier valor no incluido en una de las especificaciones de la lista Antiguo->Nuevo. Aparece en la lista Antiguo->Nuevo como ELSE (otros).

Valor nuevo. Es el valor individual en el que se recodifica cada valor o rango de valores antiguo. Los valores nuevos pueden ser numéricos o de cadena.

- **Valor.** Valor en el que se va a recodificar uno o más valores antiguos. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos del valor antiguo.

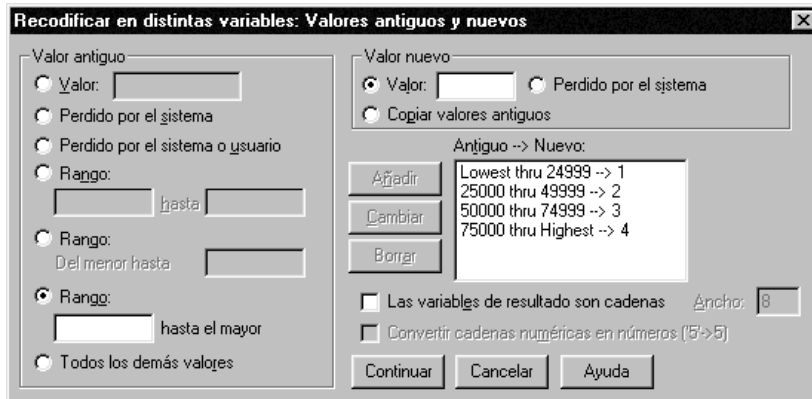
- **Perdido por el sistema.** Recodifica el valor antiguo especificado como valor perdido por el sistema. El valor perdido por el sistema no se utiliza en los cálculos. Además, los casos con valor perdido por el sistema se excluyen de muchos procedimientos. No se encuentra disponible para las variables de cadena.
- **Copiar los valores antiguos.** Retiene el valor antiguo. Si algunos de los valores no requieren la recodificación, utilice esta opción para incluir los valores antiguos. Cualquier valor antiguo que no sea especificado no se incluye en la nueva variable, y los casos con esos valores se asignan al valor perdido por el sistema en la nueva variable.

Las variables de los resultados son cadenas. Define la nueva variable recodificada como variable de cadena (alfanumérica). La variable antigua puede ser numérica o de cadena.

Convertir cadenas numéricas en números. Convierte los valores de cadena que contienen números en valores numéricos. A las cadenas que contengan cualquier cosa que no sean números y un carácter de signo opcional (+ ó -) se les asignará el valor perdido por el sistema.

Antiguo→Nuevo. Contiene la lista de especificaciones que se va a utilizar para recodificar la variable o las variables. Puede añadir, cambiar y borrar las especificaciones que desee. La lista se ordena automáticamente basándose en la especificación del valor antiguo y siguiendo este orden: valores únicos, valores perdidos, rangos y todos los demás valores. Si cambia una especificación de recodificación en la lista, el procedimiento volverá a ordenar la lista automáticamente, si fuera necesario, para mantener este orden.

Figura 7-11
Cuadro de diálogo Valores antiguos y nuevos



Asignar rangos a los casos

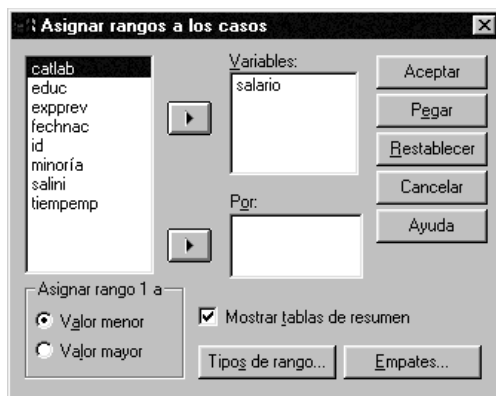
El cuadro de diálogo Asignar rangos a los casos le permite crear nuevas variables que contienen rangos, puntuaciones de Savage y normales, y los valores de los percentiles para las variables numéricas.

Los nombres de las nuevas variables y las etiquetas de variable descriptivas se generan automáticamente en función del nombre de la variable original y de las medidas seleccionadas. Una tabla de resumen presenta una lista de las variables originales, las nuevas variables y las etiquetas de variable.

Si lo desea, tiene la posibilidad de:

- Asignar los rangos a los casos en orden ascendente o descendente.
- Organizar los rangos en subgrupos seleccionando una o más variables de agrupación para la lista Por. Los rangos se calculan dentro de cada grupo, y los grupos se definen mediante la combinación de los valores de las variables de agrupación. Por ejemplo, si selecciona *sexo* y *minoría* como variables de agrupación, los rangos se calcularán para cada combinación de *sexo* y *minoría*.

Figura 7-12
Cuadro de diálogo *Asignar rangos a los casos*



Para asignar rangos a los casos

- ▶ Elija en los menús:
Transformar
Asignar rangos a casos...
- ▶ Seleccione la variable o variables a las que desee asignar los rangos. Sólo se pueden asignar rangos a las variables numéricas.

Si lo desea, puede asignar los rangos a los casos en orden ascendente o descendente y organizar los rangos por subgrupos.

Asignar rangos a los casos: Tipos

Puede seleccionar diversos métodos de asignación de rangos. En cada método se crea una variable diferente de asignación de rangos. Los métodos de asignación de rangos incluyen rangos sencillos, puntuaciones de Savage, rangos fraccionales y percentiles. También puede crear rangos basados en estimaciones de la proporción y puntuaciones normales.

Rango. El rango simple. El valor de la nueva variable es igual a su rango.

Puntuaciones de Savage. La nueva variable contiene puntuaciones de Savage basadas en una distribución exponencial.

Rango fraccional. El valor de la nueva variable es igual al rango dividido por la suma de las ponderaciones de los casos no perdidos.

Rango fraccional como porcentaje. Cada rango se divide por el número de casos que tienes valores válidos y se multiplica por 100.

Suma de ponderaciones de los casos. El valor de la nueva variable es igual a la suma de las ponderaciones de los casos. La nueva variable es una constante para todos los casos del mismo grupo.

Ntiles. Los rangos se basan en los grupos percentiles, con cada uno de los grupos conteniendo aproximadamente el mismo número de casos. Por ejemplo, 4 Ntiles asignará un rango 1 a los casos por debajo del percentil 25, 2 a los casos entre los percentiles 25 y 50, 3 a los casos entre los percentiles 50 y 75, y 4 a los casos por encima del percentil 75.

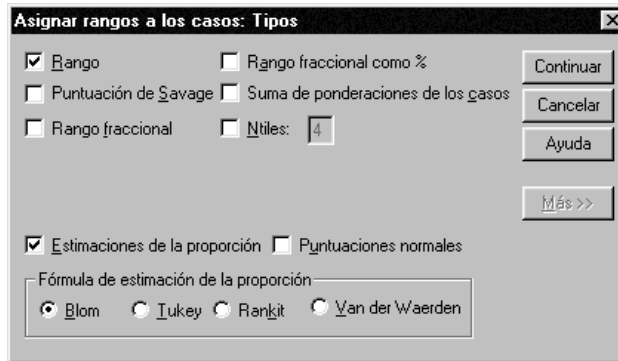
Estimaciones de proporción. Estimaciones de la proporción acumulada de la distribución que corresponde a un rango particular.

Puntuaciones normales. Puntuaciones z correspondientes a la proporción acumulada estimada.

Fórmula de estimación de la proporción. Para estimaciones de la proporción y puntuaciones normales se pueden seleccionar las fórmulas de estimación de la proporción: Blom, Tukey, Rankit o Van der Waerden.

- **Blom.** Crea nuevas variables de ordenación (rangos) que se basan en estimaciones de la proporción, las cuales utilizan la fórmula $(r-3/8) / (w+1/4)$, donde r es el rango y w es la suma de las ponderaciones de los casos.
- **Tukey.** Utiliza la fórmula $(r-1/3) / (w+1/3)$, donde r es el rango y w es la suma de las ponderaciones de los casos.
- **Rankit.** Utiliza la fórmula $(r-1/2) / w$, donde w es el número de observaciones y r es el rango, que va de 1 a w .
- **Van der Waerden.** La transformación de Van de Waerden, definida por la fórmula $r/(w+1)$, donde w es la suma de las ponderaciones de los casos y r es el rango, cuyo valor va de 1 a w .

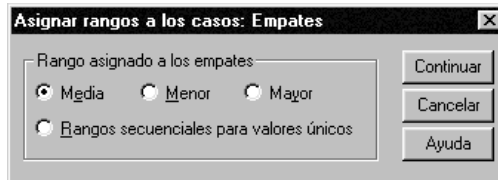
Figura 7-13
Cuadro de diálogo *Asignar rangos a los casos: Tipos*



Asignar rangos a los casos: Empates

Este cuadro de diálogo controla el método de asignación de rangos a los casos con el mismo valor en la variable original.

Figura 7-14
Cuadro de diálogo *Asignar rangos a los casos: Empates*



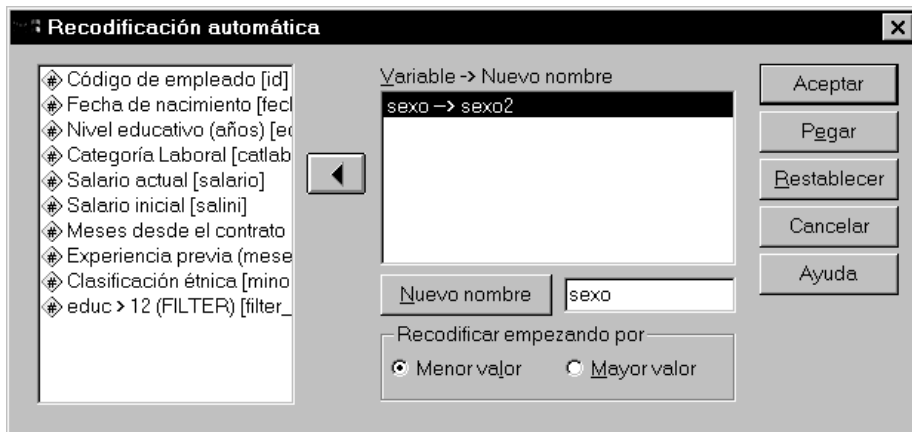
La tabla siguiente muestra cómo los distintos métodos asignan rangos a los valores empatados.

Valor	Media	Bajo	Alto	Secuencia
10	1	1	1	1
15	3	2	4	2
15	3	2	4	2
15	3	2	4	2
16	5	5	5	3
20	6	6	6	4

Recodificación automática

El cuadro de diálogo Recodificación automática le permite convertir los valores numéricos y de cadena en valores enteros consecutivos. Si los códigos de la categoría no son secuenciales, las casillas vacías resultantes reducen el rendimiento e incrementan los requisitos de memoria de muchos procedimientos. Además, algunos procedimientos no pueden utilizar variables de cadena y otros requieren valores enteros consecutivos para los niveles de los factores.

Figura 7-15
Cuadro de diálogo Recodificación automática



- La nueva variable, o variables, creadas por la recodificación automática conservan todas las etiquetas de variable y de valor definidas de la variable antigua. Para los valores que no tienen una etiqueta de valor ya definida se utiliza el valor original como etiqueta del valor recodificado. Una tabla muestra los valores antiguos, los nuevos y las etiquetas de valor.
- Los valores de cadena se recodifican por orden alfabético, con las mayúsculas antes que las minúsculas.
- Los valores perdidos se recodifican como valores perdidos mayores que cualquier valor no perdido y conservando el orden. Por ejemplo, si la variable original posee 10 valores no perdidos, el valor perdido mínimo se recodificará como 11, y el valor 11 será un valor perdido para la nueva variable.

Usar el mismo esquema de recodificación para todas las variables. Esta opción le permite aplicar un único esquema de recodificación para todas las variables seleccionadas, lo que genera un esquema de codificación coherente para todas las variables nuevas.

Si selecciona esta opción, se aplican las siguientes reglas y limitaciones:

- Todas las variables deben ser del mismo tipo (numéricas o de cadena).
- Todos los valores observados para todas las variables seleccionadas se utilizan para crear un orden de valores para recodificar en enteros consecutivos.
- Los valores perdidos definidos por el usuario para las variables nuevas se basan en la primera variable de la lista con valores perdidos definidos por el usuario. El resto de los valores de las demás variables originales, excepto los valores perdidos del sistema, se consideran válidos.

Trate los valores de cadena en blanco como valores perdidos definidos por el usuario.

En el caso de las variables de cadena, los valores en blanco o nulos no son tratados como **valores perdidos del sistema**. Esta opción recodifica automáticamente las cadenas en blanco en un valor **perdido definido por el usuario** mayor que el valor no perdido máximo.

Plantillas

Puede guardar el esquema de recodificación automática en un archivo de plantilla y, a continuación, aplicarlo a otras variables y otros archivos de datos.

Por ejemplo, puede tener un número considerable de códigos de producto alfanuméricos que se registran automáticamente en enteros cada mes, pero algunos meses se añaden códigos de productos nuevos al esquema de recodificación original. Si guarda el esquema original en una plantilla y, a continuación, la aplica a los datos nuevos que contienen el nuevo conjunto de códigos, todos los códigos nuevos encontrados en los datos se recodifican automáticamente en valores superiores al último valor de la plantilla para conservar el esquema de recodificación automática original de los códigos de productos originales.

Guardar plantilla Guarda el esquema de recodificación automática para las variables seleccionadas en un archivo de plantilla externo.

- La plantilla contiene información que asigna los valores no perdidos originales a los valores recodificados.
- En la plantilla sólo se guarda la información para los valores no perdidos. La información sobre los valores perdidos definidos por el usuario no se conserva.
- Si ha seleccionado varias variables para su recodificación, pero no ha optado por utilizar el mismo esquema de recodificación automática para todas las variables o no va a aplicar una plantilla existente como parte de la recodificación automática, la plantilla se basará en la primera variable de la lista.
- Si ha seleccionado varias variables para su recodificación, y también ha seleccionado Usar el mismo esquema de recodificación para todas las variables y/o Aplicar plantilla, la plantilla contendrá el esquema de recodificación automática combinado para todas las variables.

Aplicar plantilla. Aplica una plantilla de recodificación automática previamente guardada a las variables seleccionadas para la recodificación, añade los valores adicionales encontrados en las variables al final del esquema y conserva la relación entre los valores originales y recodificados automáticamente almacenados en el esquema guardado.

- Todas las variables seleccionadas para la recodificación deben ser del mismo tipo (numéricas o de cadena) y dicho tipo debe coincidir con el tipo definido en la plantilla.
- Las plantillas no pueden contener información sobre los valores perdidos definidos por el usuario. Los valores perdidos definidos por el usuario para las variables de destino se basan en la primera variable de la lista de variables originales con valores perdidos definidos por el usuario. El resto de los valores de las demás variables originales, excepto los valores perdidos del sistema, se consideran válidos.
- Las asignaciones de valores de la plantilla se aplican en primer lugar. Los valores restantes se recodifican en valores superiores al último valor de la plantilla, con los valores perdidos definidos por el usuario (basados en la primera variable de la lista con valores perdidos definidos por el usuario) recodificados en valores superiores al último valor válido.
- Si ha seleccionado diversas variables para su recodificación automática, la plantilla se aplica en primer lugar, seguida de una recodificación automática común combinada para todos los valores adicionales encontrados en las variables

seleccionadas, lo que resulta en un único esquema de recodificación automática para todas las variables seleccionadas.

Para recodificar valores numéricos o de cadena en valores enteros consecutivos

- ▶ Elija en los menús:
Transformar
Recodificación automática...
- ▶ Seleccione la variable o variables que desee recodificar.
- ▶ Para cada variable seleccionada, introduzca un nombre para la nueva variable y pulse en Nuevo nombre.

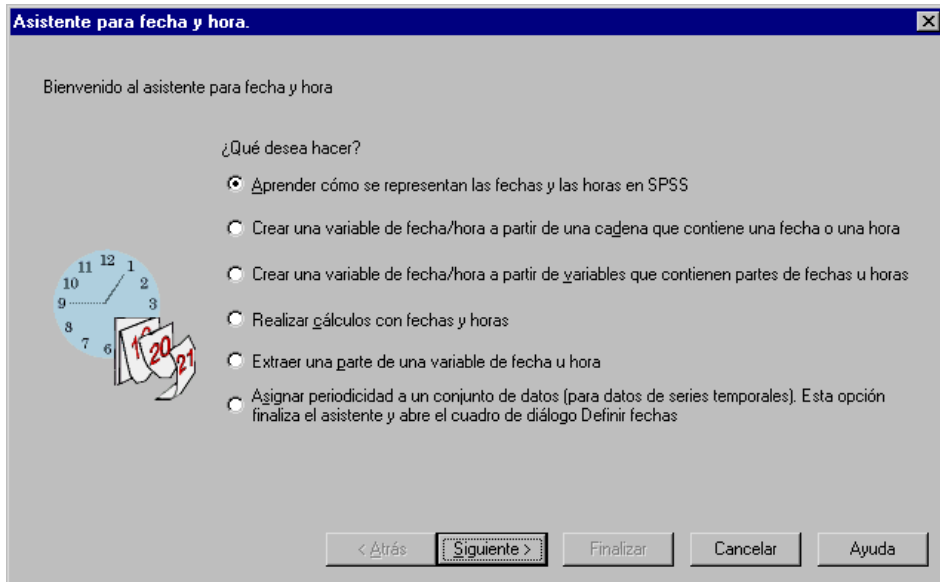
Asistente para fecha y hora

El Asistente para fecha y hora simplifica ciertas tareas comunes asociadas a las variables de fecha y hora.

Para usar el Asistente para fecha y hora

- ▶ Elija en los menús:
Transformar
Fecha/hora...
- ▶ Seleccione la tarea que desee realizar y siga los pasos para definir la tarea.

Figura 7-16
Pantalla principal del Asistente para fecha y hora



- **Aprender cómo se representan las fechas y las horas en SPSS.** Esta opción ofrece una pantalla en la que se presenta una breve descripción de las variables de fecha/hora en SPSS. El botón Ayuda también proporciona un vínculo para obtener información más detallada.
- **Crear una variable de fecha/hora a partir de una cadena que contiene una fecha o una hora.** Use esta opción para crear una variable de fecha/hora a partir de una variable de cadena. Por ejemplo, dispone de una variable de cadena que representa fechas con el formato mm/dd/aaaa y desea crear una variable de fecha/hora a partir de ella.
- **Crear una variable de fecha/hora fusionando variables que contengan partes diferentes de la fecha u hora.** Esta opción permite construir una variable de fecha/hora a partir de un conjunto de variables existentes. Por ejemplo, dispone de una variable que representa el mes (como un número entero), una segunda que representa el día del mes y una tercera que representa el año. Se pueden combinar estas variables en una única variable de fecha/hora.

- **Calcular con fechas y horas.** Use esta opción para añadir o sustraer valores a variables de fecha/hora. Por ejemplo, puede calcular la duración de un proceso sustrayendo una variable que represente la hora de comienzo del proceso de otra variable que represente la hora de finalización del proceso.
- **Extraer una parte de una variable de fecha/hora.** Esta opción permite extraer parte de una variable de fecha/hora, como el día del mes de una variable de fecha/hora, con el formato mm/dd/aaaa.
- **Asignar periodicidad a un conjunto de datos.** Esta opción presenta el cuadro de diálogo Definir fechas, que se usa para crear variables de fecha/hora compuestas por un conjunto de fechas secuenciales. Esta característica se usa generalmente para asociar fechas con datos de serie temporal.

Nota: Las tareas se desactivan cuando el conjunto de datos carece de los tipos de variables necesarias para completar la tarea. Por ejemplo, si el conjunto de datos no contiene variables de cadena, la tarea de creación de una variable de fecha/hora a partir de una cadena no se aplica y se desactiva.

Fechas y horas en SPSS

Las variables que representan fechas y horas en SPSS tienen un tipo de variable numérico, con formatos de presentación que se corresponden con los formatos específicos de fecha/hora. Estas variables se denominan generalmente variables de fecha/hora. SPSS distingue entre variables de fecha/hora que realmente representan fechas y aquellas que representan una duración temporal independiente de cualquier fecha, como 20 horas, 10 minutos y 15 segundos. Éstas últimas se denominan generalmente variables de duración, mientras que las primeras se conocen como variables de fecha o de fecha/hora. Para obtener una lista completa de los formatos de presentación, consulte el apartado sobre fecha y hora en la sección de elementos universales de la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Variables de fecha y de fecha/hora. Las variables de fecha tienen un formato que representa una fecha, como mm/dd/aaaa. Las variables de fecha/hora tienen un formato que representa una fecha y una hora, como dd-mmm-aaaa hh:mm:ss. Internamente, las variables de fecha y de fecha/hora se almacenan como el número de segundos transcurridos desde el 14 de octubre de 1582. Las variables de fecha y de fecha/hora se denominan a menudo variables con formato de fecha.

- Las especificaciones de año reconocidas son tanto de dos como de cuatro dígitos. Por defecto, los años representados por dos dígitos representan un intervalo que comienza 69 años antes de la fecha actual y finaliza 30 años después de la fecha actual. Este intervalo está determinado por la configuración de las Opciones y se puede modificar (en el menú Edición, seleccione Opciones y pulse en la pestaña Datos).
- Los delimitadores que se pueden usar en los formatos de día-mes-año son guiones, puntos, comas, barras y espacios en blanco.
- Los meses se pueden representar en dígitos, números romanos o abreviaturas de tres caracteres, y se pueden escribir con el nombre completo. Los nombres de los meses expresados con abreviaturas de tres letras y nombres completos deben estar en inglés, ya que no se reconocen los nombres de meses en otros idiomas.

Variables de duración. Las variables de duración tienen un formato que representa una duración de tiempo, como hh:mm. Se almacenan internamente como segundos sin hacer referencia a ninguna fecha en particular.

- En las especificaciones de tiempo (se aplican a las variables de fecha/hora y de duración), los dos puntos se pueden usar como delimitadores entre horas, minutos y segundos. Las horas y los minutos son valores necesarios, pero los segundos son opcionales. Para separar los segundos de las fracciones de segundo, es necesario utilizar un punto. Las horas pueden tener una magnitud ilimitada, pero el valor máximo de los minutos es 59 y el de los segundos, 59.999...

Fecha y hora actuales. La variable del sistema *\$TIME* contiene la fecha y hora actuales. Representa el número de segundos transcurridos desde el 14 de octubre de 1582 hasta la fecha y la hora en que se ejecute el comando de transformación que la use.

Creación de una variable de fecha/hora a partir de una cadena

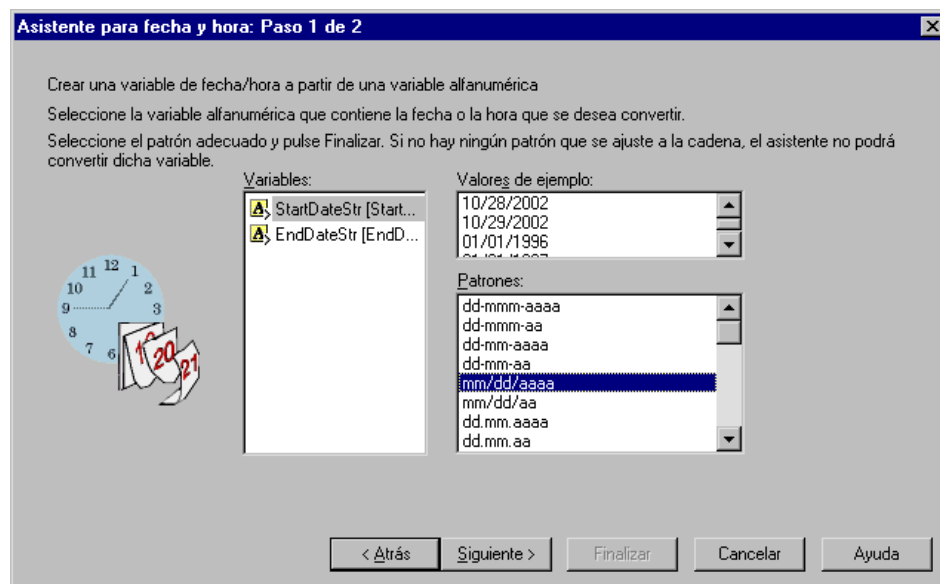
Para crear una variable de fecha/hora a partir de una variable de cadena:

- ▶ Seleccione Crear una variable de fecha/hora a partir de una variable de cadena que contenga una fecha u hora en la pantalla principal del Asistente para fecha y hora.

Selección de una variable de cadena para convertir en una variable de fecha/hora

Figura 7-17

Creación de una variable de fecha/hora a partir de una variable de cadena, paso 1

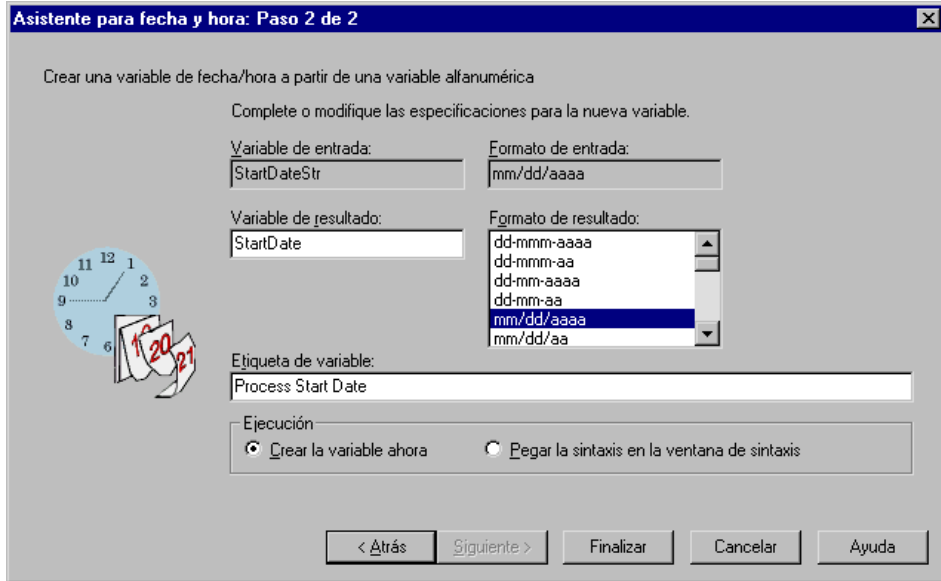


- ▶ En la lista Variables, seleccione la variable de cadena que desee convertir. Observe que la lista sólo contiene variables de cadena.
- ▶ En la lista Patrones, seleccione el patrón que coincida con el modo en que la variable de cadena representa las fechas. La lista Valores de ejemplo muestra los valores reales de la variable seleccionada en el archivo de datos. Los valores de la variable de cadena que no se ajusten al patrón seleccionado darán como resultado un valor perdido del sistema para la nueva variable.

Especificación del resultado de convertir la variable de cadena en variable de fecha/hora

Figura 7-18

Creación de una variable de fecha/hora a partir de una variable de cadena, paso 2



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.

Si lo desea, puede:

- Seleccionar un formato de fecha/hora para la nueva variable en la lista Formato de resultado.
- Asignar una etiqueta de variable descriptiva a la nueva variable.

Creación de una variable de fecha/hora a partir de un conjunto de variables

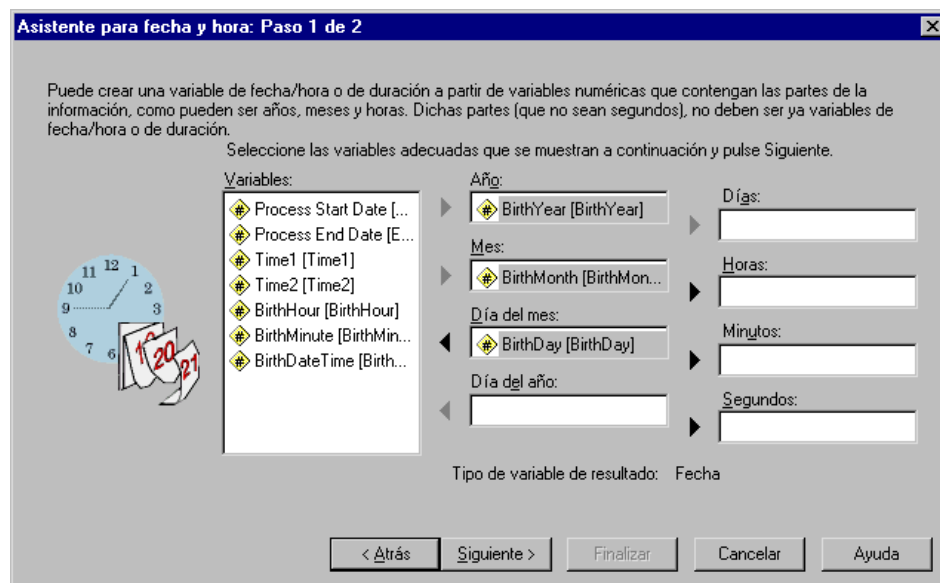
Para fusionar un conjunto de variables existentes en una única variable de fecha/hora:

- ▶ Seleccione Crear una variable de fecha/hora fusionando variables que contengan partes diferentes de la fecha u hora en la pantalla principal del Asistente para fecha y hora.

Selección de variables que fusionar en una única variable de fecha/hora

Figura 7-19

Creación de una variable de fecha/hora a partir de un conjunto de variables, paso 1



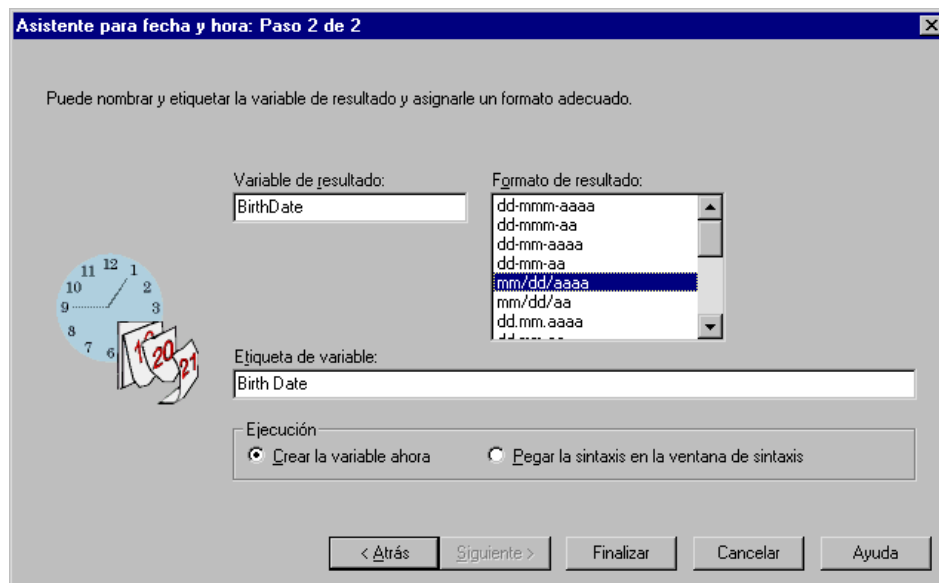
- ▶ Seleccione las variables que representen las diferentes partes de la fecha/hora.
 - Algunas combinaciones de selecciones no están permitidas. Por ejemplo, la creación de una variable de fecha/hora a partir de un valor de Año y Día del mes no es válida porque, una vez seleccionado Año, es necesario especificar una fecha completa.
 - No se puede utilizar una variable de fecha/hora existente como una de las partes de la variable de fecha/hora final que se está creando. Las variables que componen las partes de la nueva variable de fecha/hora deben ser números enteros. La excepción es el uso permitido de una variable de fecha/hora existente como la parte de los segundos de la nueva variable. Puesto que se permite el uso de fracciones de segundos, las variables utilizadas para los segundos no tiene que ser obligatoriamente un número entero.
 - Los valores de cualquier parte de la nueva variable que no se ajusten al rango permitido darán como resultado un valor perdido del sistema para la nueva variable. Esto podría suceder, por ejemplo, si se usa inadvertidamente una variable que representa un día del mes como valor de Mes. Puesto que el rango

válido para los meses en SPSS es 1–13, todos los casos en que el valor del día del mes pertenezca al rango 14–31 se considerarán valores perdidos del sistema para la nueva variable.

Especificación de variable de fecha/hora creada fusionando variables

Figura 7-20

Creación de una variable de fecha/hora a partir de un conjunto de variables, paso 2



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- ▶ Seleccione un formato de fecha/hora de la lista Formato de resultado.
Si lo desea, puede:
 - Asignar una etiqueta de variable descriptiva a la nueva variable.

Adición o sustracción de valores a partir de variables de fecha/hora

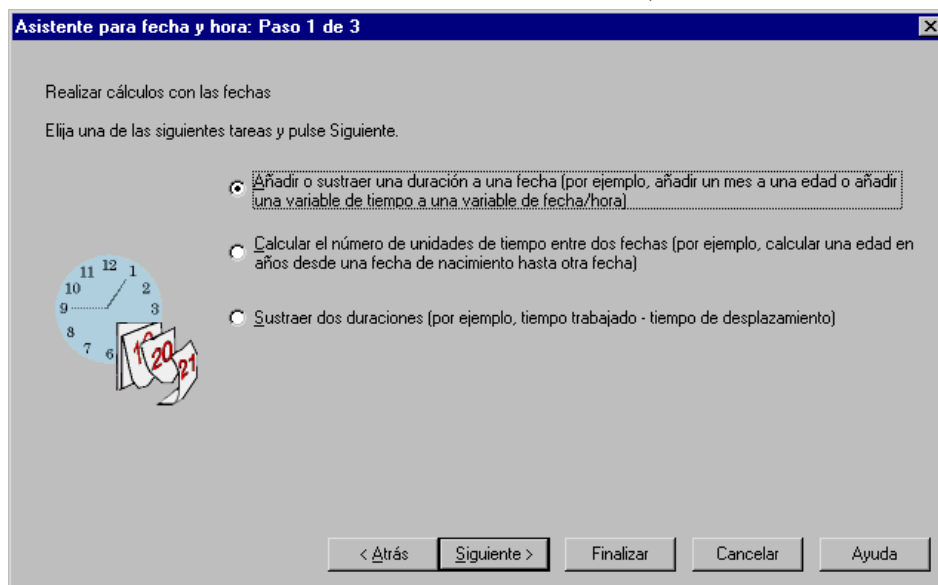
Para añadir o sustraer valores a variables de fecha/hora:

- Seleccione Calcular con fechas y horas en la pantalla principal del Asistente para fecha y hora.

Selección del tipo de cálculo que realizar con las variables de fecha/hora

Figura 7-21

Adición o sustracción de valores de variables de fecha/hora, paso 1



- **Añadir o sustraer una duración a una fecha.** Use esta opción para añadir o sustraer valores a una variable con formato de fecha. Si lo desea, puede añadir o sustraer duraciones que sean valores fijos, como 10 días, o los valores de una variable numérica (por ejemplo, una variable que represente años).

- **Calcular el número de unidades de tiempo entre dos fechas.** Use esta opción para obtener la diferencia entre dos fechas medidas en una unidad seleccionada. Por ejemplo, puede obtener el número de años o el número de días que separan dos fechas.
- **Sustraer dos duraciones.** Use esta opción para obtener la diferencia entre dos variables con formatos de duración, como hh:mm o hh:mm:ss.

Nota: Las tareas se desactivan cuando el conjunto de datos carece de los tipos de variables necesarias para completar la tarea. Por ejemplo, si el conjunto de datos no contiene dos variables con formatos de duración, la tarea de sustracción de dos duraciones no se aplica y se desactiva.

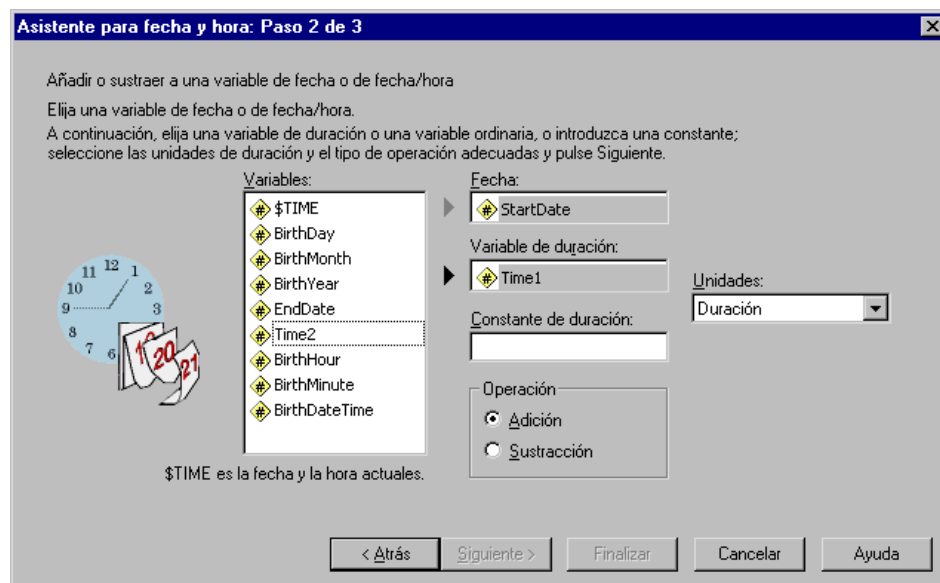
Adición o sustracción de una duración a una fecha

Para añadir o sustraer una duración a una variable con formato de fecha:

- ▶ Seleccione Añadir o sustraer una duración a una fecha en la pantalla del Asistente para fecha y hora denominada Realizar cálculos con las fechas.

Selección de variable de fecha/hora y duración que añadir o sustraer

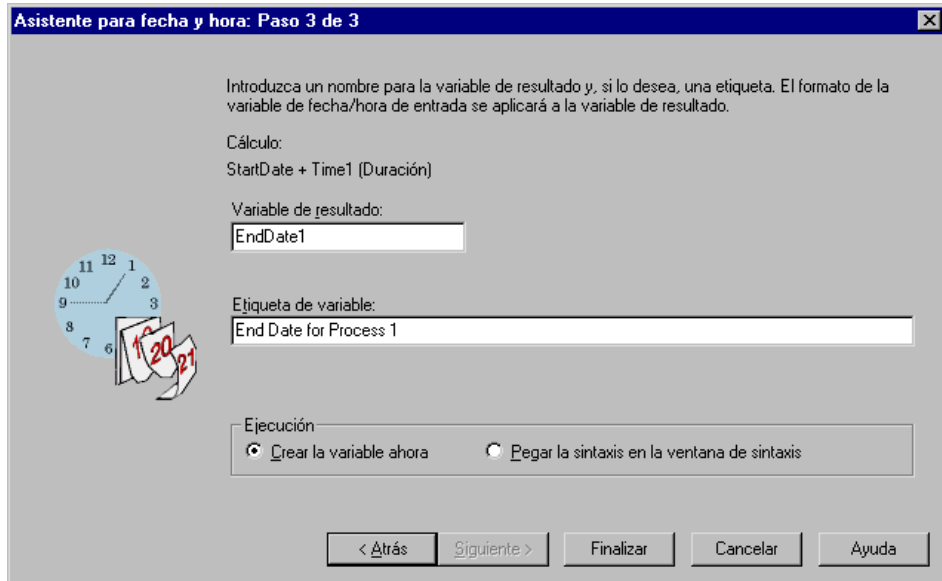
Figura 7-22
Adición o sustracción de duración, paso 2



- ▶ Seleccione una variable de fecha (u hora).
- ▶ Seleccione una variable de duración o especifique un valor para Constante de duración. Las variables utilizadas para las duraciones no pueden ser variables de fecha o de fecha/hora. Pueden ser variables de duración o variables numéricas simples.
- ▶ Seleccione la unidad que represente la duración en la lista desplegable. Seleccione Duración si se usa una variable y ésta tiene el formato de una duración, como hh:mm o hh:mm:ss.

Especificación de los resultados de la adición o sustracción de una duración a una variable de fecha/hora

Figura 7-23
Adición o sustracción de duración, paso 3



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.

Si lo desea, puede:

- Asignar una etiqueta de variable descriptiva a la nueva variable.

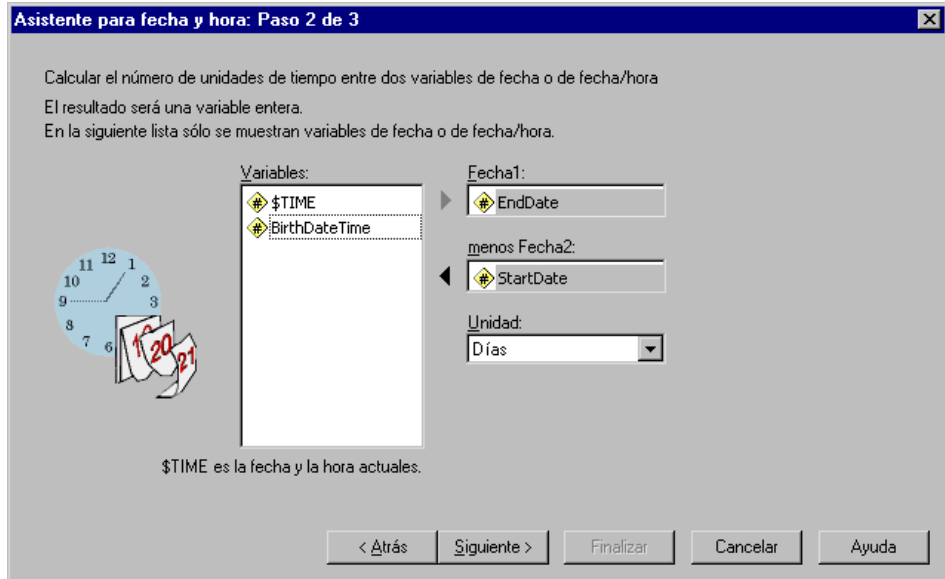
Sustracción de variables con formato de fecha

Para sustraer dos variables con formato de fecha:

- ▶ Seleccione Calcular el número de unidades de tiempo entre dos fechas en la pantalla del Asistente para fecha y hora denominada Realizar cálculos con las fechas.

Selección de variables con formato de fecha que sustraer

Figura 7-24
Sustracción de fechas, paso 2



- ▶ Seleccione las variables que se van a sustraer.
- ▶ Seleccione la unidad del resultado en la lista desplegable.

Especificación del resultado de la sustracción de dos variables con formato de fecha

Figura 7-25
Sustracción de fechas, paso 3

Asistente para fecha y hora: Paso 3 de 3

Cálculo: StartDate - EndDate

Variable de resultado: DurDays

Unidades: Días

Etiqueta de variable: Duration in Days

Ejecución

Crear la variable ahora Pegar la sintaxis en la ventana de sintaxis

< Atrás Siguiete > Finalizar Cancelar Ayuda

- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.

Si lo desea, puede:

- Asignar una etiqueta de variable descriptiva a la nueva variable.

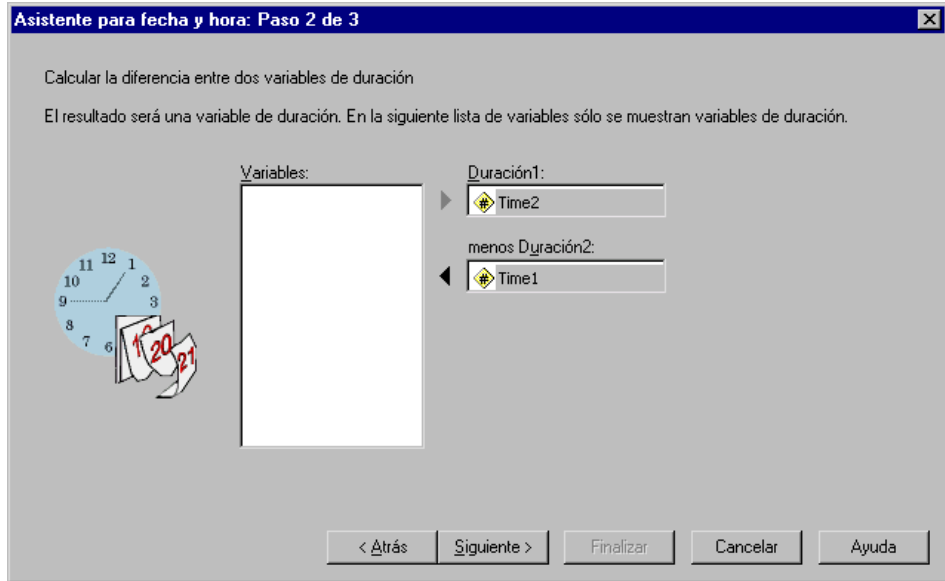
Sustracción de variables de duración

Para sustraer dos variables de duración:

- ▶ Seleccione Sustraer dos duraciones en la pantalla del Asistente para fecha y hora denominada Realizar cálculos con las fechas.

Selección de las variables de duración que sustraer

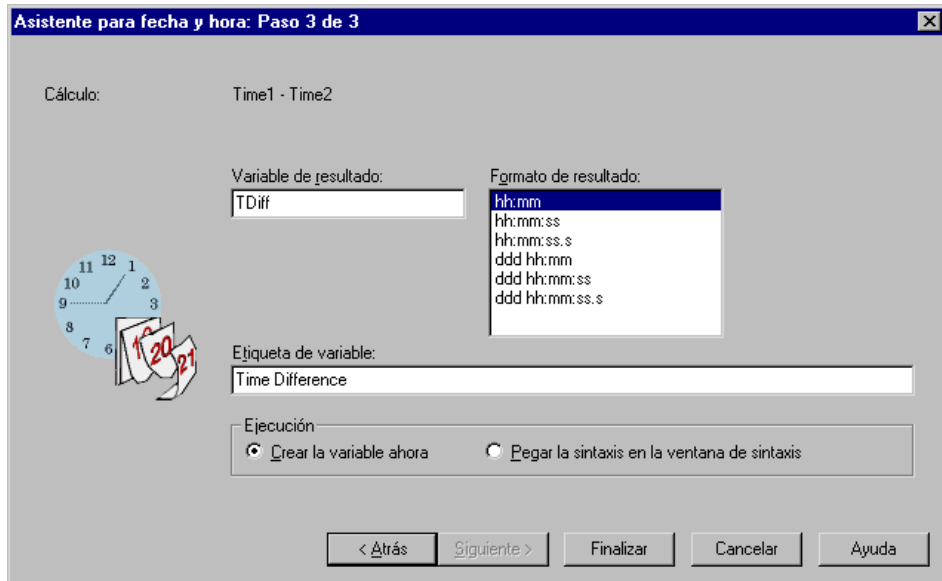
Figura 7-26
Sustracción de duraciones, paso 2



- ▶ Seleccione las variables que se van a sustraer.

Especificación del resultado de la sustracción de dos variables de duración

Figura 7-27
Sustracción de duraciones, paso 3



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- ▶ Seleccione un formato de duración de la lista Formato de resultado.
Si lo desea, puede:
 - Asignar una etiqueta de variable descriptiva a la nueva variable.

Extracción de parte de una variable de fecha/hora

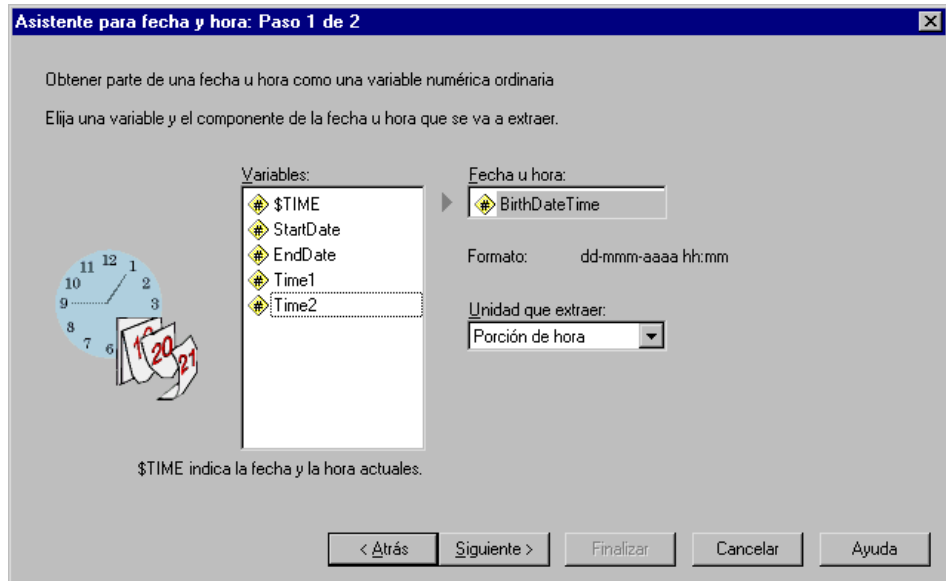
Para extraer un componente, como puede ser el año, de una variable de fecha/hora:

- ▶ Seleccione Extraer una parte de una variable de fecha u hora en la pantalla principal del Asistente para fecha y hora.

Selección de componente que extraer de una variable de fecha/hora

Figura 7-28

Extracción de parte de una variable de fecha/hora, paso 1

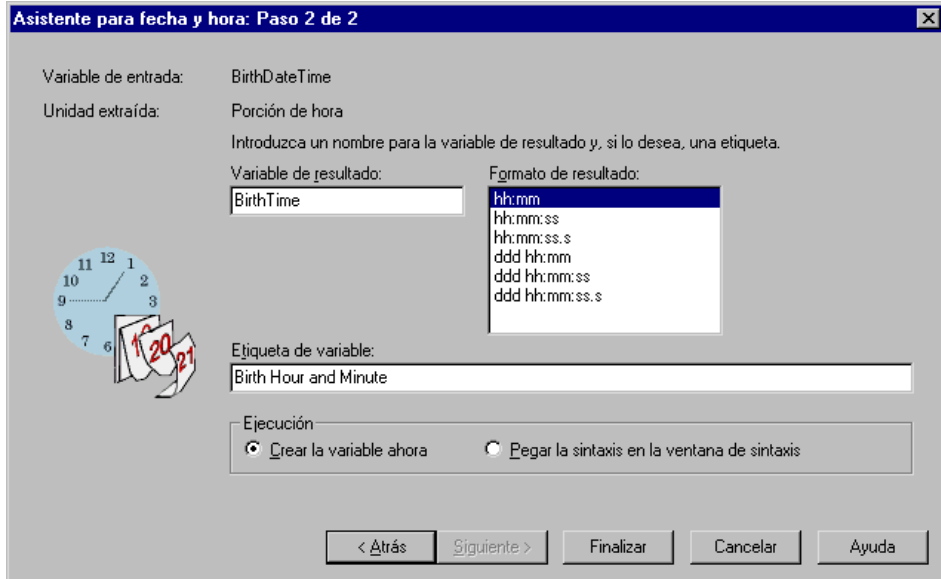


- Seleccione la variable que contiene la parte de fecha u hora que desee extraer.
- En la lista desplegable, seleccione la parte de la variable que se va a extraer. Si lo desea, puede extraer información de fechas que no sea explícitamente parte de la fecha que se muestra, por ejemplo, un día de la semana.

Especificación del resultado de la extracción de un componente de una variable de fecha/hora

Figura 7-29

Extracción de parte de una variable de fecha/hora, paso 2



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- ▶ Si está extrayendo la parte de fecha o de hora de una variable de fecha/hora, debe seleccionar un formato de la lista Formato de resultado. En los casos en que el formato de resultado no es necesario se desactivará la lista Formato de resultado.

Si lo desea, puede:

- Asignar una etiqueta de variable descriptiva a la nueva variable.

Transformaciones de los datos de serie temporal

Se incluyen diversas transformaciones de datos de gran utilidad en los análisis de series temporales:

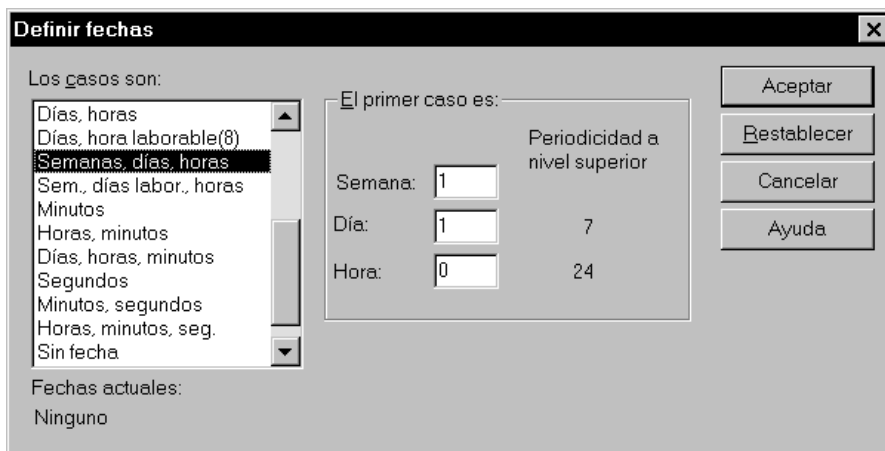
- Generar variables de fecha para establecer la periodicidad y distinguir entre los períodos históricos, de validación y de predicción.
- Elaborar nuevas variables de series temporales como funciones de variables de series temporales existentes.
- Reemplazar valores definidos como perdidos por el usuario y perdidos del sistema con estimaciones basadas en uno de los diversos métodos existentes.

Una **serie temporal** se obtiene midiendo una variable (o un conjunto de variables) de manera regular a lo largo de un período de tiempo. Las transformaciones de los datos de serie temporal suponen una estructura de archivo de datos en la que cada caso (fila) representa un conjunto de observaciones para un momento diferente y la duración del tiempo entre los casos es uniforme.

Definir fechas

El cuadro de diálogo Definir fechas genera variables de fecha que se pueden utilizar para establecer la periodicidad de una **serie temporal** y para etiquetar los resultados de los análisis de series temporales.

Figura 7-30
Cuadro de diálogo Definir fechas



Los casos son. Define el intervalo de tiempo utilizado para generar las fechas.

- Sin fecha elimina las variables de fecha definidas anteriormente. Se eliminarán todas las variables con los nombres siguientes: *año_*, *trimestre_*, *mes_*, *semana_*, *día_*, *hora_*, *minuto_*, *segundo_* y *fecha_*.
- Personalizado indica la presencia de variables de fecha personalizadas, creadas con la sintaxis de comandos (por ejemplo, una semana de cuatro días laborables). Este elemento simplemente refleja el estado actual del archivo de datos de trabajo. Su selección en la lista no produce ningún efecto. (Consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*) si desea obtener más información sobre el uso del comando `DATE` para crear variables de fecha personalizadas.)

El primer caso es. Define el valor de la fecha inicial, que se asigna al primer caso. A los casos subsiguientes se les asignan valores secuenciales, basándose en el intervalo de tiempo.

Periodicidad a nivel superior. Indica la variación cíclica repetitiva, como el número de meses de un año o el número de días de la semana. El valor mostrado indica el valor máximo que se puede introducir.

Para cada componente utilizado para definir la fecha, se crea una nueva variable numérica. Los nombres de las nuevas variables terminan con un carácter de subrayado. A partir de los componentes también se crea una variable de cadena descriptiva, *fecha_*. Por ejemplo, si selecciona Semanas, días, horas, se crearán cuatro nuevas variables: *semana_*, *día_*, *hora_* y *fecha_*.

Si ya se han definido variables de fecha, éstas serán reemplazadas cuando se definan nuevas variables de fecha con los mismos nombres que las existentes.

Para definir fechas para los datos de serie temporal

- ▶ Elija en los menús:
 - Datos
 - Definir fechas...
- ▶ Seleccione un intervalo de tiempo en la lista Los casos son.
- ▶ Introduzca el valor o los valores que definen la fecha inicial en El primer caso es, que determina la fecha asignada al primer caso.

Variables de fecha frente a variables con formato de fecha

Las variables de fecha creadas con Definir fechas no deben confundirse con las variables con formato de fecha, que se definen en Vista de variables del Editor de datos. Las variables de fecha se emplean para establecer la periodicidad de los datos de serie temporal; mientras que las variables con formato de fecha representan fechas y horas mostradas en varios formatos de fecha y hora. Las variables de fecha son números enteros sencillos que representan el número de días, semanas, horas, etc., a partir de un punto inicial especificado por el usuario. Internamente, la mayoría de las variables con formato de fecha se almacenan como el número de segundos transcurridos desde el 14 de octubre de 1582.

Crear serie temporal

El cuadro de diálogo Crear serie temporal crea nuevas variables basadas en funciones de variables de **series temporales** numéricas existentes. Estos valores transformados son de gran utilidad en muchos procedimientos de análisis de series temporales.

Los nombres por defecto de las nuevas variables se componen de los seis primeros caracteres de las variables existentes utilizadas para crearlas, seguidos por un carácter de subrayado y un número secuencial. Por ejemplo, para la variable *precio*, el nombre de la nueva variable sería *precio_1*. Las nuevas variables conservarían cualquier etiqueta de valor definida de las variables originales.

Las funciones disponibles para crear variables de series temporales incluyen las funciones de diferencias, medias móviles, medianas móviles, retardo y adelanto.

Figura 7-31
Cuadro de diálogo Crear serie temporal



Para crear una nueva variable de serie temporal

- ▶ Elija en los menús:
Transformar
Crear serie temporal...
- ▶ Seleccione la función de serie temporal que desea utilizar para transformar la variable o variables originales.
- ▶ Seleccione la variable o variables a partir de las cuales desee crear nuevas variables de serie temporal. Sólo se pueden utilizar variables numéricas.

Si lo desea, tiene la posibilidad de:

- Introducir nombres de variables, para omitir los nombres por defecto de las nuevas variables.
- Cambiar la función para una variable seleccionada.

Funciones de transformación de series temporales

Diferencia. Diferencia no estacional entre valores sucesivos de la serie. El orden es el número de valores previos utilizados para calcular la diferencia. Dado que se pierde una observación para cada orden de diferencia, aparecerán valores perdidos del sistema al comienzo de la serie. Por ejemplo, si el orden de diferencia es 2, los primeros dos casos tendrán el valor perdido del sistema para la nueva variable.

Diferencia estacional. Diferencia los valores de la serie respecto a los valores de la propia serie distanciados un orden (un lapso) de valores constante. El orden se basa en la periodicidad definida actualmente. Para calcular diferencias estacionales debe haber definido variables de fecha (menú Datos, Definir fechas) que incluyan un componente estacional (como por ejemplo los meses del año). El orden es el número de períodos estacionales utilizados para calcular la diferencia. El número de casos con el valor perdido del sistema al comienzo de la serie es igual a la periodicidad multiplicada por el orden de la diferencia estacional. Por ejemplo, si la periodicidad actual es 12 y el orden es 2, los primeros 24 casos tendrán el valor perdido del sistema para la nueva variable.

Media móvil centrada. Se utiliza el promedio de un rango de los valores de la serie, que rodean e incluyen al valor actual. La amplitud es el número de valores de la serie utilizados para calcular el promedio. Si la amplitud es par, la media móvil se calcula con el promedio de cada par de medias no centradas. Número de casos con el valor perdido del sistema al comienzo y al final de la serie para una amplitud de n es igual a $n/2$ para los valores de la amplitud par y para los valores de la amplitud impar. Por ejemplo, si la amplitud es 5, el número de casos con el valor perdido del sistema al comienzo y al final de la serie es 2.

Media móvil anterior. Se utiliza el promedio de un rango de las observaciones precedentes. La amplitud es el número de valores precedentes de la serie utilizados para calcular el promedio. El número de casos con el valor perdido del sistema al comienzo de la serie es igual al valor de la amplitud.

Medianas móviles. Se utiliza la mediana de un rango de los valores de la serie, que rodean e incluyen al valor actual. La amplitud es el número de valores de la serie utilizados para calcular la mediana. Si la amplitud es par, la mediana se calcula con el promedio de cada par de medianas no centradas. Número de casos con el valor perdido del sistema al comienzo y al final de la serie para una amplitud de n es igual a $n/2$ para los valores de la amplitud par y para los valores de la amplitud impar. Por

ejemplo, si la amplitud es 5, el número de casos con el valor perdido del sistema al comienzo y al final de la serie es 2.

Suma acumulada. Cada valor de la serie se sustituye por la suma acumulada de los valores precedentes, incluyendo el valor actual.

Retardo. Cada valor de la serie se sustituye por el valor del caso precedente, en el orden especificado. El orden especifica a qué distancia se encuentra el caso precedente. El número de casos con el valor perdido del sistema al comienzo de la serie es igual al valor del orden.

Adelanto. Cada valor de la serie se sustituye por el valor de un caso posterior, en el orden especificado. El orden especifica a qué distancia se encuentra el caso posterior. El número de casos con el valor perdido del sistema al final de la serie es igual al valor del orden.

Suavizado. Los nuevos valores de la serie se basan en un suavizador de datos compuesto. El suavizador comienza con una mediana móvil de 4, que se centra por una mediana móvil de 2. A continuación, se vuelven a suavizar estos valores aplicando una mediana móvil de 5, una mediana móvil de 3 y los promedios ponderados móviles (hanning). Los residuos se calculan sustrayendo la serie suavizada de la serie original. Después se repite todo el proceso sobre los residuos calculados. Por último, los residuos suavizados se calculan sustrayendo los valores suavizados obtenidos la primera vez que se realizó el proceso. A esto se le denomina a veces suavizado T4253H.

Reemplazar los valores perdidos

Las observaciones perdidas pueden causar problemas en los análisis y algunas medidas de series temporales no se pueden calcular si hay valores perdidos en la serie. En ocasiones el valor para una observación concreta no se conoce. Además, los datos perdidos pueden ser el resultado de lo siguiente:

- Cada grado de diferenciación reduce la longitud de una serie en 1.
- Cada grado de diferenciación estacional reduce la longitud de una serie en una estación.

- Si genera una serie nueva que contenga predicciones que sobrepasen el final de la serie existente (al pulsar en el botón Guardar y realizar las selecciones adecuadas), la serie original y la serie residual generada incluirán datos perdidos para las observaciones nuevas.
- Algunas transformaciones (por ejemplo, la transformación logarítmica) generan datos perdidos para determinados valores de la serie original.

Los valores perdidos al principio o fin de una serie no suponen un problema especial; sencillamente acortan la longitud útil de la serie. Las discontinuidades que aparecen en mitad de una serie (datos *incrustados* perdidos) pueden ser un problema mucho más grave. El alcance del problema depende del procedimiento analítico que se utilice.

El cuadro de diálogo Reemplazar valores perdidos crea nuevas variables de **series temporales** a partir de otras existentes, reemplazando los valores perdidos por estimaciones calculadas mediante uno de los distintos métodos posibles. Los nombres por defecto de las nuevas variables se componen de los seis primeros caracteres de las variables existentes utilizadas para crearlas, seguidos por un carácter de subrayado y un número secuencial. Por ejemplo, para la variable *precio*, el nombre de la nueva variable sería *precio_1*. Las nuevas variables conservarían cualquier etiqueta de valor definida de las variables originales.

Figura 7-32

Cuadro de diálogo Reemplazar los valores perdidos



Para reemplazar los valores perdidos para las variables de series temporales

- ▶ Elija en los menús:
 - Transformar
 - Reemplazar valores perdidos...
- ▶ Seleccione el método de estimación que desee utilizar para reemplazar los valores perdidos.
- ▶ Seleccione la variable o variables para las que desea reemplazar los valores perdidos.
Si lo desea, tiene la posibilidad de:
 - Introducir nombres de variables, para omitir los nombres por defecto de las nuevas variables.
 - Cambiar el método de estimación para una variable seleccionada.

Métodos de estimación para reemplazar los valores perdidos

Media de la serie. Sustituye los valores perdidos con la media de la serie completa.

Media de puntos adyacentes. Sustituye los valores perdidos por la media de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos, por encima y por debajo del valor perdido, utilizados para calcular la media.

Mediana de puntos adyacentes. Sustituye los valores perdidos por la mediana de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos, por encima y por debajo del valor perdido, utilizados para calcular la mediana.

Interpolación lineal. Sustituye los valores perdidos utilizando una interpolación lineal. Se utilizan para la interpolación el último valor válido antes del valor perdido y el primer valor válido después del valor perdido. Si el primer o el último caso de la serie tiene un valor perdido, el valor perdido no se sustituye.

Tendencia lineal en el punto. Reemplaza los valores perdidos de la serie por la tendencia lineal en ese punto. Se hace una regresión de la serie existente sobre una variable índice escalada de 1 a n . Los valores perdidos se sustituyen por sus valores pronosticados.

Puntuación de datos con modelos predictivos

El proceso de aplicar un modelo predictivo a un conjunto de datos se denomina **puntuación** de los datos. SPSS, Clementine y AnswerTree disponen de procedimientos para la creación de modelos predictivos como los modelos de regresión, conglomeración, árbol y red neuronal. Una vez creado un modelo, sus especificaciones se pueden guardar como un archivo XML que contenga toda la información necesaria para reconstruirlo. La versión de servidor de SPSS permite leer un archivo de modelo XML y aplicar el modelo a un archivo de datos.

Ejemplo. El riesgo de una solicitud de crédito se evalúa basándose en diversos aspectos del solicitante y el préstamo en cuestión. La puntuación del crédito obtenida a partir del modelo de riesgo se usa para aceptar o rechazar la solicitud de préstamo.

La puntuación se trata como una transformación de los datos. El modelo se expresa internamente como un conjunto de transformaciones numéricas que se deben aplicar a un determinado conjunto de variables, las variables predictoras especificadas en el modelo, con el fin de obtener un resultado predictivo. En este sentido, el proceso de puntuación de los datos con un modelo dado es, inherentemente, igual que la aplicación de cualquier función, como puede ser una función de raíz cuadrada, a un conjunto de datos.

El método de puntuación sólo está disponible en la versión de servidor de SPSS y requiere el uso de sintaxis de comandos de SPSS. Los comandos necesarios se pueden introducir en la ventana del editor de sintaxis y ejecutar de forma interactiva por parte de usuarios que trabajen en el análisis en modo distribuido. El conjunto de comandos también se puede guardar en un archivo de sintaxis de comandos y enviar a SPSS Batch Facility, una versión ejecutable independiente de SPSS que se distribuye junto con la versión de servidor de SPSS. En el caso de archivos de datos de gran tamaño, probablemente sea preciso usar SPSS Batch Facility. Para obtener información acerca del uso de SPSS Batch Facility, consulte el manual del usuario de SPSS Batch Facility (*SPSS Batch Facility User's Guide*), que se incluye como archivo PDF en el CD del producto de la versión de servidor de SPSS.

La sintaxis necesaria para la puntuación incluye el comando `MODEL HANDLE` y la función `ApplyModel` o `StrApplyModel`.

- El comando `MODEL HANDLE` se usa para leer el archivo XML que contiene las especificaciones del modelo. Almacena en caché las especificaciones del modelo y asocia un nombre exclusivo con el modelo en caché. Para obtener más detalles,

consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

- Las funciones `ApplyModel` o `StrApplyModel` se usan con el comando `COMPUTE` para aplicar el modelo. Para obtener más detalles, consulte las expresiones de puntuación en el apartado sobre expresiones de transformación de la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Gestión y transformación de los archivos

Los archivos de datos no siempre están organizados de la forma ideal para las necesidades específicas del usuario. Puede que le interese combinar archivos de datos, organizar los datos en un orden diferente, seleccionar un subconjunto de casos o cambiar la unidad de análisis agrupando casos. Entre la amplia gama de posibilidades de transformación de archivos disponibles se encuentran las siguientes:

Ordenar datos. Puede ordenar los casos en función del valor de una o más variables.

Transponer casos y variables. El formato de archivo de datos de SPSS lee las filas como casos y las columnas como variables. Para los archivos de datos en los que el orden está invertido, se pueden intercambiar las filas y las columnas para leer los datos en el formato correcto.

Fundir archivos. Puede fundir dos o más archivos de datos. Es posible combinar archivos con las mismas variables pero con casos distintos, o con los mismos casos pero variables diferentes.

Seleccionar subconjuntos de casos. Puede restringir el análisis a un subconjunto de casos o efectuar análisis simultáneos de subconjuntos diferentes.

Agregar datos. Puede cambiar la unidad de análisis agregando casos basados en el valor de una o más variables de agrupación.

Ponderar datos. Puede ponderar los casos para un análisis basado en el valor de una variable de ponderación.

Reestructurar datos. Puede reestructurar los datos para crear un único caso (registro) a partir de varios casos o crear varios casos a partir de un único caso.

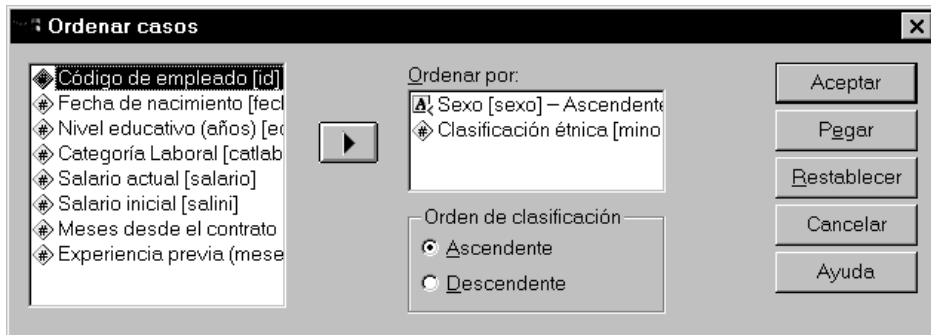
Ordenar casos

Este cuadro de diálogo ordena los casos (las filas) del archivo de datos basándose en los valores de una o más variables de ordenación. Puede ordenar los casos en orden ascendente o descendente.

- Si selecciona más de una variable de ordenación, los casos se ordenarán por variable dentro de las categorías de la variable anterior de la lista Ordenar por. Por ejemplo, si selecciona *Sexo* como la primera variable de ordenación y *Minoría* como la segunda, los casos se ordenarán por minorías dentro de cada categoría de sexo.
- Para las variables de cadena, las letras mayúsculas preceden a las minúsculas correspondientes en orden. Por ejemplo, el valor de cadena “Sí” precede a “sí”.

Figura 8-1

Cuadro de diálogo Ordenar casos



Para ordenar casos

- ▶ Elija en los menús:
Datos
Ordenar casos...
- ▶ Seleccione una o más variables de ordenación.

Transponer

Transponer crea un archivo de datos nuevo en el que se transponen las filas y las columnas del archivo de datos original de manera que los casos (las filas) se convierten en variables, y las variables (las columnas) se convierten en casos. También crea automáticamente nombres de variable y presenta una lista de dichos nombres.

- Se crea automáticamente una nueva variable de cadena, *case_lbl*, que contiene el nombre de variable original.
- Si el archivo de datos de trabajo contiene una variable de identificación o de nombre con valores únicos, podrá utilizarla como variable de nombre: sus valores se emplearán como nombres de variable en el archivo de datos transpuesto. Si se trata de una variable numérica, los nombres de variable comenzarán por la letra *V*, seguida de un valor numérico.
- Los valores perdidos definidos por el usuario se convierten en el valor perdido del sistema en el archivo de datos transpuesto. Para conservar cualquiera de estos valores, se debe cambiar la definición de los valores perdidos en la vista de variables del Editor de datos.

Para transponer variables y casos

- ▶ Elija en los menús:
 - Datos
 - Transponer...
- ▶ Seleccione la variable o variables que desee transponer en casos.

Fusión de archivos de datos

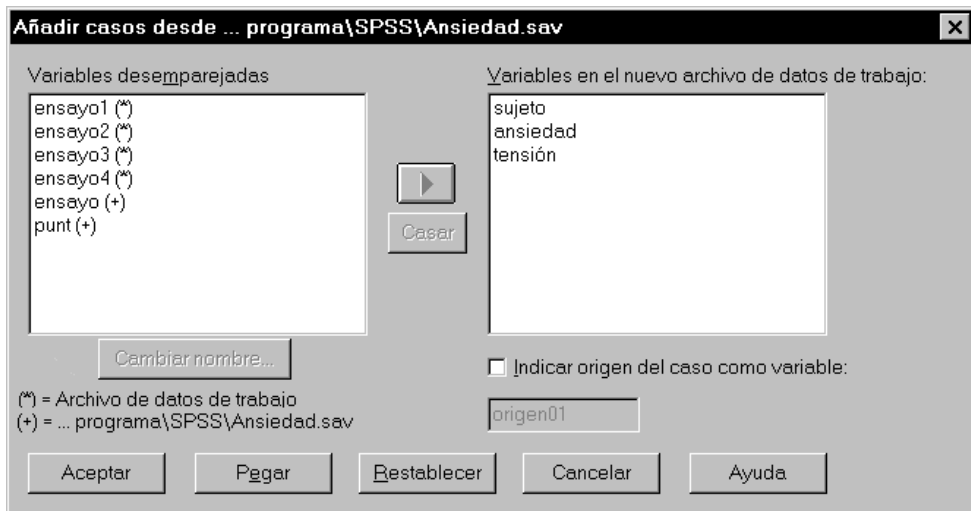
Es posible unir los datos de dos archivos de dos maneras diferentes. Tiene la posibilidad de:

- Fundir archivos que contengan las mismas variables pero casos diferentes.
- Fundir archivos que contengan los mismos casos pero variables diferentes.

Añadir casos

Añadir casos fusiona el archivo de datos de trabajo con otro archivo de datos que contiene las mismas variables pero diferentes casos. Por ejemplo, podría registrar la misma información de los clientes de dos zonas de venta diferentes y conservar los datos de cada zona en archivos distintos.

Figura 8-2
Cuadro de diálogo Añadir casos



Variables desemparejadas. Muestra las variables que se van a excluir del nuevo archivo de datos fusionado. Las variables del archivo de trabajo se identifican mediante un asterisco (*); las del archivo externo presentan un signo más (+). Las del archivo externo presentan un signo más (+). Por defecto, la lista contiene:

- Las variables de cualquiera de los archivos de datos que no coincidan con un nombre de variable del otro archivo. Puede crear pares a partir de variables desemparejadas e incluirlos en el nuevo archivo fusionado.
- Las variables definidas como datos numéricos en un archivo y como datos de cadena en el otro. Las variables numéricas no pueden fusionarse con variables de cadena.
- Variables de cadena de longitud diferente. La longitud definida de una variable de cadena debe ser la misma en ambos archivos de datos.

Variables en el nuevo archivo de datos de trabajo. Variables que se van a incluir en el nuevo archivo de datos fusionado. Por defecto, la lista incluye todas las variables que coinciden en el nombre y el tipo de datos (numéricos o de cadena).

- Puede eliminar de la lista las variables que no desee incluir en el archivo fusionado.
- Las variables desemparejadas incluidas en el archivo fusionado contendrán los datos perdidos para los casos del archivo que no contiene esa variable.

Indicar origen del caso como variable. Indica, para cada caso, el archivo de datos de origen. Esta variable toma un valor 0 para los casos del archivo de datos de trabajo y un valor 1 para los casos del archivo de datos externo.

Para fundir archivos de datos con las mismas variables y casos diferentes

- ▶ Abra uno de los archivos de datos. Los casos de este archivo aparecerán primero en el nuevo archivo de datos fusionado.
- ▶ Elija en los menús:
 - Datos
 - Fundir archivos
 - Añadir casos...
- ▶ Seleccione el archivo de datos que desee fundir con el archivo de datos abierto.
- ▶ Elimine de la lista Variables en el nuevo archivo de datos de trabajo las variables que no desee incluir.
- ▶ Añada parejas de variables de la lista Variables desemparejadas que representen la misma información registrada con nombres diferentes en los dos archivos. Por ejemplo, la fecha de nacimiento podría tener el nombre de variable *fechnac* en un archivo y *nacfech* en el otro.

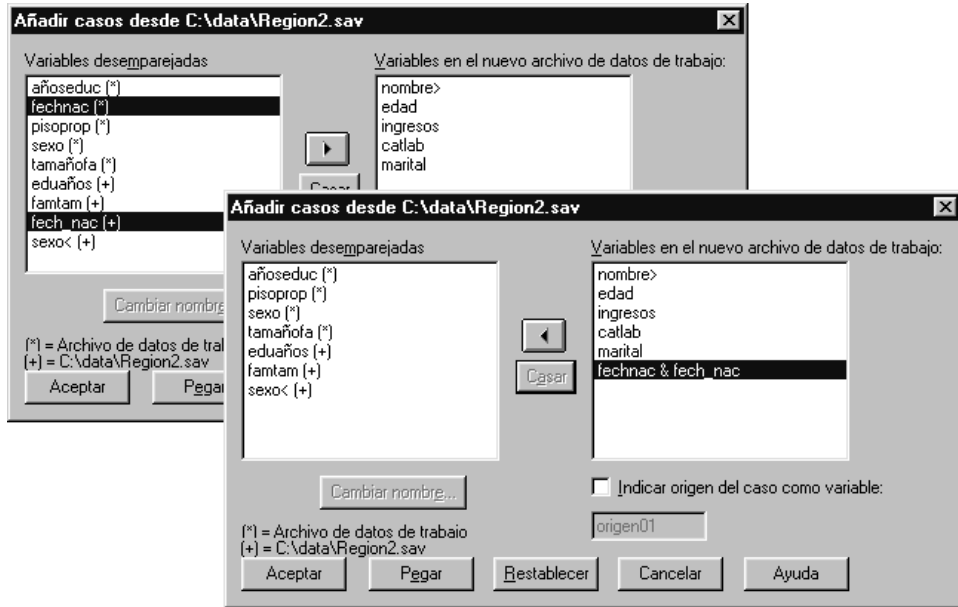
Para seleccionar una pareja de variables desemparejadas

- ▶ Pulse en una de las variables en la lista Variables desemparejadas.
- ▶ Mantenga pulsada la tecla Ctrl mientras selecciona la otra variable de la lista con el ratón (pulse al mismo tiempo la tecla Ctrl y el botón izquierdo del ratón).

- Pulse en **Casar** para desplazar el par de variables a la lista Variables en el nuevo archivo de datos de trabajo. el nombre de variable del archivo de datos de trabajo se empleará como el nombre de variable del archivo fusionado.

Figura 8-3

Selección de parejas de variables mediante la tecla Ctrl



Añadir casos: Cambiar nombre

Puede cambiar los nombres de las variables del archivo de datos de trabajo o los del archivo externo antes de desplazarlas desde la lista de variables desparejadas a la lista de variables que se van a incluir en el archivo fusionado. Cambiar el nombre de las variables le permite:

- Utilizar el nombre de variable del archivo externo y no del archivo de trabajo para las parejas de variables.
- Incluir dos variables con el mismo nombre pero de diferentes tipos o longitudes de cadena. Por ejemplo, para incluir la variable numérica *sexo* del archivo de datos de trabajo y la variable de cadena *sexo* del archivo externo, se debe cambiar el nombre de una de ellas.

Añadir casos: Información del diccionario

Toda información del diccionario (etiquetas de variable y de valor, valores perdidos definidos por el usuario, formatos de presentación) existente en el archivo de datos de trabajo es aplicable al archivo de datos fusionado.

- Si alguna información del diccionario sobre una variable no está definida en el archivo de trabajo, se utilizará la información del diccionario del archivo de datos externo.
- Si el archivo de trabajo contiene etiquetas de valor o valores perdidos definidos por el usuario para una variable, se ignorará cualquier otra etiqueta de valor o valor perdido definido por el usuario para esa variable del archivo externo.

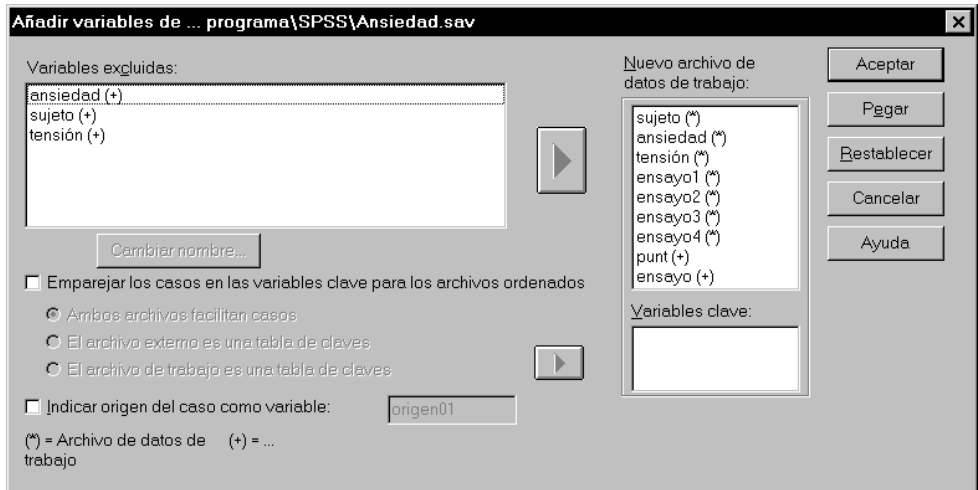
Añadir variables

Añadir variables fusiona el archivo de datos de trabajo con un archivo de datos externo que contiene los mismos casos pero variables diferentes. Por ejemplo, es posible que desee fusionar un archivo de datos que contenga los resultados previos de la prueba con otro que contenga los resultados posteriores.

- Los casos deberán tener el mismo orden en ambos archivos.
- Si se utiliza una o más variables clave para emparejar los casos, los dos archivos deben estar ordenados por orden ascendente de la variable o variables clave.
- Los nombres de las variables del segundo archivo que duplican los del archivo de trabajo se excluyen por defecto, ya que Añadir variables supone que estas variables contienen información duplicada.

Indicar origen del caso como variable. Indica, para cada caso, el archivo de datos de origen. Esta variable toma un valor 0 para los casos del archivo de datos de trabajo y un valor 1 para los casos del archivo de datos externo.

Figura 8-4
Cuadro de diálogo *Añadir variables*



Variabes excluidas. Muestra las variables que se van a excluir del nuevo archivo de datos fusionado. Por defecto, la lista contiene los nombres de variable del archivo externo que duplican los del archivo de trabajo. Las variables del archivo de trabajo se identifican mediante un asterisco (*); las del archivo externo presentan un signo más (+). Si desea incluir en el archivo fusionado una variable excluida con un nombre duplicado, cámbiele el nombre y añádala a la lista de variables que se van a incluir.

Nuevo archivo de datos de trabajo. Variables que se van a incluir en el nuevo archivo de datos fusionado. Por defecto, se incluyen en la lista todos los nombres de variable únicos que existan en ambos archivos de datos.

Variabes clave. Si algunos casos de un archivo no se emparejan con los del otro archivo (es decir, si faltan casos en un archivo), utilice las variables clave para identificar y emparejar correctamente los casos de ambos archivos. También puede utilizar las variables clave según tablas de claves.

- Las variables clave deben tener los mismos nombres en ambos archivos.

- Ambos archivos deben estar ordenados según el orden ascendente de las variables clave, y el orden de las variables de la lista Variables clave debe ser igual a su secuencia de ordenación.
- Los casos que no se emparejan con las variables clave se incluyen en el archivo fusionado, pero no se funden con los casos del otro archivo. Los casos no emparejados sólo contienen valores para las variables del archivo de procedencia; las variables del otro archivo contienen el valor perdido del sistema.

El archivo externo o el archivo de trabajo es una tabla de claves. Una tabla de claves, o **tabla de referencia**, es un archivo en el que los datos de cada “caso” se pueden aplicar a varios casos del otro archivo de datos. Por ejemplo, si un archivo contiene información sobre los diferentes miembros de la familia (como el sexo, la edad, la formación) y el otro contiene información global (como los ingresos totales, el número de miembros o la ubicación), se puede utilizar el archivo global como una tabla de referencia y aplicar los datos comunes de la familia a cada uno de sus miembros en el archivo fusionado.

Para fundir archivos con los mismos casos pero variables diferentes

- ▶ Abra uno de los archivos de datos.
- ▶ Elija en los menús:
 - Datos
 - Fundir archivos
 - Añadir variables...
- ▶ Seleccione el archivo de datos que desee fundir con el archivo de datos abierto.

Para seleccionar variables clave

- ▶ Seleccione las variables entre las variables del archivo externo (marcadas con el signo +) en la lista Variables excluidas.
- ▶ Seleccione Emparejar los casos en las variables clave para los archivos ordenados.
- ▶ Añada las variables a la lista Variables clave.

Las variables clave deben existir tanto en el archivo de trabajo como en el archivo externo. Ambos archivos deben estar ordenados según el orden ascendente de las variables clave, y el orden de las variables de la lista Variables clave debe ser igual a su secuencia de ordenación.

Añadir variables: Cambiar nombre

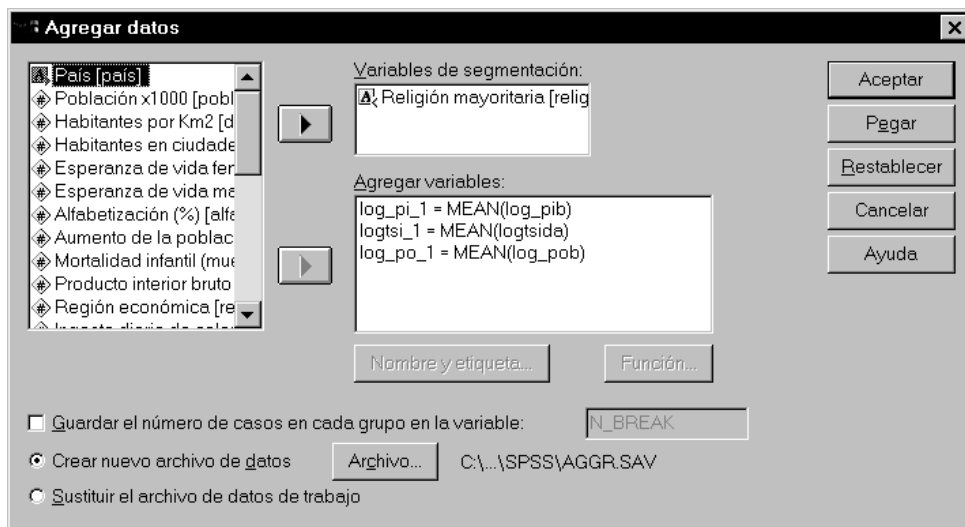
Puede cambiar los nombres de las variables del archivo de datos de trabajo o del archivo de datos externo antes de desplazarlas a la lista de variables que se van a incluir en el archivo de datos fusionado. Esta medida es especialmente útil cuando se desea incluir dos variables que tienen el mismo nombre pero información diferente en los dos archivos.

Agregar datos

Agregar datos agrega grupos de casos en el archivo de datos de trabajo para casos individuales y crea un archivo nuevo agregado o variables nuevas en el archivo de datos de trabajo que contiene los datos agregados. Los casos se agregan en función del valor de una o más variables de segmentación (agrupación).

- Si crea un archivo de datos agregado nuevo, dicho archivo de datos nuevo contiene un caso para cada grupo definido por las variables de segmentación. Por ejemplo, si hay una variable de segmentación con dos valores, el archivo de datos nuevo contiene sólo dos casos.
- Si añade variables agregadas al archivo de datos de trabajo, no se agrega el archivo de datos. Cada caso con los mismos valores de variables de segmentación recibe los mismos valores para las nuevas variables agregadas. Por ejemplo, si el sexo es la única variable de segmentación, todos los hombres reciben el mismo valor para la variable agregada nueva que representa la edad media.

Figura 8-5
Cuadro de diálogo Agregar datos



Variables de segmentación. Los casos se agrupan en función de los valores de las variables de segmentación. Cada combinación única de valores de variables de segmentación define un grupo. Al crear un archivo de datos agregados nuevo, todas las variables de segmentación se guardan en el archivo nuevo con sus nombres y la información del diccionario. La variable de segmentación puede ser tanto numérica como de cadena.

Agregar variables. Las variables de origen se utilizan con funciones agregadas para crear variables agregadas nuevas. El nombre de la variable agregada viene seguido de una etiqueta de variable opcional entrecomillada, el nombre de la función de agregación y el nombre de la variable de origen entre paréntesis. Para las funciones de agregación, las variables de origen deben ser numéricas.

Puede anular los nombres por defecto de las variables agregadas con nuevos nombres de variable, proporcionar etiquetas de variable descriptivas y cambiar las funciones empleadas para calcular los valores de los datos agregados. También puede crear una variable que contenga el número de casos en cada grupo de segmentación.

Para agregar un archivo de datos

- ▶ Elija en los menús:
Datos
Agregar...
- ▶ Seleccione una o más variables de segmentación que definan cómo deben agruparse los casos para crear datos agregados.
- ▶ Seleccione una o más variables agregadas.
- ▶ Seleccione una función de agregación para cada variable agregada.

Almacenamiento de resultados agregados

Puede añadir variables agregadas al archivo de datos de trabajo o crear un archivo de datos agregados nuevo.

- **Añadir las variables agregadas al archivo de datos de trabajo.** Las nuevas variables basadas en las funciones de agregación se añaden al archivo de datos de trabajo. El propio archivo de datos no se agrega. Cada caso con los mismos valores en las variables de segmentación recibe los mismos valores en las nuevas variables agregadas.
- **Crear nuevo archivo de datos que contenga sólo variables agregadas.** Crea un nuevo archivo de datos con los datos agregados. Por defecto, se guarda un archivo denominado `aggr.sav` en el directorio actual.
- **Sustituir el archivo de datos de trabajo sólo con variables agregadas.** Sustituye el archivo de datos de trabajo con el archivo de datos agregado. El archivo incluye las variables de segmentación que definen los casos agregados y todas las variables agregadas definidas por las funciones de agregación. Esto no afecta al archivo de datos original. El archivo de datos agregado no se guardará a menos que se guarde el archivo de datos de manera explícita (menú Archivo, Guardar o Guardar como).

Opciones de ordenación para archivos de datos grandes

En el caso de los archivos de datos muy grandes, puede resultar más útil agregar datos ordenados previamente.

El archivo ya está ordenado por las variables de segmentación. Si los datos ya han sido ordenados por los valores de las variables de segmentación, esta opción permite al procedimiento una ejecución más rápida y utilizar menos memoria. Utilice esta opción con precaución.

- Los datos se deben ordenar por valores de variables de segmentación en el mismo orden que las variables de segmentación especificadas para el procedimiento Agregar datos.
- Si va a añadir variables al archivo de datos de trabajo, seleccione sólo esta opción si los datos se han ordenado mediante valores ascendentes de las variables de segmentación.

Ordenar el archivo antes de la agregación. En situaciones muy extrañas y con archivos de datos voluminosos, puede ser necesario ordenar el archivo de datos por los valores de las variables de segmentación antes de realizar la agregación. No se recomienda esta opción a menos que se presenten problemas de memoria y/o rendimiento.

Agregar datos: Función de agregación

Este cuadro de diálogo permite especificar la función que se utilizará para calcular los valores de los datos agregados para las variables seleccionadas en la lista Agregar variables, en el cuadro de diálogo Agregar datos. Las funciones de agregación incluyen:

- Funciones de resumen, incluyendo la media, la mediana, la desviación típica y la suma.
- Número de casos, incluyendo los no ponderados, los ponderados, los no perdidos y los perdidos.
- Porcentaje o fracción de los valores por encima o por debajo de un valor especificado.
- Porcentaje o fracción de los valores dentro o fuera de un rango especificado.

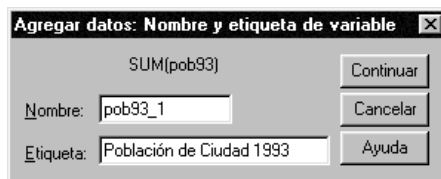
Figura 8-6
Cuadro de diálogo *Función de agregación*



Agregar datos: Nombre y etiqueta de variable

Agregar datos asigna nombres de variable por defecto a las variables agregadas al nuevo archivo de datos. Este cuadro de diálogo le permite cambiar el nombre de variable de la variable seleccionada en la lista Agregar variables y proporcionar una etiqueta de variable descriptiva. Si desea obtener más información, consulte “Nombres de variable” en Capítulo 5 en p. 86.

Figura 8-7
Cuadro de diálogo *Nombre y etiqueta de variable*

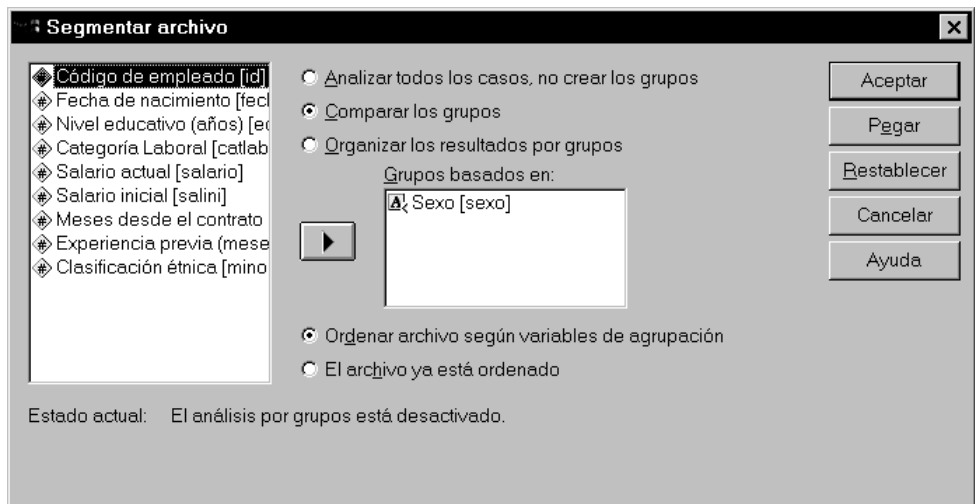


Segmentar archivo

Segmentar archivo divide el archivo de datos en distintos grupos para el análisis basándose en los valores de una o más variables de agrupación. Si selecciona varias variables de agrupación, los casos se agruparán por variable dentro de las categorías de la variable anterior de la lista Grupos basados en. Por ejemplo, si selecciona *sexo* como la primera variable de agrupación y *minoría* como la segunda, los casos se agruparán por minorías dentro de cada categoría de sexo.

- Es posible especificar hasta ocho variables de agrupación.
- Cada ocho caracteres de una variable de cadena larga (variables de cadena que superan los ocho caracteres) cuenta como una variable hasta llegar al límite de ocho variables de agrupación.
- Los casos deben ordenarse según los valores de las variables de agrupación, en el mismo orden en el que aparecen las variables en la lista Grupos basados en. Si el archivo de datos todavía no está ordenado, seleccione Ordenar archivo según variables de agrupación.

Figura 8-8
Cuadro de diálogo Segmentar archivo



Comparar los grupos. Los grupos de segmentación del archivo se presentan juntos para poder compararlos. Para las tablas pivote se crea una sola tabla y cada variable de segmentación del archivo puede desplazarse entre las dimensiones de la tabla. En

el caso de los gráficos se crea un gráfico diferente para cada grupo y se muestran juntos en el Visor.

Organizar los resultados por grupos. Los resultados de cada procedimiento se muestran por separado para cada grupo de segmentación del archivo.

Para segmentar un archivo de datos para el análisis

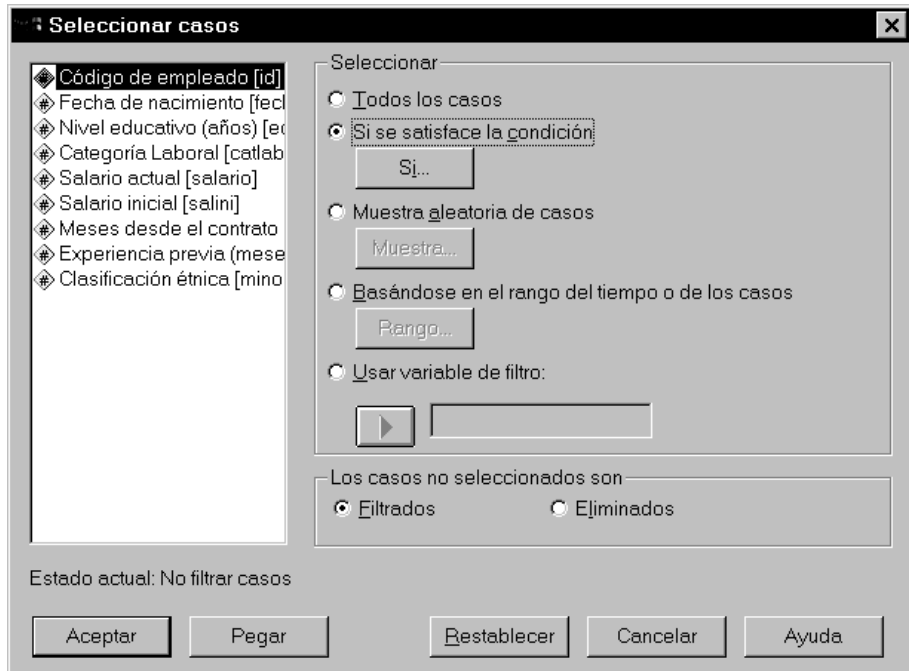
- ▶ Elija en los menús:
 - Datos
 - Segmentar archivo...
- ▶ Seleccione Comparar los grupos o Organizar los resultados por grupos.
- ▶ Seleccione una o más variables de agrupación.

Seleccionar casos

Seleccionar casos proporciona varios métodos para seleccionar un subgrupo de casos basándose en criterios que incluyen variables y expresiones complejas. También se puede seleccionar una muestra aleatoria de casos. Los criterios usados para definir un subgrupo pueden incluir:

- Valores y rangos de las variables
- Rangos de fechas y horas
- Números de caso (filas)
- Expresiones aritméticas
- Expresiones lógicas
- Funciones

Figura 8-9
Cuadro de diálogo Seleccionar casos



Todos los casos. Desactiva el filtrado y utiliza todos los casos.

Si se satisface la condición. Utiliza una expresión condicional para seleccionar los casos. Si el resultado de la expresión condicional es verdadero, el caso se selecciona. Si el resultado es falso o perdido, entonces el caso no se selecciona.

Muestra aleatoria de casos. Selecciona una muestra aleatoria basándose en un porcentaje aproximado o en un número exacto de casos.

Basándose en el rango del tiempo o de los casos. Selecciona los casos basándose en un rango de los números de caso o en un rango de las fechas/horas.

Usar variable de filtro. Utiliza como variable para el filtrado la variable numérica seleccionada del archivo de datos. Se seleccionan los casos con cualquier valor distinto del 0 o del valor perdido para la variable seleccionada.

Casos no seleccionados. Puede filtrar o eliminar los casos que no reúnan los criterios de selección. Los casos filtrados permanecen en el archivo de datos pero se excluyen del análisis. Seleccionar casos crea una variable de filtro, *filter_\$*, para indicar el estado del filtro. Los casos seleccionados tienen un valor de 1; los casos filtrados tienen un valor de 0. Estos últimos también se indican con una barra transversal sobre el número de fila en el Editor de datos. Para desactivar el filtrado e incluir todos los casos en el análisis, seleccione Todos los casos.

Si guarda el archivo de datos después de eliminar casos, no podrá recuperar los casos eliminados.

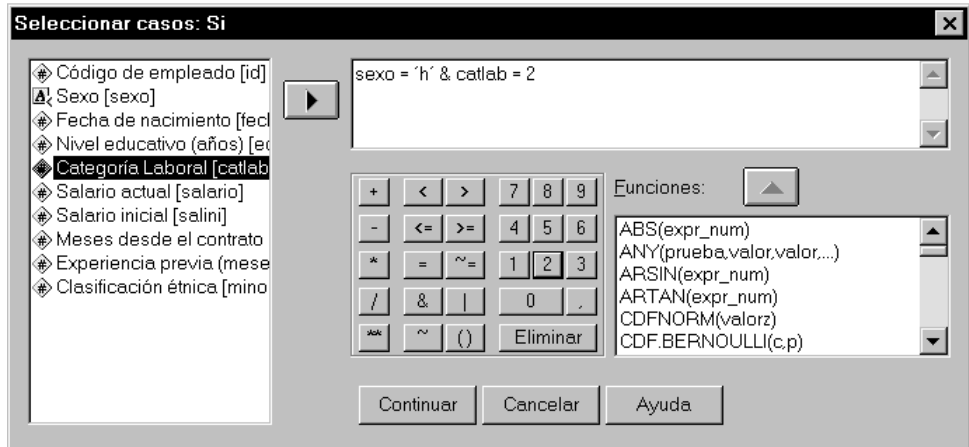
Para seleccionar un subconjunto de casos

- ▶ Elija en los menús:
 - Datos
 - Seleccionar casos...
- ▶ Seleccione uno de los métodos de selección de casos.
- ▶ Especifique los criterios para la selección de casos.

Seleccionar casos: Si la opción

Este cuadro de diálogo permite seleccionar subconjuntos de casos utilizando expresiones condicionales. Una expresión condicional devuelve un valor *verdadero*, *falso* o *perdido* para cada caso.

Figura 8-10
Cuadro de diálogo *Seleccionar casos: Si*

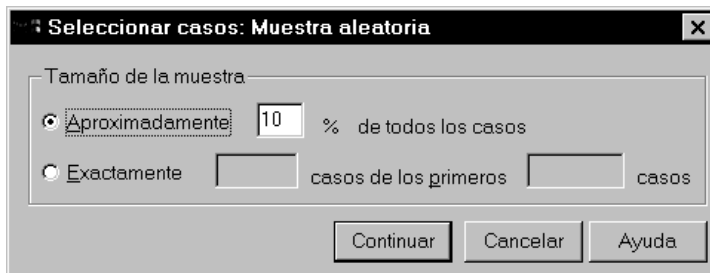


- Si el resultado de una expresión condicional es *verdadero*, se incluirá el caso en el subconjunto seleccionado.
- Si el resultado de una expresión condicional es *falso* o *perdido*, no se incluirá el caso en el subconjunto seleccionado.
- La mayoría de las expresiones condicionales utilizan al menos uno de los seis operadores de relación (<, >, <=, >=, =, and ~=) de la calculadora.
- Las expresiones condicionales pueden incluir nombres de variable, constantes, operadores aritméticos, funciones numéricas y de otros tipos, variables lógicas y operadores de relación.

Seleccionar casos: Muestra aleatoria

Este cuadro de diálogo permite seleccionar una muestra aleatoria basada en un porcentaje aproximado o en un número exacto de casos. El muestreo se realiza sin sustitución, de manera que el mismo caso no se puede seleccionar más de una vez.

Figura 8-11
Cuadro de diálogo Seleccionar casos: Muestra aleatoria



Aproximadamente. Genera una muestra aleatoria con el porcentaje aproximado de casos indicado. Dado que esta rutina toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar al especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.

Exactamente. Un número de casos especificado por el usuario. También se debe especificar el número de casos a partir de los cuales se generará la muestra. Este segundo número debe ser menor o igual que el número total de casos presentes en el archivo de datos. Si lo excede, la muestra contendrá un número menor de casos proporcional al número solicitado.

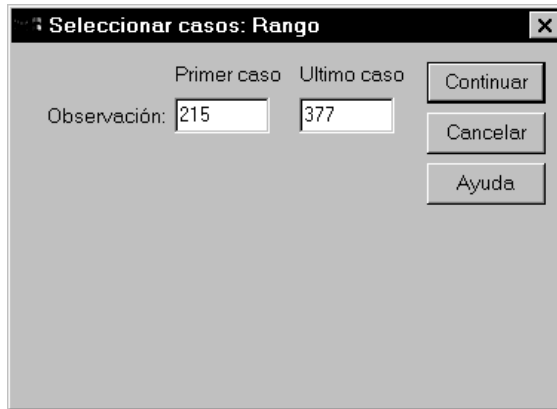
Seleccionar casos: Amplitud

Este cuadro de diálogo selecciona los casos basándose en un rango de números de caso o en un rango de fechas u horas.

- Los rangos de casos se basan en el número de fila que se muestra en el Editor de datos.
- Los rangos de fechas y horas sólo están disponibles para los **datos de serie temporal** con variables de fecha definidas (menú Datos, Definir fechas).

Figura 8-12

Cuadro de diálogo *Seleccionar casos: Rango* para rangos de casos (sin variables de fecha definidas)



The dialog box titled "Seleccionar casos: Rango" has a close button (X) in the top right corner. It contains two input fields: "Primer caso" with the value "215" and "Ultimo caso" with the value "377". Below these fields is the label "Observación:". To the right of the input fields are three buttons: "Continuar", "Cancelar", and "Ayuda".

Figura 8-13

Cuadro de diálogo *Seleccionar casos: Rango* para datos de la serie temporal con variables de fecha definidas



The dialog box titled "Seleccionar casos: Rango" has a close button (X) in the top right corner. It contains two rows of input fields. The first row has "Primer caso" with the value "1990" and "Ultimo caso" with the value "1995", with the label "Año:" to the left. The second row has "Primer caso" with the value "1" and "Ultimo caso" with the value "12", with the label "Mes:" to the left. To the right of the input fields are three buttons: "Continuar", "Cancelar", and "Ayuda".

Ponderar casos

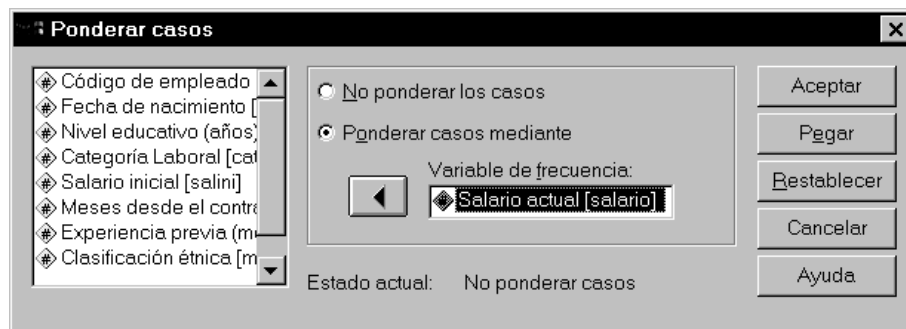
Ponderar casos proporciona a los casos diferentes ponderaciones (mediante una réplica simulada) para el análisis estadístico.

- Los valores de la variable de ponderación deben indicar el número de observaciones representadas por casos únicos en el archivo de datos.

- Los casos con valores perdidos, negativos o cero para la variable de ponderación se excluyen del análisis.
- Los valores fraccionarios son válidos; se usan exactamente donde adquieren sentido y, con mayor probabilidad, donde se tabulan los casos.

Figura 8-14

Cuadro de diálogo Ponderar casos



Si aplica una variable de ponderación, ésta seguirá vigente hasta que se seleccione otra o se desactive la ponderación. Si guarda un archivo de datos ponderado, la información de ponderación se guardará con el archivo. Puede desactivar la ponderación en cualquier momento, incluso después de haber guardado el archivo de forma ponderada.

Ponderaciones en las tablas de contingencia. El procedimiento Tablas de contingencia cuenta con diversas opciones para el tratamiento de ponderaciones de los casos. Si desea obtener más información, consulte “Tablas de contingencia: Mostrar en las casillas” en Capítulo 16 en p. 358.

Ponderaciones en los diagramas de dispersión y los histogramas. Los diagramas de dispersión y los histogramas tienen una opción para activar y desactivar las ponderaciones de los casos, pero dicha opción no afecta a los casos que tienen un valor negativo, un valor 0 o un valor perdido para la variable de ponderación. Estos casos permanecen excluidos del gráfico incluso si se desactiva la ponderación desde el gráfico.

Para ponderar casos

- ▶ Elija en los menús:
 - Datos
 - Ponderar casos...
- ▶ Seleccione Ponderar casos mediante.
- ▶ Seleccione una variable de frecuencia.

Los valores de la variable de frecuencia se utilizan como ponderaciones de los casos. Por ejemplo, un caso con un valor 3 para la variable de frecuencia representará tres casos en el archivo de datos ponderado.

Reestructuración de los datos

Utilice el Asistente de reestructuración de datos para reestructurar los datos de acuerdo con el procedimiento de SPSS que desee utilizar. El asistente sustituye el archivo actual con un archivo nuevo reestructurado. El asistente puede:

- Reestructurar variables seleccionadas en casos.
- Reestructurar casos seleccionados en variables.
- Transponer todos los datos.

Para reestructurar datos

- ▶ Elija en los menús:
 - Datos
 - Reestructurar...
- ▶ Seleccione el tipo de reestructuración que desea realizar.
- ▶ Seleccione los datos que se van a reestructurar.

Si lo desea, tiene la posibilidad de:

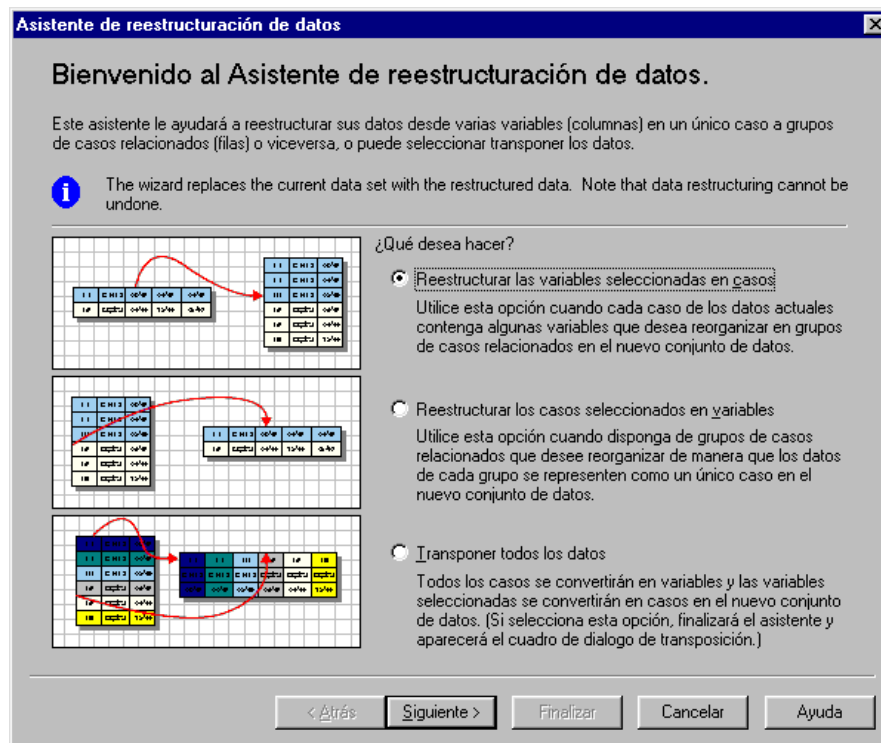
- Crear variables de identificación, que permitirán un seguimiento de un valor del nuevo archivo a partir de un valor del archivo original.
- Ordenar los datos antes de la reestructuración.

- Definir opciones para el nuevo archivo.
- Pegar la sintaxis de comandos en una ventana de sintaxis.

Asistente de reestructuración de datos: *Seleccionar tipo*

Utilice el Asistente de reestructuración de datos para reestructurar los datos. En el primer cuadro de diálogo, seleccione el tipo de reestructuración que desea llevar a cabo.

Figura 8-15
Asistente de reestructuración de datos



- **Reestructurar variables seleccionadas en casos.** Seleccione esta opción cuando disponga, en los datos, de grupos de columnas relacionadas y desee que aparezcan en el nuevo archivo de datos como grupos de filas. Si elige esta opción, el asistente mostrará los pasos para **Variables a casos**.

- **Reestructurar casos seleccionados en variables.** Seleccione esta opción cuando disponga, en los datos, de grupos de filas relacionadas y desee que aparezcan en el nuevo archivo de datos como grupos de columnas. Si elige esta opción, el asistente mostrará los pasos para **Casos a variables**.
- **Transponer todos los datos.** Seleccione esta opción cuando desee transponer los datos. Todas las filas se convertirán en columnas y todas las columnas en filas, en el nuevo archivo de datos. Esta opción cierra el Asistente de reestructuración de datos y abre el cuadro de diálogo Transponer datos.

Opciones de reestructuración de los datos

Una **variable** contiene información que se desea analizar, por ejemplo, una medida o una puntuación. Un **caso** es una observación, por ejemplo, un individuo. En una estructura de datos *simple*, cada variable es una única columna de datos y cada caso es una única fila. De manera que, por ejemplo, si estuviera midiendo las puntuaciones de un examen realizado a todos los alumnos de una clase, todos los valores de las notas aparecerían en una única columna y habría una fila para cada alumno.

Cuando se analizan datos, a menudo se está analizando cómo varía una variable en función de cierta condición. Dicha condición puede ser un tratamiento experimental específico, un grupo demográfico, un momento en el tiempo u otra cosa. En el análisis de datos, a las condiciones de interés a menudo se las denomina **factores**. Al analizar factores, se dispone de una estructura de datos *compleja*. Es posible que haya información acerca de una variable en más de una columna de datos (por ejemplo, una columna para cada nivel de un factor), o que haya información acerca de un caso en más de una fila (por ejemplo, una fila para cada nivel de un factor). El Asistente de reestructuración de datos le ayuda a reestructurar archivos con una estructura de datos compleja.

La estructura del archivo actual y la estructura que se desea en el nuevo archivo determinan las elecciones que se deben seleccionar en el asistente.

¿Cómo están organizados los datos en el archivo actual? Es posible que los datos actuales estén organizados de manera que los factores estén registrados en una variable *diferente* (como grupos de casos) o *con* la variable (como grupos de variables).

- **Grupos de casos.** ¿El archivo actual tiene registradas las variables y las condiciones en columnas diferentes? Por ejemplo:

var	factor
8	1
9	1
3	2
1	2

En este ejemplo, las dos primeras filas son un **grupo de casos** porque están relacionadas. Contienen datos para el mismo nivel del factor. En el análisis de datos de SPSS, cuando los datos están estructurados de esta manera, se hace referencia al factor como **variable de agrupación**.

- **Grupos de columnas.** ¿El archivo actual tiene registradas las variables y las condiciones en la misma columna? Por ejemplo:

var_1	var_2
8	3
9	1

En este ejemplo, las dos primeras columnas son un **grupo de variables** porque están relacionadas. Contienen datos para la misma variable, *var_1* para el nivel 1 del factor y *var_2* para el nivel 2 del factor. En el análisis de datos de SPSS, si los datos se estructuran de esta manera, el factor se suele denominar **de medidas repetidas**.

¿Cómo deben organizarse los datos en el archivo nuevo? Normalmente, la organización estará determinada por el procedimiento que se vaya a utilizar para analizar los datos.

- **Procedimientos que requieren grupos de casos.** Los datos deberán estructurarse en grupos de casos para realizar los análisis que requieran una variable de agrupación. Algunos ejemplos son: *univariante*, *multivariante* y *componentes de la varianza* de los Modelos lineales generales; Modelos mixtos; Cubos OLAP; y *muestras independientes* de las Pruebas T o Pruebas no paramétricas. Si la

estructura de datos actual es de grupos de variables y desea realizar estos análisis, seleccione Reestructurar variables seleccionadas en casos.

- **Procedimientos que requieren grupos de variables.** Los datos se deberán estructurar en grupos de variables para analizar medidas repetidas. Algunos ejemplos son: *medidas repetidas* de los Modelos lineales generales, análisis de *covariables dependientes del tiempo* del Análisis de regresión de Cox, *muestras relacionadas* de las Pruebas T o *muestras relacionadas* de las Pruebas no paramétricas. Si la estructura de datos actual es de grupos de casos y desea realizar estos análisis, seleccione Reestructurar casos seleccionados en variables.

Ejemplo de variables a casos

En este ejemplo, las puntuaciones de las pruebas están registradas en columnas diferentes para cada factor, A y B.

Figura 8-16

Datos actuales para reestructurar variables a casos

	punt_a	punt_b
1	1014.00	864.00
2	684.00	636.00
3	810.00	638.00

Se desea realizar una prueba *t* para muestras independientes. Se dispone de un grupo de columnas compuesto por *puntuación_a* y *puntuación_b*, pero no se dispone de la **variable de agrupación** que requiere el procedimiento. Seleccione Reestructurar variables seleccionadas en casos en el Asistente de reestructuración de datos, reestructure un grupo de variables en una nueva variable denominada *puntuación* y cree un índice denominado *grupo*. El nuevo archivo de datos se muestra en la siguiente imagen.

Figura 8-17

Datos nuevos y reestructurados para variables a casos

	grupo	punt
1	PUNT_A	1014.00
2	PUNT_B	864.00
3	PUNT_A	684.00
4	PUNT_B	636.00
5	PUNT_A	810.00
6	PUNT_B	638.00

Cuando se ejecute la prueba t para muestras independientes, podrá utilizar *grupo* como variable de agrupación.

Ejemplo de casos a variables

En este ejemplo, las puntuaciones de las pruebas están registradas dos veces para cada sujeto, antes y después de un tratamiento.

Figura 8-18

Datos actuales para reestructurar casos a variables

	id	punt	tiempo
1	1	1014.00	bef
2	1	864.00	aft
3	2	684.00	bef
4	2	636.00	aft

Se desea realizar una prueba t para muestras relacionadas. La estructura de datos es de grupos de casos, pero no se dispone de las **medidas repetidas** para las variables relacionadas que requiere el procedimiento. Seleccione Reestructurar casos seleccionados en variables en el Asistente de reestructuración de datos, utilice *id* para identificar los grupos de filas en los datos actuales y utilice *tiempo* para crear el grupo de variables en el nuevo archivo.

Figura 8-19

Datos nuevos y reestructurados para casos a variables

	id	ant	des
1	1	864.00	1014.00
2	2	636.00	684.00

Cuando se ejecute la prueba t de muestras relacionadas, podrá utilizar *ant* y *des* como el par de variables.

Asistente de reestructuración de datos (variables a casos): Número de grupos de variables

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

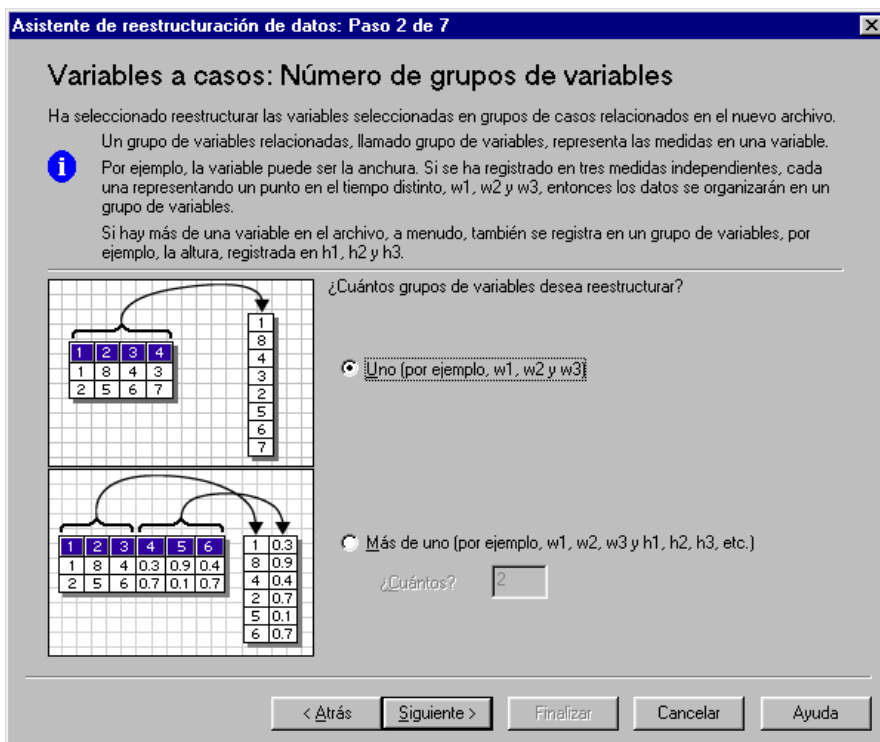
En este paso, se debe elegir el número de grupos de variables del archivo actual que se desea reestructurar en el nuevo archivo.

¿Cuántos grupos de variables hay en el archivo actual? Piense cuántos grupos de variables existen en los datos actuales. Un grupo de columnas relacionadas, llamado **grupo de variables**, registra medidas repetidas de la misma variable en distintas columnas. Por ejemplo, si en los datos actuales hay tres columnas, *c1*, *c2* y *c3*, que registran el **contorno**, entonces hay un grupo de variables. Si además hay otras tres columnas, *a1*, *a2* y *a3*, que registran la **altura**, entonces hay dos grupos de variables.

¿Cuántos grupos de variables debe haber en el archivo nuevo? Considere cuántos grupos de variables desea que estén representados en el nuevo archivo de datos, teniendo en cuenta que no es necesario reestructurar todos los grupos de variables en el nuevo archivo.

Figura 8-20

Asistente de reestructuración de datos: Número de grupos de variables



- **Uno.** El asistente creará una única variable reestructurada en el nuevo archivo a partir de un grupo de variables del archivo actual.
- **Más de uno.** El asistente creará varias variables reestructuradas en el nuevo archivo. El número que se especifique afectará al siguiente paso, en el que el asistente creará de forma automática el número especificado de nuevas variables.

Asistente de reestructuración de datos (variables a casos): Seleccione Variables

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se debe proporcionar información sobre cómo se van a utilizar las variables del archivo actual en el nuevo archivo. También se puede crear una variable que identifique las filas en el nuevo archivo.

Figura 8-21
Asistente de reestructuración de datos: Seleccione Variables



¿Cómo se deben identificar las nuevas filas? En el nuevo archivo de datos, puede crear una variable que identifique la fila del archivo de datos actual que ha sido utilizada para crear un grupo de filas nuevo. El identificador puede ser un número de caso secuencial o los valores de una variable. Utilice los controles disponibles en el apartado Identificación de grupos de casos para definir la variable de identificación utilizada en el nuevo archivo. Pulse en la casilla para cambiar el nombre de variable por defecto y para dotar a la variable de identificación de una etiqueta de variable descriptiva.

¿Qué se debe reestructurar en el nuevo archivo? En el paso anterior, se informó al asistente del número de grupos de variables que se deseaba reestructurar. El asistente creó una nueva variable para cada grupo. Los valores para el grupo de variables

aparecerán en dicha variable en el nuevo archivo. Utilice los controles en Variables que se van a transponer para definir la variable reestructurada en el nuevo archivo.

Para especificar una variable reestructurada

- ▶ Ponga las variables que componen el grupo de variables que desea transformar en la lista Variables que se van a transponer. Todas las variables del grupo deberán ser del mismo tipo (numéricas o de cadena).

Se puede incluir la misma variable más de una vez en el grupo de variables (las variables se copian de la lista origen de variables en lugar de moverlas); los valores se repetirán en el nuevo archivo.

Para especificar varias variables reestructuradas

- ▶ Seleccione la primera variable de destino que desea definir de la lista desplegable Variable de destino.
- ▶ Ponga las variables que componen el grupo de variables que desea transformar en la lista Variables que se van a transponer. Todas las variables del grupo deberán ser del mismo tipo (numéricas o de cadena). Puede incluir la misma variable más de una vez en el grupo de variables. (Las variables se copian de la lista origen de variables en lugar de moverlas, y los valores se repetirán en el nuevo archivo.)
- ▶ Seleccione la siguiente variable de destino que desea definir y repita el proceso de selección de variables para todas las variables de destino disponibles.
 - Aunque puede incluir la misma variable más de una vez en el mismo grupo de variables de destino, no puede incluir la misma variable en más de un grupo de variables de destino.
 - Cada lista de grupos de variables de destino debe contener el mismo número de variables. (Las variables que aparecen más de una vez se incluyen en el recuento).

- El número de grupos de variables de destino está determinado por el número de grupos de variables especificados en el paso anterior. Aquí puede cambiar los nombres de las variables por defecto, pero deberá volver al paso anterior para cambiar el número de grupos de variables que se van a reestructurar.
- Debe definir los grupos de variables (seleccionando variables de la lista de origen) para todas las variables de destino disponibles antes de poder pasar al siguiente paso.

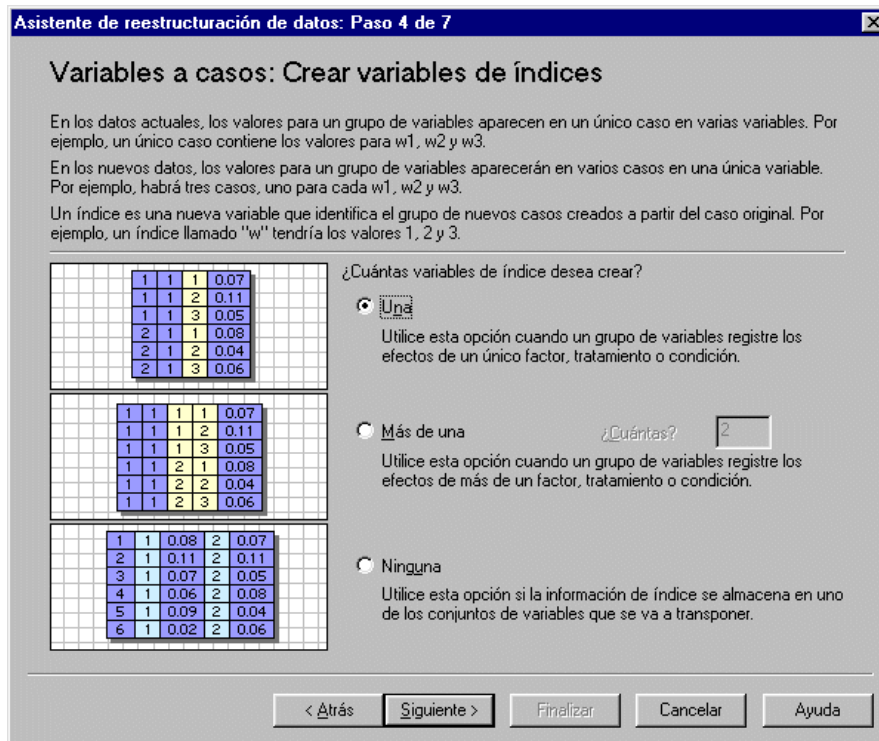
¿Qué se debe copiar en el nuevo archivo? En el nuevo archivo se pueden copiar variables que no se han reestructurado. Sus valores se propagarán en las nuevas filas. Desplace las variables que desea copiar en el nuevo archivo en la lista Variables fijas.

Asistente de reestructuración de datos (variables a casos): Crear variables de Variables

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se debe decidir si se crean variables de índice. Un índice es una nueva variable que identifica de forma secuencial un grupo de filas en función de la variable original a partir de la cual se creó la nueva fila.

Figura 8-22
Asistente de reestructuración de datos: Crear variables de índice



¿Cuántas variables de índice debe haber en el archivo nuevo? Las variables de índice se pueden utilizar como variables de agrupación en los procedimientos de SPSS. En la mayoría de los casos, es suficiente una única variable de índice; no obstante, si los grupos de variables del archivo actual reflejan varios niveles de factor, puede ser conveniente utilizar varios índices.

- **Uno.** El asistente creará una única variable de índice.
- **Más de uno.** El asistente creará varios índices y deberá introducir el número de índices que desea crear. El número especificado afectará al siguiente paso, en el que el asistente crea de forma automática el número especificado de índices.
- **Ninguno.** Seleccione esta opción si no desea crear variables de índice en el nuevo archivo.

Ejemplo de un índice para variables a casos

En los datos actuales, hay un grupo de variables, denominado *contorno*, y un factor, el *tiempo*. El contorno se ha medido en tres ocasiones y se ha registrado en *c1*, *c2* y *c3*.

Figura 8-23
Datos actuales para un índice

	sujeo	w1	w2	w3
1	1	6.70	4.30	5.70
2	2	7.10	5.90	5.60

Se va a reestructurar el grupo de variables en una única variable, *contorno*, y se va a crear un único índice numérico. Los nuevos datos se muestran en la siguiente tabla.

Figura 8-24
Datos nuevos y reestructurados con un índice

	sujeo	índice	anchura
1	1	1	6.70
2	1	2	4.30
3	1	3	5.70
4	2	1	7.10
5	2	2	5.90
6	2	3	5.60

El *Índice* comienza por 1 y se incrementa por cada variable del grupo. Vuelve a comenzar cada vez que se encuentra una fila en el archivo original. Ahora se puede utilizar *índice* en procedimientos de SPSS que requieran una variable de agrupación.

Ejemplo de dos índices para variables a casos

Cuando un grupo de variables registra más de un factor, se puede crear más de un índice; no obstante, se deben organizar los datos actuales de forma que los niveles del primer factor sean un índice primario dentro del cual varían los niveles de los siguientes factores. En los datos actuales, hay un grupo de variables, denominado *contorno*, y dos factores, *A* y *B*. Los datos se organizan de manera que los niveles del factor *B* varían dentro de los niveles del factor *A*.

Figura 8-25

Datos actuales para dos índices

	sujeto	w_a1b1	w_a1b2	w_a2b1	w_a2b2
1	1	5.50	6.40	5.80	5.90
2	2	7.40	7.10	5.60	6.70

Se va a reestructurar el grupo de variables en una única variable, *contorno*, y se van a crear dos índices. Los nuevos datos se muestran en la siguiente tabla.

Figura 8-26

Datos nuevos reestructurados con los dos índices

	sujeto	índice_a	índice_b	anchura
1	1	1	1	5.50
2	1	1	2	6.40
3	1	2	1	5.80
4	1	2	2	5.90
5	2	1	1	7.40
6	2	1	2	7.10
7	2	2	1	5.60
8	2	2	2	6.70

Asistente de reestructuración de datos (variables a casos): Crear una variable de índices

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas y crear una variable de índices.

En este paso, se debe decidir los valores que se desean para la variable de índice. Los valores pueden ser números secuenciales o los nombres de las variables en un grupo de variables original. También puede especificar un nombre y una etiqueta para la nueva variable de índice.

Figura 8-27
Asistente de reestructuración de datos: Crear una variable de índice

Asistente de reestructuración de datos: Paso 5 de 7

Variables a casos: Crear una variable de índice

Ha seleccionado crear una variable de índice. Los valores de la variable pueden ser números secuenciales o los nombres de variables en un grupo.
En la tabla, puede especificar el nombre y etiqueta de la variable de índice.

¿Qué tipo de valores de índice?

Números secuenciales
Valores de índice: 1, 2, 3

Nombres de variables
Valores de índice:

Editar el nombre y la etiqueta de la variable de índice:

	Nombre	Etiqueta	Niveles	Valores de índice
1	índice1		3	1, 2, 3

< Atrás **Siguiente** > Finalizar Cancelar Ayuda

Si desea obtener más información, consulte “Ejemplo de un índice para variables a casos” en p. 225.

- **Números secuenciales.** El asistente asignará de forma automática números secuenciales como valores índice.
- **Nombres de variable.** El asistente utilizará los nombres del grupo de variables seleccionado como valores índice. Seleccione un grupo de variables de la lista.
- **Nombres y etiquetas.** Pulse en una casilla para cambiar el nombre de variable por defecto y proporcionar una etiqueta de variable descriptiva para la variable de índice.

Asistente de reestructuración de datos (variables a casos): Crear varias variables de índices

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas y crear varias variables de índice.

En este paso, se debe especificar el número de niveles para cada variable de índice. También puede especificar un nombre y una etiqueta para la nueva variable de índice.

Figura 8-28

Asistente de reestructuración de datos: Crear Varias variables de índice

Asistente de reestructuración de datos: Paso 5 de 7

Variables a casos: Crear varias variables de índice

Ha seleccionado disponer de varias variables de índice en el nuevo archivo de datos. La primera variable de la tabla es el índice primario. Todos los índices siguientes recorren cada nivel de este índice.

En la siguiente tabla, introduzca el número de niveles para cada variable de índice. También puede editar el nombre y la etiqueta de cada variable de índice.

Nombres, etiquetas y número de niveles de las variables de índice:

	Nombre	Etiqueta	Niveles	Valores de índice
1	índice1			
2	índice2			

Número total de niveles combinados [product:4]

i Recuerde que el producto del número de niveles de todas las variables de índice debe ser igual al número 'Total' que aparece bajo la tabla. Este número equivale al número de variables que se van a transponer.

< Atrás Siguiete > Finalizar Cancelar Ayuda

Si desea obtener más información, consulte “Ejemplo de dos índices para variables a casos” en p. 225.

¿Cuántos niveles hay registrados en el archivo actual? Calcule los niveles de factor que hay registrados en los datos actuales. Un **nivel** define un grupo de casos que experimentan las mismas condiciones. Si hay varios factores, los datos actuales se deben organizar de manera que los niveles del primer factor sean un índice primario dentro del cual varían los niveles de los siguientes factores.

¿Cuántos niveles debe haber en el archivo nuevo? Introduzca el número de niveles para cada índice. Los valores para varias variables de índice son siempre números secuenciales. Los valores comienzan en 1 y se incrementan con cada nivel. El primer índice se incrementa más despacio y el último más deprisa.

Número total de niveles combinados. No se puede crear más niveles de los que existen en los datos actuales. Como los datos reestructurados contendrán una fila por cada combinación de tratamientos, el asistente realizará una comprobación del número de niveles que se crean. Comparará el producto de los niveles creados con el número de variables del grupo de variables. Deben coincidir.

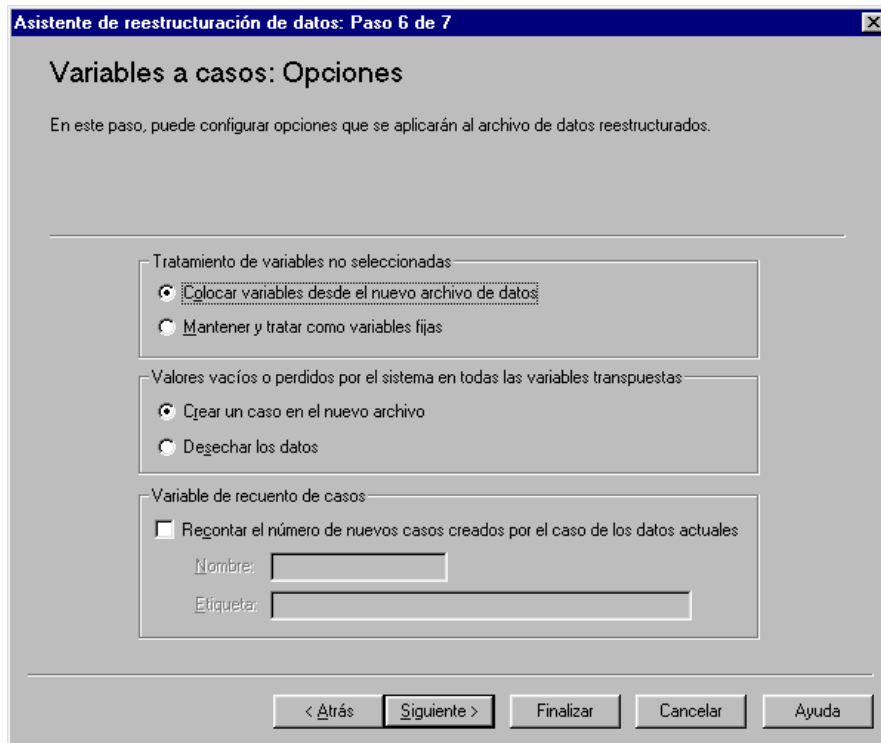
Nombres y etiquetas. Pulse en una casilla para cambiar el nombre de variable por defecto y proporcionar una etiqueta de variable descriptiva para las variables de índice.

Asistente de reestructuración de datos (variables a casos): Opciones

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se deben especificar las opciones para el nuevo archivo reestructurado.

Figura 8-29
Asistente de reestructuración de datos: Opciones



¿Desea eliminar las variables no seleccionadas? En el paso de selección de variables (paso 3), se seleccionaron los grupos de variables que se iban a reestructurar, las variables que se iban a copiar y una variable de identificación de los datos actuales. Los datos de las variables seleccionadas aparecerán en el nuevo archivo. Si hay más variables en los datos actuales, puede elegir descartarlas o conservarlas.

¿Desea conservar los datos perdidos? El asistente comprueba cada nueva fila potencial en busca de valores nulos. Un **valor nulo** es un valor en blanco o perdido por el sistema. Se puede elegir entre conservar o descartar las filas que contienen sólo valores nulos.

¿Desea crear una variable de recuento? El asistente puede crear una **variable de recuento** en el nuevo archivo. Dicha variable contiene el número de nuevas filas generadas por una fila de los datos actuales. Una variable de recuento puede ser de

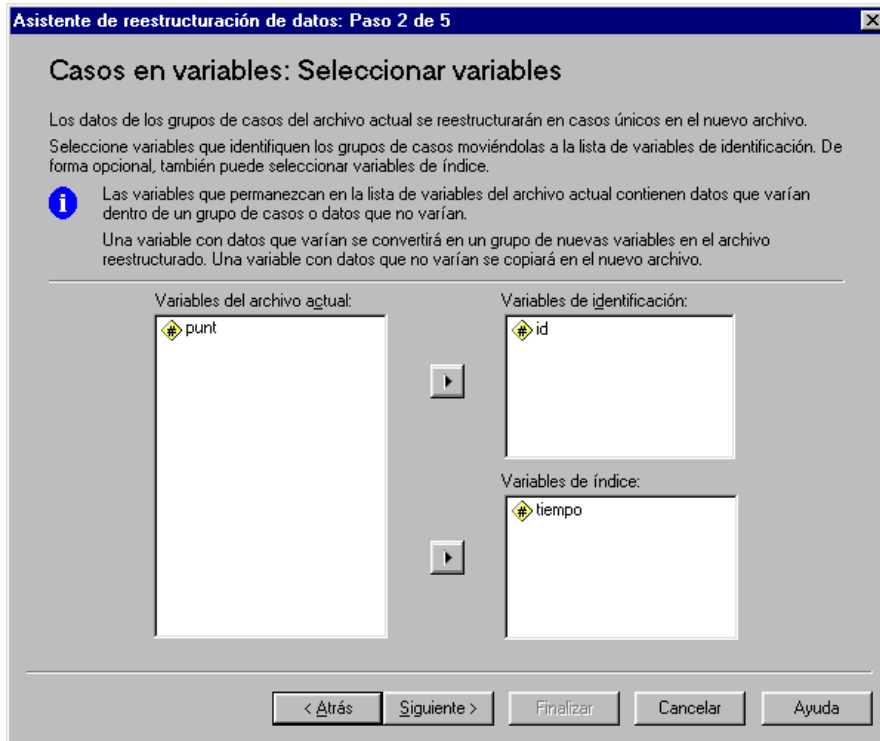
gran utilidad si decide descartar del nuevo archivo los valores nulos, ya que esto conlleva la generación de un número distinto de nuevas filas por una fila dada de los datos actuales. Pulse en una casilla para cambiar el nombre de variable por defecto y proporcionar una etiqueta de variable descriptiva para la variable de recuento.

Asistente de reestructuración de datos (casos a variables): Seleccione Variables

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de casos en columnas.

En este paso, se debe proporcionar información sobre cómo se van a utilizar las variables del archivo actual en el nuevo archivo.

Figura 8-30
Asistente de reestructuración de datos: Seleccione Variables



¿Qué identifica los grupos de casos en los datos actuales? Un grupo de casos es un grupo de filas relacionadas porque miden la misma unidad de observación, por ejemplo, un individuo o una institución. El asistente necesita conocer cuáles son las variables del archivo actual que identifican los grupos de casos para que se pueda consolidar cada grupo en una única fila del nuevo archivo. Desplace las variables que identifican grupos de casos en el archivo actual a la lista de Variables de identificación. Las variables que se utilizan para segmentar el archivo de datos actual se utilizan de forma automática para identificar los grupos de casos. Cada vez que se encuentra una nueva combinación de valores de identificación, el asistente creará una nueva fila, de manera que los casos del archivo actual deberán ordenarse en función de los valores de las variables de identificación, en el mismo orden en el que aparecen las variables en la lista Variables de identificación. Si el archivo de datos actual no está aún ordenado, podrá hacerlo en el siguiente paso.

¿Cómo deben crearse los nuevos grupos de variables en el archivo nuevo? En los datos originales, una variable aparece en una única columna. En el nuevo archivo de datos, dicha variable aparecerá en varias columnas. Las **variables de índice** son variables existentes en los datos actuales que el asistente deberá utilizar para crear las nuevas columnas. Los datos reestructurados contendrán una nueva variable por cada valor único contenido en dichas columnas. Desplace a la lista Variables de índice las variables que se deben utilizar para formar los nuevos grupos de variables. Cuando el asistente ofrezca opciones, también puede elegir ordenar las nuevas columnas por el índice.

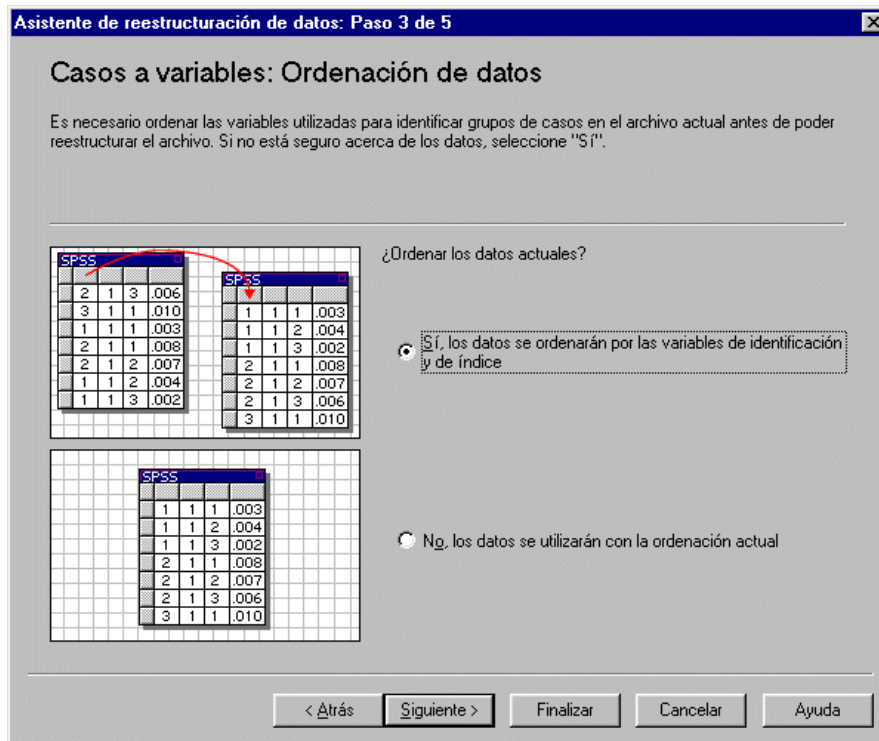
¿Qué sucede con las demás columnas? El asistente decide de forma automática lo que hay que hacer con las variables que quedan en la lista Archivo actual. Comprueba cada variable para ver si los valores de los datos varían dentro de un grupo de casos. Si hay alguna variación, el asistente reestructurará los valores en un grupo de variables en el nuevo archivo. Si no la hay, el asistente copiará los valores en el nuevo archivo.

Asistente de reestructuración de datos (casos a variables): Ordenar Datos

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de casos en columnas.

En este paso, debe decidir si se ordena el archivo actual antes de reestructurarlo. Cada vez que el asistente se encuentra una nueva combinación de valores de identificación, se crea una nueva fila, por lo tanto, es importante que los datos estén ordenados por las variables que identifican los grupos de casos.

Figura 8-31
Asistente de reestructuración de datos: Ordenar Datos



¿Cómo están ordenadas las filas en el archivo actual? Tenga en cuenta la ordenación de los datos actuales y cuáles son las variables que se están utilizando para identificar grupos de casos (especificadas en el paso anterior).

- **Sí.** El asistente ordenará de forma automática los datos actuales en función de la variable de identificación, con el mismo orden en el que aparecen las variables en la lista Variables de identificación en el paso anterior. Seleccione esta opción cuando los datos no estén ordenados en función de las variables de identificación o cuando no esté seguro. Esta opción requiere una lectura adicional de los

datos, pero garantiza que las filas estén correctamente ordenadas antes de la reestructuración.

- **No.** El asistente no ordenará los datos actuales. Seleccione esta opción cuando esté seguro de que los datos actuales están ordenados en función de las variables que identifican los grupos de casos.

Asistente de reestructuración de datos (casos a variables): Opciones

Nota: El asistente presenta este paso si se ha seleccionado reestructurar grupos de casos en columnas.

En este paso, se deben especificar las opciones para el nuevo archivo reestructurado.

Figura 8-32

Asistente de reestructuración de datos: Opciones

Asistente de reestructuración de datos: Paso 4 de 5

Casos a variables: Opciones

En este paso, puede configurar opciones que se aplicarán al archivo de datos reestructurados.

Orden de los nuevos grupos de variables:

- Agrupar por variable original (por ejemplo: w1 w2 w3, h1 h2 h3)
- Agrupar por índice (por ejemplo: w1 h1, w2 h2, w3 h3)

Variable de recuento de casos:

- Preguntar el número de casos en los datos actuales utilizados para crear un nuevo caso

Nombre:

Etiqueta:

< Atrás Siguiente > Finalizar Cancelar Ayuda

¿Cómo deben ordenarse los nuevos grupos de variables en el archivo nuevo?

- **Por variable.** El asistente agrupa juntas las nuevas variables creadas a partir de una variable original.
- **Por índice.** El asistente agrupa las variables en función de los valores de las variables de índice.

Ejemplo. Las variables que se van a reestructurar son *w* y *h*, y el índice es *mes*:

```
w      h      mes
```

La agrupación por variable dará como resultado:

```
w.ene  w.feb  h.ene
```

La agrupación por índice dará como resultado:

```
w.ene  h.ene  w.feb
```

¿Desea crear una variable de recuento? El asistente puede crear una variable de recuento en el nuevo archivo. Dicha variable contendrá el número de filas de los datos actuales que se utilizaron para crear una fila en el nuevo archivo de datos.

¿Desea crear variables indicadoras? El asistente puede utilizar las variables de índice para crear **variables indicadoras** en el nuevo archivo de datos. Creará una nueva variable por cada valor único de la variable de índice. Las variables indicadoras indican la presencia o ausencia de un valor para un caso. Una variable indicadora toma el valor 1 si el caso tiene un valor; en caso contrario, vale 0.

Ejemplo. La variable de índice es *producto*. Registra los productos que ha comprado un cliente. Los datos originales son:

cliente	producto
1	pollo
1	huevos
2	huevos
3	pollo

La creación de una variable indicadora da como resultado una nueva variable para cada valor único de *producto*. Los datos reestructurados son:

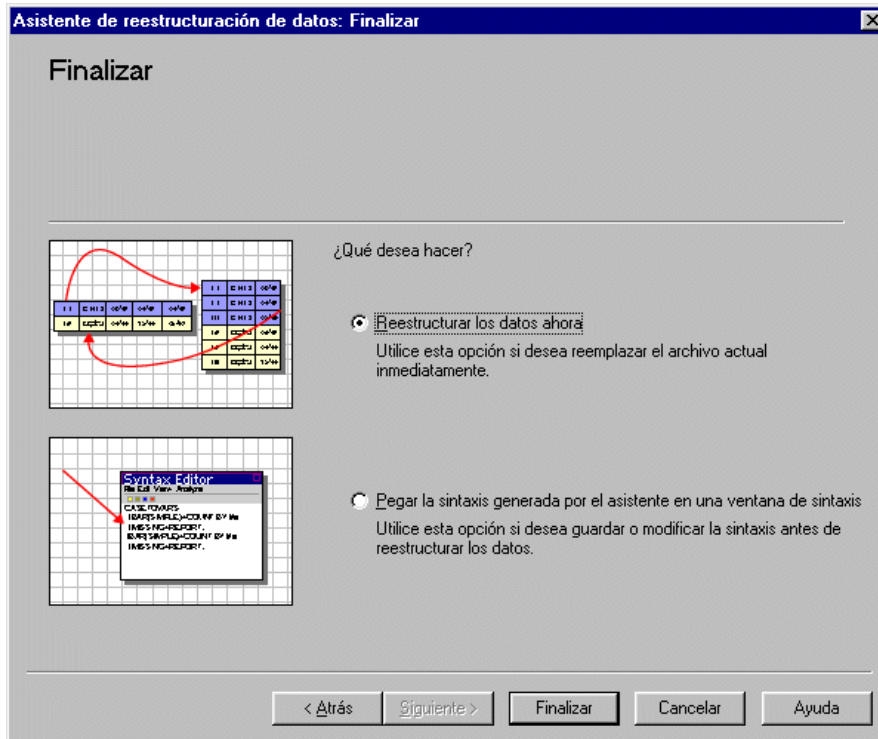
cliente	indpollo	indhuevos
1	1	1
2	0	1
3	1	0

En este ejemplo, se pueden utilizar los datos reestructurados para obtener recuentos de frecuencias de los productos que compran los clientes.

Asistente de reestructuración de datos: Finalizar

Este es el paso final del Asistente de reestructuración de datos. Debe decidir qué hacer con las especificaciones.

Figura 8-33
Asistente de reestructuración de datos: Finalizar



- **Reestructurar los datos ahora.** El asistente creará el nuevo archivo de datos reestructurado. Seleccione esta opción si desea reemplazar el archivo actual inmediatamente. *Nota:* Si los datos originales están ponderados, los nuevos datos también lo estarán, a menos que la variable utilizada como ponderación se reestructure o se elimine del nuevo archivo.
- **Pegar la sintaxis.** El asistente pegará la sintaxis que ha generado en una ventana de sintaxis. Seleccione esta opción si no está preparado para reemplazar el archivo actual, si desea modificar la sintaxis o si desea guardarla para utilizarla en el futuro.

Trabajo con resultados

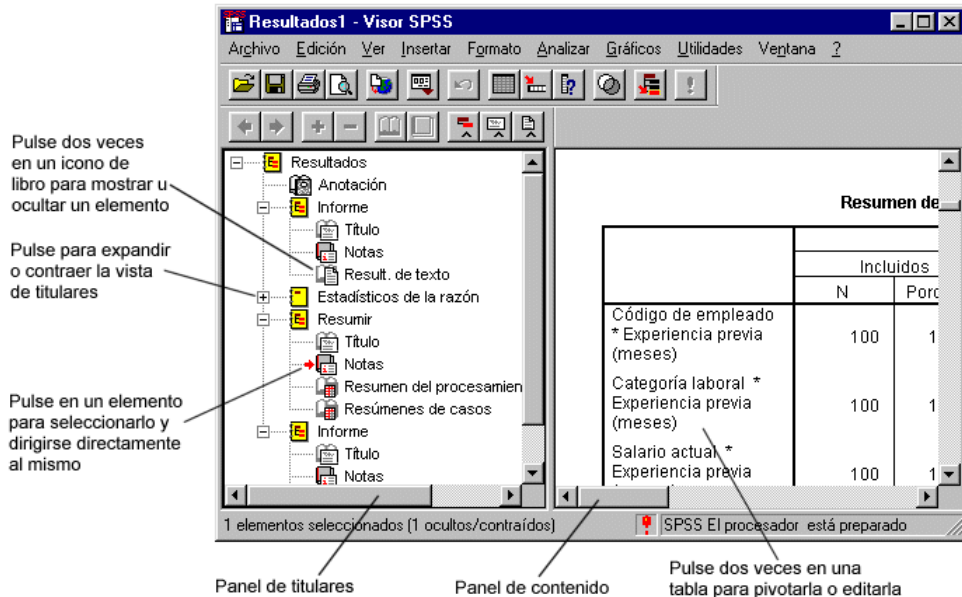
Cuando ejecute un procedimiento, los resultados se mostrarán en una ventana llamada Visor. Desde esta ventana puede desplazarse con facilidad a cualquier parte de los resultados que desee ver. También puede modificar los resultados y crear un documento que contenga exactamente los resultados que desee, organizados y con los formatos adecuados.

Visor

Los resultados se muestran en el Visor. Puede utilizar el Visor para:

- Examinar los resultados.
- Mostrar u ocultar tablas y gráficos seleccionados.
- Cambiar el orden de presentación de los resultados moviendo los elementos seleccionados.
- Mover elementos entre el Visor y otras aplicaciones.

Figura 9-1
Visor



El Visor se divide en dos paneles:

- El panel izquierdo muestra una vista de titulares del contenido.
- El panel derecho contiene tablas estadísticas, gráficos y resultados de texto.

Puede utilizar las barras de desplazamiento para examinar los resultados o bien pulsar en un elemento de los titulares para ir directamente a la tabla o gráfico correspondientes. Puede pulsar y arrastrar el borde derecho del panel de titulares para cambiar la anchura del mismo.

Uso del Visor de borrador

Si prefiere ver los resultados en formato de sólo texto en lugar de como tablas pivote interactivas, utilice el Visor de borrador.

Para utilizar el Visor de borrador:

- ▶ Seleccione en cualquiera de las ventanas las siguientes opciones de los menús:

- Edición
- Opciones...

- ▶ En la pestaña General, pulse en Borrador para determinar el tipo de resultados.
- ▶ Para modificar las opciones de formato de los resultados del Visor de borrador, pulse en la pestaña Visor de borrador.

Si desea obtener más información, consulte “Visor de borrador” en Capítulo 10 en p. 275. También puede ir a la función Ayuda para aprender más:

- ▶ Seleccione en cualquiera de las ventanas las siguientes opciones de los menús:

- ?
- Temas

- ▶ Pulse en la pestaña Índice de la ventana de Ayuda en pantalla.
- ▶ Escriba visor de borrador y pulse dos veces en la entrada del índice.

Mostrar y ocultar resultados

En el Visor, puede mostrar y ocultar de forma selectiva las tablas o los resultados individuales de todo un procedimiento. Resulta de utilidad cuando desea reducir la cantidad de resultados visibles en el panel de contenidos.

Ocultación de gráficos y tablas

- ▶ En el panel de titulares del Visor, pulse dos veces en su icono de libro.

o

- ▶ Pulse en el elemento para seleccionarlo.

- ▶ Elija en los menús:

Ver
Ocult.

o

- ▶ Pulse en el icono de libro cerrado (Ocultar) de la barra de herramientas de titulares.

El icono de libro abierto (Mostrar) se convierte en el icono activo, indicando que el elemento no está oculto.

Para ocultar todos los resultados de un procedimiento

- ▶ Pulse en el cuadro situado a la izquierda del nombre del procedimiento en el panel de titulares.

Se ocultarán todos los resultados del procedimiento y se contraerá la presentación de titulares.

Desplazamiento, eliminación y copia de resultados

Los resultados se pueden reorganizar copiando, moviendo o eliminando un elemento o un grupo de elementos.

Para desplazar resultados en el Visor

- ▶ Pulse en un elemento del panel de titulares o de contenidos para seleccionarlo. Mantenga pulsada la tecla Mayús para seleccionar múltiples elementos o la tecla Ctrl para seleccionar elementos no contiguos)
- ▶ Utilice el ratón para pulsar y arrastrar elementos seleccionados (mantenga pulsado el botón del ratón mientras arrastra).
- ▶ Suelte el botón del ratón en el elemento situado justamente encima de la ubicación donde desea colocar los elementos movidos.

También puede mover elementos con las opciones Cortar y Pegar detrás del menú Edición.

Para eliminar resultados en el Visor

- ▶ Pulse en un elemento del panel de titulares o de contenidos para seleccionarlo. Mantenga pulsada la tecla Mayús para seleccionar múltiples elementos o la tecla Ctrl para seleccionar elementos no contiguos)
- ▶ Pulse Supr.
o
- ▶ Elija en los menús:
Edición
Borrar

Para copiar resultados en el Visor

- ▶ Pulse en los elementos del panel de titulares o de contenidos para seleccionarlos. Mantenga pulsada la tecla Mayús para seleccionar múltiples elementos o la tecla Ctrl para seleccionar elementos no contiguos)
- ▶ Mantenga pulsada la tecla Ctrl mientras utiliza el ratón para pulsar y arrastrar los elementos seleccionados (mantenga pulsado el botón del ratón mientras arrastra).
- ▶ Suelte el botón del ratón para colocar los elementos en la ubicación que desee.
También puede copiar elementos con las opciones Copiar y Pegar detrás del menú Edición.

Cambio de la alineación

Por defecto, todos los resultados están alineados inicialmente a la izquierda. Puede cambiar la alineación inicial (menú Edición, comando Opciones, pestaña Visor) o modificar la alineación de los elementos seleccionados en cualquier momento.

Cambio de la alineación de los resultados

- ▶ Seleccione los elementos que desee alinear (pulse en los elementos del panel de titulares o de contenidos; mantenga pulsada la tecla Mayús o la tecla Ctrl si desea seleccionar múltiples elementos).
- ▶ Elija en los menús:
 - Formato
 - Alineación izquierda

Otras opciones de alineación son Centrado y Alineación derecha.

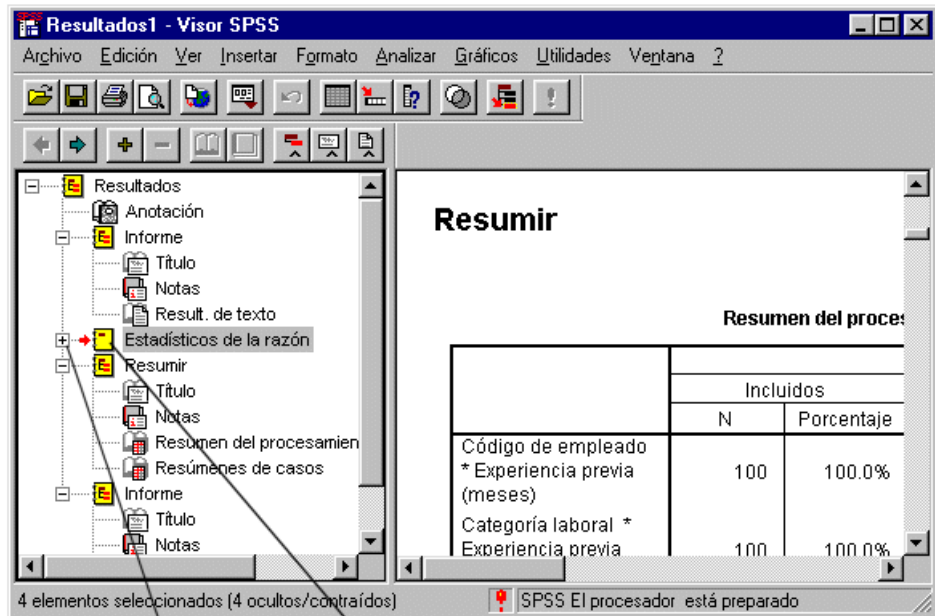
Nota: Todos los resultados aparecen alineados a la izquierda en el Visor. Los ajustes de alineación sólo afectan a los resultados. Los elementos con alineación centrada y a la derecha se identifican por un pequeño símbolo situado encima y a la izquierda del elemento.

Resultados del Visor

El panel de titulares proporciona una tabla de contenidos del documento del Visor. Utilice este panel para navegar por los resultados y controlar su presentación. La mayoría de las acciones en dicho panel tienen su efecto correspondiente en el panel de contenidos.

- Si se selecciona un elemento en el panel de titulares, se seleccionará y mostrará el elemento correspondiente en el panel de contenidos.
- Si se mueve un elemento en el panel de titulares, también se moverá el elemento correspondiente en el panel de contenidos.
- Si se contrae la vista de titulares, se ocultarán los resultados de todos los elementos en los niveles contraídos.

Figura 9-2
Vista de titulares contraída y resultados ocultos



Pulse aquí para expandir o contraer la vista de titulares

Resultado de la vista de titulares contraída y oculta en el panel de contenidos

Control de la presentación de los titulares. Para controlar la presentación de titulares, puede:

- Expandir y contraer la presentación de titulares.
- Cambiar el nivel de los titulares para los elementos seleccionados.
- Cambiar el tamaño de los elementos en la presentación de titulares.
- Cambiar la fuente utilizada en la presentación de titulares.

Contracción y expansión de la vista de titulares

- ▶ Pulse en el cuadro situado a la izquierda del elemento de los titulares que desee contraer o expandir.

o

- ▶ Pulse en el elemento de los titulares.

- ▶ Elija en los menús:

Ver
Contraer

o

Ver
Expandir

Cambio del nivel de titulares

- ▶ Pulse en el elemento del panel de titulares para seleccionarlo.
- ▶ Pulse en la flecha izquierda de la barra de herramientas de titulares para ascender el elemento (mueva el elemento hacia la izquierda).
- ▶ Pulse en la flecha derecha de la barra de herramientas de titulares para degradar el elemento (mueva el elemento hacia la derecha).

o

Elija en los menús:

Edición
Resaltado
Ascender

o

Edición
Resaltado
Descender

El cambio del nivel de titulares es particularmente útil después de mover elementos en el nivel de titulares. El desplazamiento de elementos puede cambiar el nivel de titulares de los elementos seleccionados y puede utilizar los botones de flecha

izquierda y derecha de la barra de herramientas de los titulares para restaurar el nivel de titulares original.

Para cambiar el tamaño de elementos de los titulares

- ▶ Elija en los menús:
 - Ver
 - Tamaño de los titulares
 - Pequeño

Otras opciones son Mediano y Grande.

Cambiará el tamaño de los iconos y del texto asociado.

Para cambiar la fuente de los titulares

- ▶ Elija en los menús:
 - Ver
 - Fuente de los titulares
- ▶ Seleccione una fuente.

Adición de elementos al Visor

En el Visor puede añadir elementos tales como títulos, nuevo texto, gráficos o material de otras aplicaciones.

Para añadir un nuevo elemento de título o de texto al contenido del Visor

Pueden añadirse al Visor elementos de texto que no estén conectados a una tabla o a un gráfico.

- ▶ Pulse en la tabla, en el gráfico o en el otro objeto que precederá al título o al texto.

- ▶ Elija en los menús:
Insertar
 Nuevo título

- o*
- Insertar
 Nuevo texto

- ▶ Pulse dos veces en el nuevo objeto.

- ▶ Introduzca el texto que desee en esta ubicación.

Para añadir un archivo de texto

- ▶ En el panel de titulares o en el panel de contenidos del Visor, pulse en la tabla, en el gráfico o en otro objeto que vaya a preceder al texto.

 - ▶ Elija en los menús:
Insertar
 Archivo de texto...

 - ▶ Seleccione un archivo de texto.
- Para editar el texto, pulse en él dos veces.

Uso de los resultados en otras aplicaciones

Las tablas pivote y los gráficos pueden copiarse y pegarse en otra aplicación para Windows, como puede ser un procesador de textos o una hoja de cálculo. Puede pegar las tablas pivote o los gráficos en varios formatos, incluyendo los siguientes:

Objeto incrustado. Puede incrustar las tablas pivote y los gráficos interactivos en las aplicaciones que admitan objetos Active-X. Después de pegar la tabla, puede activarla en su lugar pulsando dos veces en ella y editarla posteriormente como si estuviera en el Visor.

Imagen (metarchivo). Puede pegar las tablas pivote, los resultados de texto y los gráficos como imágenes de metarchivo. El formato de la imagen puede cambiarse de tamaño en la otra aplicación y a veces determinados aspectos de la edición se pueden

realizar con las características que proporcione la otra aplicación. Las tablas pivote pegadas como imágenes conservan todas las características de bordes y fuente.

RTF (formato de texto enriquecido). Las tablas pivote pueden pegarse en otras aplicaciones con formato RTF. En la mayoría de las aplicaciones, esta opción pega la tabla pivote como una tabla que puede editarse en ellas.

Mapa de bits. Los gráficos pueden pegarse en otras aplicaciones como mapas de bits.

BIFF. El contenido de una tabla se puede pegar en una hoja de cálculo y conservar la precisión numérica.

Texto. El contenido de una tabla se puede copiar y pegar como texto. Esto puede ser útil en aplicaciones como el correo electrónico, en las que sólo se puede aceptar o transmitir texto.

Para copiar una tabla o un gráfico

- ▶ Seleccione la tabla o gráfico que va a copiar.
- ▶ Elija en los menús:
 - Edición
 - Copiar

Para copiar y pegar los resultados en otra aplicación

- ▶ Copie los resultados en el Visor.
- ▶ Elija en los menús de la aplicación de destino:
 - Edición
 - Pegar
- o*
 - Edición
 - Pegado especial...

Pegar. Los resultados se copian al Portapapeles en múltiples formatos. Cada aplicación determinará el mejor formato para la operación de pegado. En la mayoría de las aplicaciones, el comando de pegado pegará los resultados como una imagen (metarchivo). En los procesadores de textos, este comando pegará las tablas pivote

con formato RTF, el cual las pega como tablas. En las aplicaciones de hoja de cálculo, las tablas pivote de pegan con formato BIFF. Los gráficos se pegan como metarchivos.

Pegado especial. Los resultados se pegan al Portapapeles en múltiples formatos. Pegado especial le permite seleccionar el formato que desea de la lista de formatos disponibles en la aplicación de destino.

Para incrustar una tabla en otra aplicación

Es posible incrustar tablas pivote y gráficos interactivos en otras aplicaciones con formato ActiveX. El objeto incrustado puede activarse en su lugar pulsando dos veces en él, pudiendo editarse y pivotarse a continuación como en el Visor.

Si tiene aplicaciones que admitan objetos ActiveX:

- ▶ Ejecute el archivo *objs-on.bat* situado en el directorio donde se instaló SPSS. (Pulse dos veces en el archivo para ejecutarlo.)

Esta acción activará la incrustación ActiveX para las tablas pivote. El archivo *objs-off.bat* desactiva la incrustación ActiveX.

Para incrustar una tabla pivote o un gráfico interactivo en otra aplicación:

- ▶ Copie la tabla en el Visor.
- ▶ Elija en los menús de la aplicación de destino:
 - Edición
 - Pegado especial...
- ▶ Seleccione en la lista Objeto Tabla pivote de SPSS u Objeto de control de gráficos de SPSS.

La aplicación de destino debe admitir los objetos ActiveX. Consulte la documentación que se suministra con la aplicación para obtener información sobre la compatibilidad con ActiveX. Algunas aplicaciones que no admiten ActiveX pueden aceptar en un principio las tablas pivote ActiveX, sin embargo pueden presentar posteriormente un comportamiento inestable. No confíe en los objetos incrustados hasta que no haya comprobado la estabilidad de la aplicación con los objetos ActiveX incrustados.

Para pegar una tabla pivote o un gráfico como imagen (Metarchivo)

- ▶ Copie la tabla o el gráfico en el Visor.
- ▶ Elija en los menús de la aplicación de destino:
 - Edición
 - Pegado especial...
- ▶ Seleccione Imagen en la lista que aparece.

El elemento se pegará como metarchivo. Sólo la capa y las columnas visibles cuando se copió el elemento estarán disponibles en el metarchivo. Las otras capas o las columnas ocultas no estarán disponibles.

Para pegar una tabla pivote como una tabla (RTF)

- ▶ Copie la tabla pivote en el Visor.
- ▶ Elija en los menús de la aplicación de destino:
 - Edición
 - Pegado especial...
- ▶ En la lista que aparece, seleccione Texto con formato (RTF) o Formato de texto enriquecido.

Al realizar esta operación, la tabla pivote se pegará como una tabla. Sólo se pegarán en la tabla la capa y las columnas que estaban visibles cuando se copió el elemento. Las otras capas o las columnas ocultas no estarán disponibles. Con este formato sólo es posible copiar y pegar una tabla pivote cada vez.

Para pegar una tabla pivote como texto

- ▶ Copie la tabla en el Visor.
- ▶ Elija en los menús de la aplicación de destino:
 - Edición
 - Pegado especial...
- ▶ Seleccione Texto sin formato en la lista.

El texto sin formato de la tabla pivote contiene tabuladores entre columnas. Las columnas se pueden alinear ajustando los tabuladores en la otra aplicación.

Copiado y pegado de múltiples elementos en otra aplicación

- ▶ Seleccione las tablas y/o gráficos que vaya a copiar. Mantenga pulsada la tecla Mayús o Ctrl para seleccionar múltiples elementos.
- ▶ Elija en los menús:
 - Edición
 - Copiar objetos
- ▶ En la aplicación de destino, elija en los menús:
 - Edición
 - Pegar

Nota: Utilice la opción Copiar objetos sólo para copiar varios elementos del Visor en otra aplicación. Para copiar y pegar dentro de los documentos del Visor (por ejemplo, entre dos ventanas del Visor), utilice el comando Copiar del menú Edición.

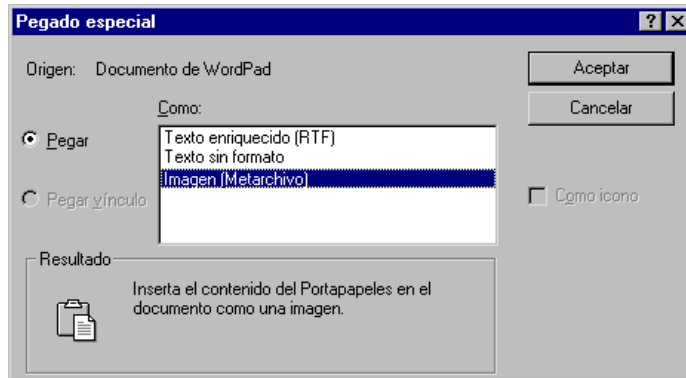
Pegado de objetos en el Visor

Es posible pegar objetos de otras aplicaciones en el Visor. Puede utilizar Pegar debajo o Pegado especial. Cualquiera de estos tipos de pegado coloca el nuevo objeto después del objeto actualmente seleccionado en el Visor. Utilice Pegado especial cuando desee seleccionar el formato del objeto pegado.

Pegado especial

Pegado especial permite seleccionar el formato de un objeto copiado que se pega en el Visor. Se muestran en una lista los posibles tipos de archivo para el objeto en el Portapapeles. El objeto se insertará en el Visor después del objeto actualmente seleccionado.

Figura 9-3
Cuadro de diálogo Pegado especial



Para pegar un objeto de otra aplicación en el Visor

- ▶ Copie el objeto de la otra aplicación.
- ▶ En el panel de titulares o en el panel de contenidos del Visor, pulse en la tabla, el gráfico u otro objeto que precederá al objeto que se dispone a pegar.
- ▶ Elija en los menús:
 - Edición
 - Pegado especial...
- ▶ En la lista, seleccione el formato para el objeto.

Exportar resultados

Exportar resultados guarda las tablas pivote y los resultados de texto en formato HTML, de texto, Word/RTF, Excel y PowerPoint (requiere PowerPoint 97 o posterior) y guarda los gráficos en una amplia variedad de formatos comunes utilizados por otras aplicaciones.

Documento de salida. Exporta cualquier combinación de tablas pivote, resultados de texto y gráficos.

- Para los formatos HTML y de texto, los gráficos se exportan en el formato de exportación de gráficos seleccionado actualmente. Si se trata de documentos HTML, los gráficos se incrustan mediante referencias y deben exportarse a un formato adecuado para incluirlos en los documentos HTML. En el caso de documentos de texto se inserta una línea en el archivo de texto para cada gráfico, indicando el nombre de archivo del gráfico exportado.
- Para los formatos Word o RTF, los gráficos se exportan en formato de metarchivo de Windows y se incrustan en el documento de Word.
- No se incluyen gráficos en los documentos de Excel.
- Para el formato PowerPoint, los gráficos se exportan en formato TIFF y se incrustan en el archivo de PowerPoint.

Documento de resultados (sin gráficos). Exporta tablas pivote y resultados de texto. Se ignoran todos los gráficos que se muestran en el Visor.

Sólo gráficos. Los formatos de exportación disponibles incluyen: metarchivo de Windows (WMF), metarchivo mejorado (EMF), mapa de bits de Windows (BMP), PostScript encapsulado (EPS), JPEG, TIFF, PNG y PICT de Macintosh.

Exportar qué. Permite exportar todos los objetos del Visor, todos los objetos visibles o sólo los objetos seleccionados.

Formato de exportación. Para los documentos de resultados, las opciones disponibles son HTML, texto, Word/RTF, Excel y PowerPoint; para los formatos HTML y de texto, los gráficos se exportan en el formato de gráfico seleccionado actualmente en el cuadro de diálogo Opciones para el formato seleccionado. Para Sólo gráficos, seleccione un formato de exportación del gráfico en la lista desplegable. Para exportar documentos de resultados, tablas pivote y texto, siga las instrucciones siguientes:

- **Archivo HTML (*.htm).** Las tablas pivote se exportan como tablas HTML. Los resultados de texto se exportan como formato previo de HTML.
- **Archivo de texto (*.txt).** Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Todos los resultados de texto se exportan en formato separado por espacios.
- **Archivo de Excel (*.xls).** Las filas, columnas y casillas de tablas pivote se exportan como filas, columnas y casillas de Excel, con todos los atributos de formato intactos (por ejemplo, bordes de casilla, estilos de fuente, colores de fondo, etc.). Los resultados de texto se exportan con todos los atributos de fuente intactos.

Cada línea del resultado de texto constituye una fila del archivo de Excel y se incluye todo su contenido en una sola casilla.

- **Archivo de Word/RTF (*.doc).** Las tablas pivote se exportan como tablas de Word, con todos los atributos de formato intactos (por ejemplo, bordes de casilla, estilos de fuente, colores de fondo, etc.). Los resultados de texto se exportan en formato RTF. Los resultados de texto siempre se muestran en SPSS con un tipo de letra de paso fijo (monoespaciadas) y se exportan con los mismos atributos de fuente. Es necesario utilizar una fuente de paso fijo para que los resultados de texto separado por espacios queden correctamente alineados; por defecto, la mayoría de los navegadores de Web utilizan una fuente de paso fijo para el texto con formato previo.
- **Archivo PowerPoint (*.ppt).** Las tablas pivote se exportan como tablas de Word y se incrustan en diapositivas independientes del archivo de PowerPoint (con una diapositiva para cada tabla pivote). Se conservan todos los atributos de formato de la tabla (por ejemplo, bordes de casillas, estilos de fuente, colores de fondo, etc.) Los resultados de texto se exportan en formato RTF. Los resultados de texto siempre se muestran en SPSS con un tipo de letra de paso fijo (monoespaciadas) y se exportan con los mismos atributos de fuente. Es necesario utilizar una fuente de paso fijo para que los resultados de texto separado por espacios queden correctamente alineados; por defecto, la mayoría de los navegadores de Web utilizan una fuente de paso fijo para el texto con formato previo.

Sistema de gestión de resultados. Puede exportar automáticamente además todos los resultados o tipos de resultados especificados por el usuario como texto, HTML, XML y archivos de datos en formato SPSS. Si desea obtener más información, consulte “Sistema de gestión de resultados” en Capítulo 47 en p. 695.

Exportación de resultados

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
 - Archivo
 - Exportar...
- ▶ Introduzca un nombre de archivo (o prefijo para los gráficos) y seleccione un formato de exportación.

Figura 9-4
Cuadro de diálogo *Exportar resultados*

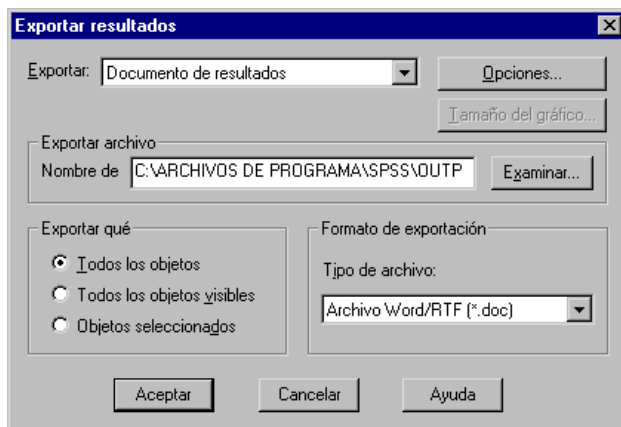
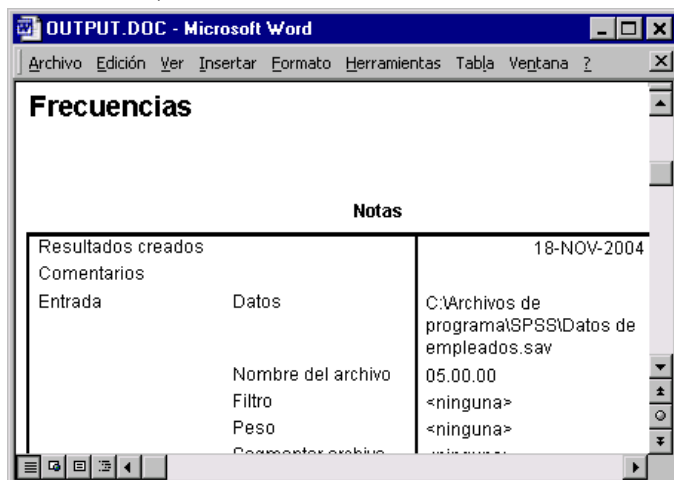


Figura 9-5
Resultados exportados en formato Word/RTF



Opciones de HTML, Word/RTF y Excel

Este cuadro de diálogo controla la inclusión de notas y textos al pie para los documentos exportados en los formatos HTML, Word/RTF y Excel, así como las opciones de exportación de gráficos para los documentos en HTML y la gestión de tablas pivote de varias capas.

Formato de imagen. Controla el formato de exportación de los gráficos y otros ajustes, incluido el tamaño de los gráficos para documentos HTML. Para los formatos Word o RTF, todos los gráficos se exportan en formato de metarchivo de Windows (WMF). No se incluyen gráficos para los documentos de Excel.

Mantener las notas y el texto al pie al exportar. Seleccione esta casilla para incluir las notas y el texto al pie en la exportación de tablas pivote.

Exportar todas las capas. Seleccione esta casilla para exportar todas las capas de una tabla pivote de varias capas. Si no se selecciona, sólo se exporta la capa superior.

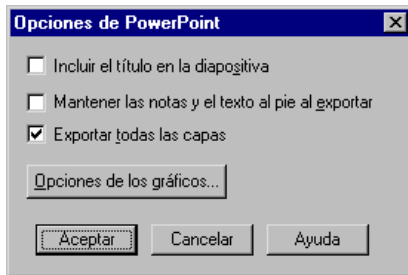
Para establecer los ajustes de exportación de HTML, Word/RTF y Excel

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
 - Archivo
 - Exportar...
- ▶ Seleccione Archivo HTML, Archivo Word/RTF o Archivo Excel como formato de exportación.
- ▶ Pulse en Opciones.

Opciones de PowerPoint

Opciones de PowerPoint controla la inclusión de los títulos de diapositivas, la inclusión de notas y textos al pie para las tablas pivote, la gestión de tablas pivote de varias capas y las opciones par los gráficos exportados a PowerPoint.

Figura 9-6
Cuadro de diálogo Opciones de PowerPoint



Incluir el título en la diapositiva. Seleccione esta casilla para incluir un título en cada diapositiva creada mediante la exportación. Cada diapositiva contiene un único elemento exportado del Visor. El título se genera a partir de la entrada del titular para el elemento en el panel de titulares del Visor.

Mantener las notas y el texto al pie al exportar. Seleccione esta casilla para incluir las notas y el texto al pie en la exportación de tablas pivote.

Exportar todas las capas. Seleccione esta casilla para exportar todas las capas de una tabla pivote de varias capas; cada capa se coloca en una diapositiva distinta y todas las capas tienen el mismo título. Si no se selecciona, sólo se exporta la capa superior.

Opciones del texto

El cuadro de diálogo Opciones del texto controla las opciones de formato de los gráficos, de los resultados de texto y de las tablas pivote y la inclusión de notas y textos al pie para los documentos exportados en formato de texto.

Figura 9-7
Cuadro de diálogo *Opciones del texto*

Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Para el formato separado por tabuladores, si una casilla no está vacía, se imprimirá su contenido y un tabulador. Si una casilla está vacía, se imprimirá un tabulador.

Todos los resultados de texto se exportan en formato separado por espacios. Todos los resultados separados por espacios requieren una fuente de paso fijo (monoespaciada) para su correcta alineación.

Formato de casilla. En las tablas pivote separadas por espacios se eliminan por defecto todas las líneas de separación y se establece el ancho mediante la etiqueta o el valor más largo de la columna. Para limitar el ancho de las columnas y ajustar las etiquetas largas, especifique un número de caracteres para el ancho de la columna. Este ajuste sólo afecta a las tablas pivote.

Separador de casillas. En las tablas pivote separadas por espacios puede especificar los caracteres empleados para crear los bordes de las casillas.

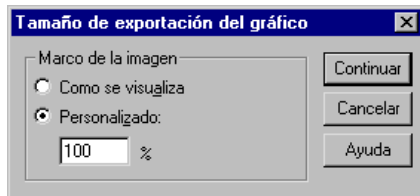
Formato de imagen. Controla el formato de exportación de los gráficos y otros ajustes, incluyendo el tamaño de los gráficos.

Insertar salto de página entre tablas. Inserta un avance de página o salto de página entre cada tabla. Para las tablas pivote de varias capas, inserta un salto de página entre cada capa.

Opciones de tamaño de gráfico

Tamaño del gráfico controla el tamaño de los gráficos exportados. Si especifica un porcentaje personalizado, podrá reducir o ampliar el tamaño del gráfico exportado hasta un porcentaje máximo de 200 por ciento.

Figura 9-8
Cuadro de diálogo Tamaño de exportación del gráfico



Definición del tamaño de los gráficos exportados

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
 - Archivo
 - Exportar...
- ▶ En el caso de los documentos de resultados, pulse en Opciones, seleccione el formato de exportación y pulse en Tamaño del gráfico.
- ▶ En el caso de Sólo gráficos, seleccione el formato de exportación y pulse en Tamaño del gráfico.

Opciones de exportación de gráficos JPEG

Profundidad de color. Los gráficos JPEG se pueden exportar como color verdadero (24 bits) o escala de 256 grises.

Espacio de color. Espacio de color hace referencia al modo en que los colores están codificados en la imagen. El modelo de color YUV es una forma de codificación de color, utilizada normalmente para la transmisión por vídeo digital y MPEG. El acrónimo significa señal Y, señal U, señal V. El componente Y especifica la escala de grises o luminancia, y los componentes U y V se corresponden con la crominancia (información de color).

Las relaciones representan las tasas de muestreo de cada componente. Si se reducen las tasas de muestreo U y V, se reduce a su vez el tamaño de los archivos (y también la calidad). Espacio de color determina el grado de “pérdida” de los colores de la imagen exportada. YUV 4:4:4 implica una menor pérdida, mientras que YUV 4:2:2 y YUV 4:1:1 representan un equilibrio decreciente entre el tamaño de los archivos (espacio en disco) y la calidad de los colores representados.

Codificación progresiva. Activa la imagen para realizar la carga por etapas, mostrándola al principio a baja resolución y aumentando a continuación la calidad conforme continúa cargándose la imagen.

Configuración de calidad de compresión. Controla la relación de compresión de la calidad de la imagen. Cuanto mejor es la calidad de la imagen, mayor es el tamaño del archivo exportado.

Operaciones con el color. Se encuentran disponibles las siguientes operaciones:

- **Invertir.** Cada píxel se guarda como la inversa del color original.
- **Corrección de gamma.** Ajusta la intensidad de los colores en el gráfico exportado mediante el cambio de la constante gamma utilizada para asignar los valores de intensidad. Se puede utilizar fundamentalmente para iluminar u oscurecer la imagen de mapa de bits. El valor puede estar comprendido entre 0,10 (el más oscuro) y 6,5 (el más claro).

Opciones de exportación de gráficos BMP y PICT

Profundidad de color. Determina el número de colores del gráfico exportado. Un gráfico guardado con cualquier profundidad tendrá un mínimo del número de colores utilizados y un máximo del número de colores permitidos por la profundidad. Por ejemplo, si el gráfico incluye tres colores (rojo, blanco y negro) y se guarda como un gráfico de 16 colores, permanecerá como gráfico de tres colores.

- Si el número de colores del gráfico es superior al número de colores para dicha profundidad, los colores se interpolarán para reproducir los colores del gráfico.
- Profundidad de la pantalla actual es el número de colores mostrados actualmente en el monitor del ordenador.

Operaciones con el color. Se encuentran disponibles las siguientes operaciones:

- **Invertir.** Cada píxel se guarda como la inversa del color original.
- **Corrección de gamma.** Ajusta la intensidad de los colores en el gráfico exportado mediante el cambio de la constante gamma utilizada para asignar los valores de intensidad. Se puede utilizar fundamentalmente para iluminar u oscurecer la imagen de mapa de bits. El valor puede estar comprendido entre 0,10 (el más oscuro) y 6,5 (el más claro).

Utilizar compresión RLE. (Sólo BMP). Se trata de una técnica de compresión de menor pérdida admitida por los formatos de archivo comunes de Windows. La compresión de menor pérdida implica que la calidad de la imagen no se sacrifica en detrimento de archivos de menor tamaño.

Opciones de exportación de gráficos PNG y TIFF

Profundidad de color. Determina el número de colores del gráfico exportado. Un gráfico guardado con cualquier profundidad tendrá un mínimo del número de colores utilizados y un máximo del número de colores permitidos por la profundidad. Por ejemplo, si el gráfico incluye tres colores (rojo, blanco y negro) y se guarda como un gráfico de 16 colores, permanecerá como gráfico de tres colores.

- Si el número de colores del gráfico es superior al número de colores para dicha profundidad, los colores se interpolarán para reproducir los colores del gráfico.
- Profundidad de la pantalla actual es el número de colores mostrados actualmente en el monitor del ordenador.

Operaciones con el color. Se encuentran disponibles las siguientes operaciones:

- **Invertir.** Cada píxel se guarda como la inversa del color origina.
- **Corrección de gamma.** Ajusta la intensidad de los colores en el gráfico exportado mediante el cambio de la constante gamma utilizada para asignar los valores de intensidad. Se puede utilizar fundamentalmente para iluminar u oscurecer la imagen de mapa de bits. El valor puede estar comprendido entre 0,10 (el más oscuro) y 6,5 (el más claro).

Transparencia. Permite seleccionar un color que se mostrará como transparente en el gráfico exportado. Se encuentra disponible sólo con una exportación de color verdadero de 32 bits. Introduzca valores enteros entre 0 y 255 para cada color. El valor por defecto para cada color es 255, y se crea un color blanco transparente por defecto.

Formato. (Sólo TIFF). Permite definir el espacio de color y comprimir el gráfico exportado. Todas las profundidades de color están disponibles en color RVA. Sólo se encuentra disponible un color verdadero de 24 y 32 bits para CMYK. Con la opción YCbCr, sólo se encuentra disponible un color verdadero de 24 bits.

Opciones de exportación de gráficos EPS

Árboles, mapas y gráficos interactivos

Para los árboles (opción Árboles de clasificación), los mapas (opción Mapas) y para los gráficos interactivos (menú Gráficos, submenú Interactivo) se encuentran disponibles las siguientes opciones de EPS:

Vista previa de las imágenes. Permite guardar una imagen de vista previa en la imagen EPS. Una imagen de vista previa se utiliza principalmente si se ubica un archivo EPS en otro documento. Muchas aplicaciones no pueden mostrar ninguna imagen EPS en la pantalla, pero sí pueden mostrar la vista previa guardada con la imagen. La vista previa de la imagen puede ser WMF (de menor tamaño y más escalable) o TIFF

(portátil y admitida por otras plataformas). Compruebe la aplicación en que desea incluir el gráfico EPS para ver qué formato de vista previa admite.

Fuentes. Controla el tratamiento de fuentes TrueType en las imágenes EPS.

- **Incrustar como TrueType nativo.** Incrusta la mayoría de los datos de fuentes en el EPS. La fuente PostScript resultante se denomina fuente de tipo 42 (Type 42). *Nota:* No todas las impresoras PostScript tienen controladores de nivel 3 que puedan leer las fuentes de tipo 42.
- **Convertir en fuentes PostScript.** Convierte las fuentes TrueType en fuentes PostScript (Tipo 1) basadas en la familia de fuentes. Por ejemplo, Times New Roman se convierte en Times y Arial se convierte en Helvetica. *Nota:* Este formato no se recomienda para los gráficos interactivos que utilizan la fuente del marcador de SPSS (por ejemplo, diagramas de dispersión) dado que no existe ningún equivalente PostScript representativo para los símbolos del marcador TrueType de SPSS.
- **Sustituir fuentes con curvas.** Convierte las fuentes TrueType en datos de curvas PostScript. El texto ya no se puede editar como texto en las aplicaciones que pueden editar los gráficos EPS. Se produce asimismo una pérdida de calidad, pero esta opción resulta útil si dispone de una impresora PostScript que no admite las fuentes del tipo 42 (Type 42) y necesita conservar símbolos especiales TrueType, como es el caso de los marcadores utilizados en los diagramas de dispersión interactivos.

Otros gráficos

Para el resto de gráficos, están disponibles las siguientes opciones EPS:

Incluir vista preliminar TIFF. Guarda una vista preliminar con la imagen EPS image en formato TIFF para su visualización en aplicaciones que no pueden mostrar imágenes EPS en la pantalla.

Fuentes. Controla el tratamiento de fuentes en las imágenes EPS.

- **Sustituir fuentes con curvas.** Convierte las fuentes en datos de curvas PostScript. El texto ya no se puede editar como texto en las aplicaciones que pueden editar los gráficos EPS. Esta opción es útil si las fuentes utilizadas en el gráfico no están disponibles en el dispositivo de los resultados.
- **Utilizar referencias de fuentes.** Si las fuentes utilizadas en el gráfico están disponibles en el dispositivo de resultados, se utilizan las fuentes. En caso contrario, el dispositivo de resultados utiliza fuentes alternativas.

Opciones de exportación de gráficos WMF

Válido para Aldus. Proporciona cierto grado de independencia al dispositivo (el mismo tamaño físico que cuando se abre a 96 x 120 ppp), pero no todas las aplicaciones admiten este formato.

Windows estándar. Lo admiten la mayoría de las aplicaciones que pueden mostrar metarchivos de Windows.

Configuración de las opciones de exportación de gráficos

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
 - Archivo
 - Exportar...
- ▶ En el caso de los documentos de resultados, pulse en Opciones, seleccione el formato de exportación y pulse en Opciones de los gráficos.
- ▶ En caso de Sólo gráficos, seleccione el formato de exportación y pulse en Opciones.

Impresión de documentos del Visor

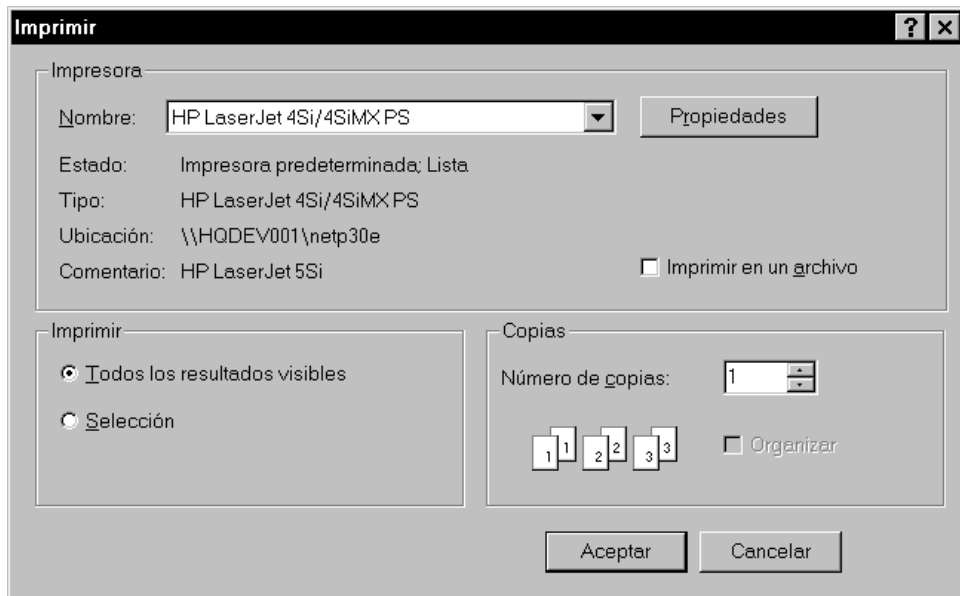
Hay dos opciones para imprimir el contenido de la ventana del Visor:

Todos los resultados visibles. Se imprimen sólo los elementos mostrados actualmente en el panel de contenidos. No se imprimen los elementos ocultos (los elementos con un icono de libro cerrado en el panel de titulares o los ocultos en las capas de titulares contraídas).

Selección. Se imprimen sólo los elementos actualmente seleccionados en los paneles de titulares y de contenidos.

Figura 9-9

Cuadro de diálogo Imprimir del Visor



Impresión de resultados y gráficos

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).

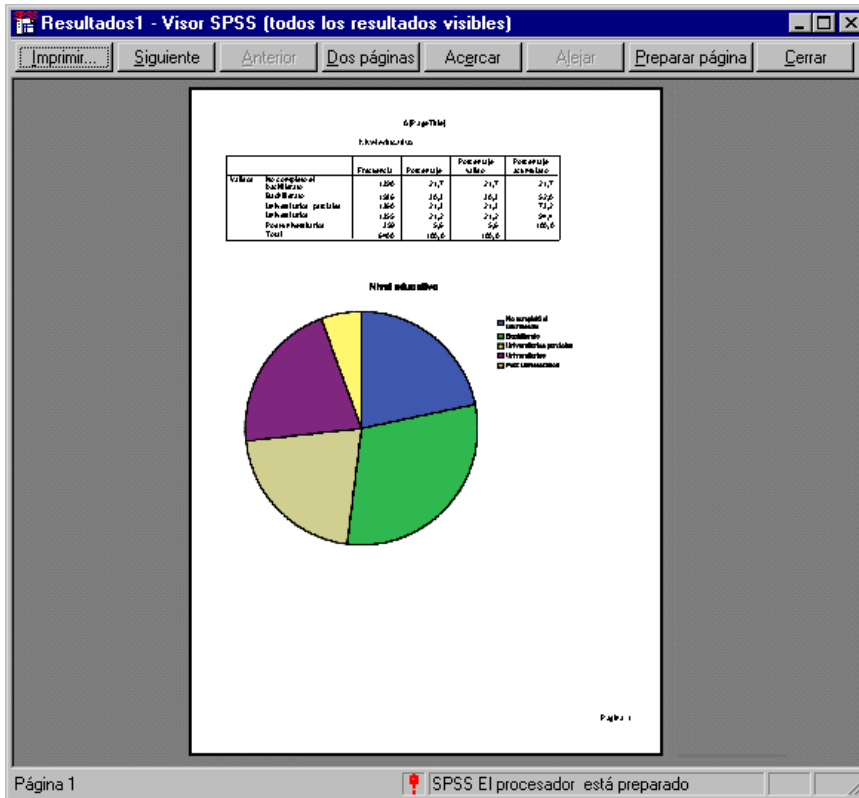
- ▶ Elija en los menús:
 - Archivo
 - Imprimir...
- ▶ Seleccione los ajustes de impresión que desee.
- ▶ Pulse en Aceptar para imprimir.

Presentación preliminar

Presentación preliminar muestra lo que se imprimirá en cada página de los documentos del Visor. Es recomendable comprobar la presentación preliminar antes de imprimir un documento del Visor, ya que muestra elementos que no puedan verse a simple vista en el panel de contenidos; entre ellos:

- Los saltos de página
- Las capas ocultas de las tablas pivote
- Las rupturas de las tablas anchas
- Los resultados completos de las tablas de gran tamaño
- Los encabezados y pies impresos en cada página

Figura 9-10
Presentación preliminar



Si se han seleccionado resultados en el Visor, la presentación preliminar sólo mostrará estos resultados. Si desea ver una presentación preliminar de todos los resultados, asegúrese de que no haya nada seleccionado en el Visor.

Para ver una presentación preliminar

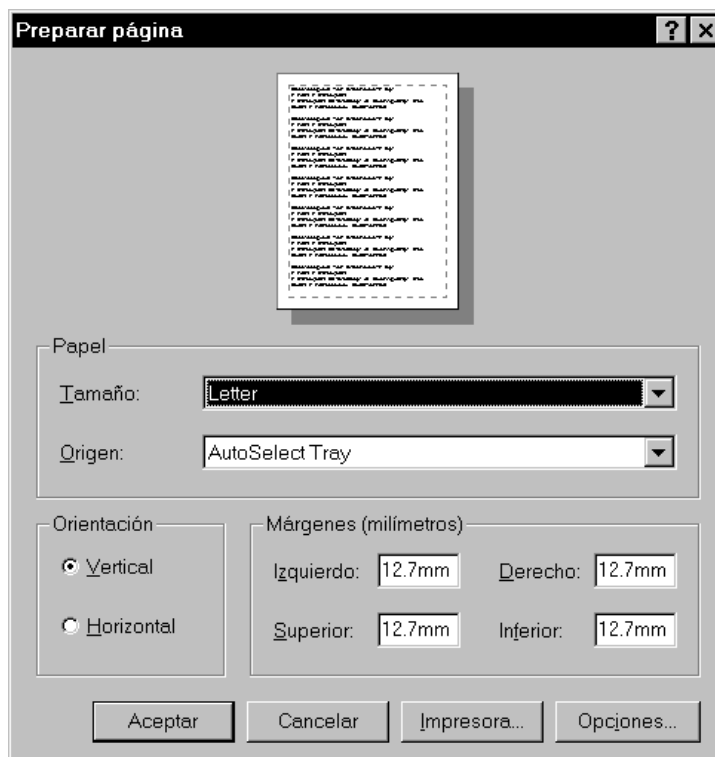
- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
 - Archivo
 - Presentación preliminar

Configuración de página

Con Preparar página, se puede controlar:

- El tamaño y la orientación del papel
- Los márgenes de la página
- Los encabezados y pies de página
- Numeración de páginas
- El tamaño impreso para los gráficos

Figura 9-11
Cuadro de diálogo Preparar página



La configuración de Preparar página se guarda con el documento del Visor. Preparar página solamente afecta a la configuración para la impresión de documentos del Visor. Esta configuración no tiene ningún efecto en la impresión de los datos del Editor de datos ni en la sintaxis de una ventana de sintaxis.

Para modificar la preparación de página

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
 - Archivo
 - Preparar página
- ▶ Cambie los ajustes y pulse en Aceptar.

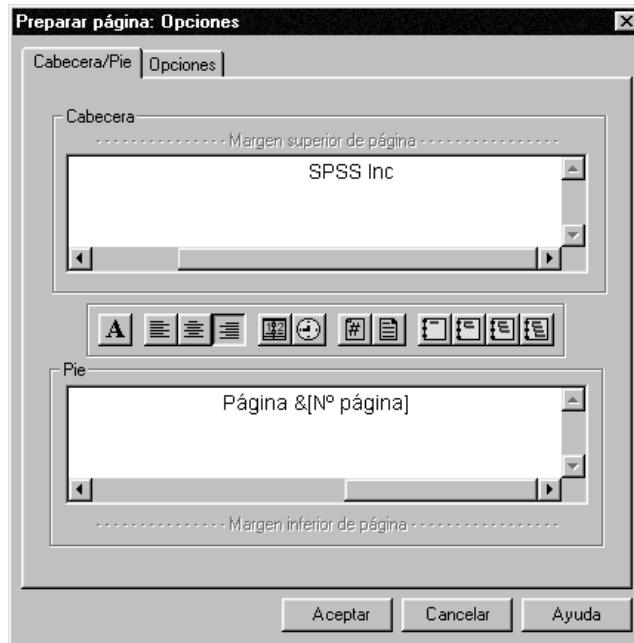
Opciones de Preparar página: Encabezados y pies

Los encabezados y los pies son la información que se imprime en la parte superior e inferior de cada página. Puede introducir cualquier texto que desee utilizar como encabezados y pies. También puede utilizar la barra de herramientas, situada en medio del cuadro de diálogo, para insertar:

- La fecha y la hora
- Los números de páginas
- El nombre del archivo del Visor
- Las etiquetas de los encabezados de los titulares
- Los títulos y subtítulos de página.

Figura 9-12

Pestaña Cabecera/Pie del cuadro de diálogo Opciones de Preparar página:



Las etiquetas de las cabeceras de los titulares indican el primer, el segundo, el tercer y/o el cuarto nivel de cabecera del titular para el primer elemento en cada página.

Los títulos y los subtítulos de página imprimen los títulos y subtítulos de página actuales. Los títulos y los subtítulos de página se crean con la opción Insertar nuevo título de página del menú Insertar del Visor o con los comandos `TITLE` y `SUBTITLE` de la sintaxis de comandos. Si no ha especificado ningún título ni subtítulo de página, este ajuste no se tendrá en cuenta.

Nota: Las características de las fuentes de los nuevos títulos y subtítulos de página se controlan en la pestaña Visor del cuadro de diálogo Opciones (menú Edición). También se pueden cambiar las características de los títulos y subtítulos de página existentes editándolos en el Visor.

Utilice Presentación preliminar en el menú Archivo para ver cómo aparecerán los encabezados y los pies en la página impresa.

Opciones de Preparar página: Opciones

Este cuadro de diálogo controla el tamaño de los gráficos impresos, el espacio entre los elementos de resultado impresos y la numeración de las páginas.

Tamaño del gráfico impreso. Controla el tamaño del gráfico impreso relativo al tamaño de la página definido. La relación de aspecto de los gráficos (proporción anchura-altura) no se ve afectada por el tamaño del gráfico impreso. El tamaño global impreso de un gráfico está limitado tanto por su altura como por su anchura. Una vez que los bordes externos de un gráfico alcancen los bordes izquierdo y derecho de la página, el tamaño del gráfico no se podrá aumentar más para completar la altura de página adicional.

Espacio entre los elementos. Controla el espacio entre los elementos impresos. Cada tabla pivote, gráfico y objeto de texto es un elemento diferente. Este ajuste no afecta a la presentación de los elementos en el Visor.

Numerar las páginas empezando por. Numera las páginas secuencialmente empezando por el número especificado.

Figura 9-13
Pestaña Opciones del cuadro de diálogo Opciones de Preparar página:



Almacenamiento de resultados

El contenido del Visor puede guardarse en un documento del Visor. El documento incluirá los dos paneles de la ventana del Visor (titulares y contenidos).

Para guardar un documento del Visor

- ▶ Seleccione en los menús de la ventana del Visor:
 - Archivo
 - Guardar
- ▶ Escriba el nombre del documento y pulse en Guardar.

Para guardar los resultados en otros formatos externos (por ejemplo, HTML o texto), utilice el comando Exportar del menú Archivo. (No está disponible en el programa Smart Viewer independiente.)

Opción Guardar con contraseña

Guardar con contraseña le permite proteger con contraseña los archivos en el Visor.

Contraseña. La contraseña tiene en cuenta las mayúsculas y las minúsculas y puede tener hasta dieciséis caracteres. Si asigna una contraseña, el archivo no podrá verse hasta que se introduzca la contraseña.

Código OEM. Deje este campo en blanco a menos que tenga un acuerdo contractual con SPSS Inc. para distribuir SmartViewer. Con el contrato se proporciona un código de licencia OEM.

Para guardar los archivos del Visor con una contraseña

- ▶ Seleccione en los menús de la ventana del Visor:
 - Archivo
 - Guardar con contraseña...
- ▶ Escriba la contraseña.
- ▶ Vuelva a introducir la contraseña para confirmarla y pulse en Aceptar.
- ▶ Escriba un nombre de archivo en el cuadro de diálogo Guardar como.
- ▶ Pulse en Guardar.

Nota: Deje en blanco el espacio destinado al código OEM a menos que tenga un acuerdo contractual con SPSS Inc. para distribuir Smart Viewer.

Visor de borrador

El Visor de borrador ofrece los resultados a modo de borrador, incluyendo:

- Resultados de sólo texto (en lugar de tablas pivote)
- Gráficos como imágenes de metarchivo (en lugar de objetos de gráfico)

Puede editar los resultados de texto en el Visor de borrador, cambiar el tamaño de los gráficos y pegar tanto los resultados de texto como los gráficos en otras aplicaciones. Sin embargo, no es posible editar los gráficos ni disponer de las funciones interactivas de las tablas pivote y los gráficos.

Figura 10-1
Ventana Visor de borrador

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Este	120	30,8	30,8	30,8
	Central	161	41,3	41,3	72,1
	Oeste	109	27,9	27,9	100,0
	Total	390	100,0	100,0	

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Este	120	30,8	30,8	30,8
	Central	161	41,3	41,3	72,1
	Oeste	109	27,9	27,9	100,0
	Total	390	100,0	100,0	

Para crear resultados de borrador

- ▶ Elija en los menús:
 - Archivo
 - Nuevo
 - Resultados de borrador

- ▶ Para convertir los resultados de borrador en el tipo de resultado por defecto, elija en los menús:
 - Edición
 - Opciones...

- ▶ Pulse en la pestaña General.

- ▶ Seleccione Borrador en Tipo de visor al inicio.

Nota: Los nuevos resultados se muestran siempre en la ventana designada del Visor. Si tiene abiertas una ventana del Visor y una del Visor de borrador, la **ventana designada** será la que se haya abierto más recientemente o la designada mediante la herramienta Designar ventana (el signo de exclamación) de la barra de herramientas.

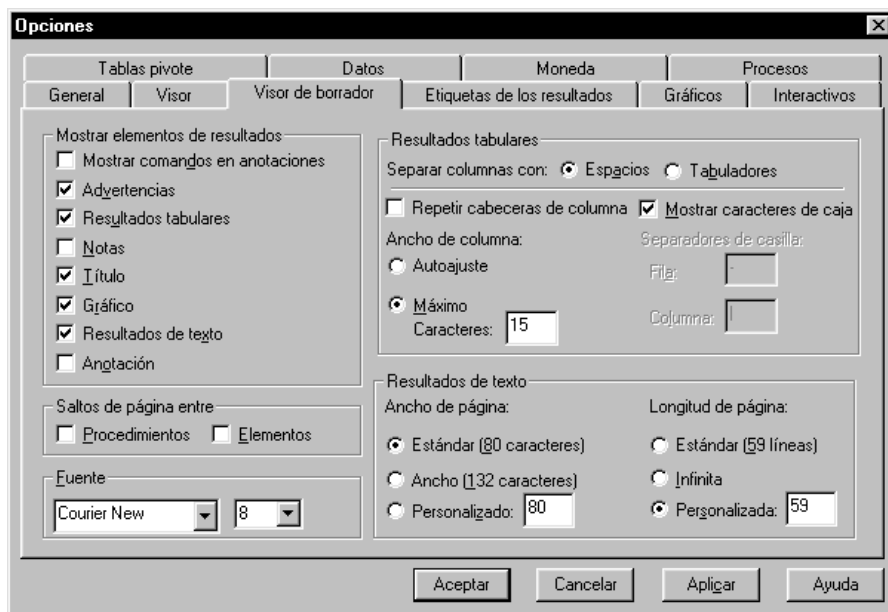
Control del formato de los resultados de borrador

Los resultados que aparecen como tablas pivote en el Visor se convierten a resultados de texto en el Visor de borrador. La configuración por defecto de las tablas pivote convertidas es la siguiente:

- El ancho de cada columna se ajusta al de la etiqueta de la columna y las etiquetas no se dividen en varias líneas.
- La alineación se controla mediante espacios en lugar de tabuladores.
- Los caracteres de caja de la fuente SPSS Marker Set se utilizan para los separadores de filas y columnas.
- Si los caracteres de caja estuvieran desactivados, se empleará el carácter de línea vertical (|) para los separadores de columna y los guiones (–) para los separadores de filas.

Puede controlar el formato de los nuevos resultados de borrador mediante las opciones del Visor de borrador (menú Edición, Opciones, pestaña Visor de borrador).

Figura 10-2
Opciones del Visor de borrador



Ancho de columna. Para reducir el ancho de las tablas que contienen etiquetas de gran longitud, seleccione la opción de máximo de caracteres en Ancho de columna. Las etiquetas que exceden el ancho especificado se ajustan a este valor máximo aplicando líneas de separación.

Figura 10-3

Resultados de borrador antes y después de establecer el ancho máximo de columna

Tabla con Autoajuste en la anchura de las columnas

		Frecuencia	Porcentaje	Porcentaje válido	Porcenta
Válidos	Productos para consumo	179	45,9	45,9	45,9
	Productos financieros	211	54,1	54,1	100,0
	Total	390	100,0	100,0	

Tabla con anchura máxima de columnas fijada en 12 caracteres

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Productos para consumo	179	45,9	45,9	45,9
	Productos financieros	211	54,1	54,1	100,0
	Total	390	100,0	100,0	

Separadores de filas y de columnas. En vez de los caracteres de caja para los bordes de filas y columnas, se pueden utilizar las opciones de Separadores de casillas para controlar la presentación de los separadores de filas y columnas en los resultados de borrador. Puede especificar diferentes separadores de casillas o introducir espacios en blanco si no desea emplear ningún carácter para marcar las filas y las columnas. Tendrá que anular la selección de Mostrar caracteres de caja para especificar los separadores de casillas.

Figura 10-4

Resultados de borrador antes y después de establecer los separadores de casillas

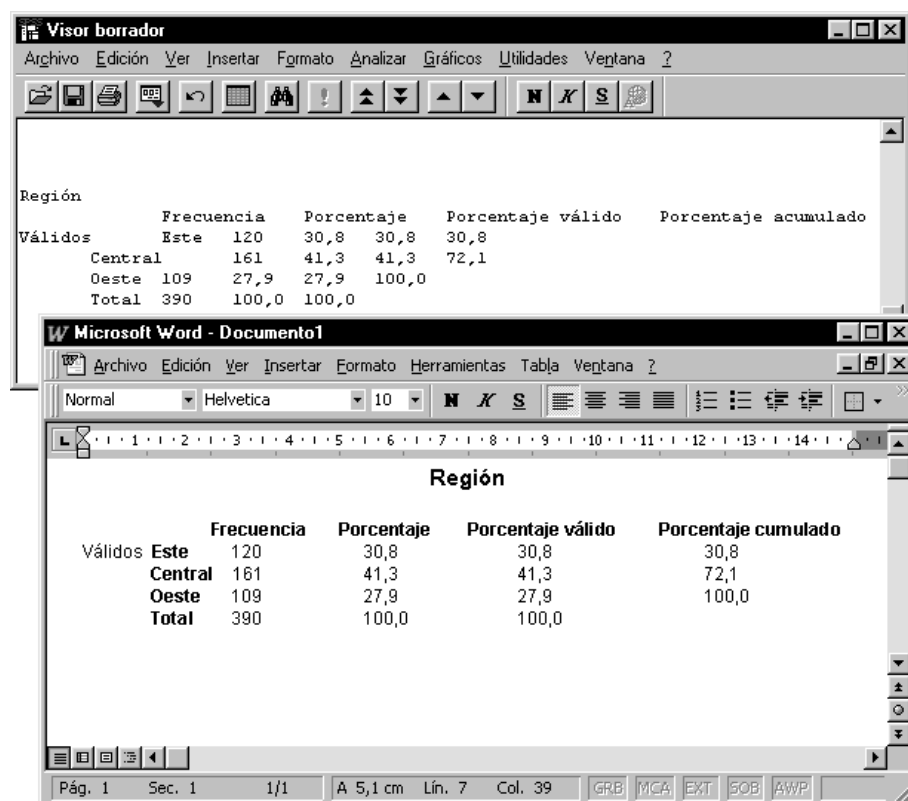
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Este	120	30,8	30,8
	Central	161	41,3	72,1
	Oeste	109	27,9	100,0
	Total	390	100,0	

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Este	120	30,8	30,8
	Central	161	41,3	72,1
	Oeste	109	27,9	100,0
	Total	390	100,0	

Columnas separadas por espacios frente a columnas separadas por tabuladores. El Visor de borrador se ha diseñado para presentar los resultados separados por espacios en una fuente de paso fijo (monoespaciada). Si desea pegar los resultados de borrador en otra aplicación, deberá utilizar este tipo de fuente para que las columnas separadas por espacios queden bien alineadas. Si selecciona Tabuladores como separadores de columnas, podrá utilizar cualquier fuente que desee en la otra aplicación y definir los tabuladores para alinear los resultados correctamente. Sin embargo, en el Visor de borrador, los resultados separados por tabuladores no se alinearán correctamente.

Figura 10-5

Resultados separados por tabuladores en el Visor de borrador y formateados en un procesador de textos.



Para establecer las opciones del Visor de borrador

- ▶ Elija en los menús:
Edición
Opciones...
- ▶ Pulse en la pestaña Visor de borrador.
- ▶ Seleccione la configuración que desee.
- ▶ Pulse en Aceptar o Aplicar.

Las opciones de presentación de resultados del Visor de borrador sólo afectan a los resultados producidos tras el cambio de las especificaciones. Los cambios en la configuración no afectan al resultado mostrado en el Visor de borrador.

Atributos de fuente

Es posible modificar los atributos de fuente (como fuente, tamaño y estilo) de los resultados de texto del Visor de borrador. En cualquier caso, si utiliza caracteres de caja para los bordes de las filas y las columnas, para que la alineación del texto separado por espacio en las columnas sea correcta, será necesaria una fuente de paso fijo (monoespaciada), por ejemplo, Courier. Asimismo, la modificación de otros atributos de fuente como el tamaño y el estilo (por ejemplo, negrita o cursiva) sólo en una parte de la tabla puede afectar también a la alineación de las columnas.

Bordes de filas y columnas. Los bordes de línea sólidos para las filas y las columnas utiliza, por defecto, la fuente SPSS Marker Set. Las demás fuentes no permiten los caracteres de dibujo de líneas utilizados para trazar los bordes.

Para cambiar las fuentes en el Visor de borrador

- ▶ Seleccione el texto al que desee aplicar el cambio.
- ▶ En los menús del Visor de borrador, seleccione:
 - Formato
 - Fuente...
- ▶ Seleccione los atributos de fuente que desee aplicar al texto seleccionado.

Para imprimir resultados de borrador

- ▶ En los menús del Visor de borrador, seleccione:
 - Archivo
 - Imprimir...

Para imprimir sólo una parte seleccionada de los resultados de borrador:

- ▶ Seleccione los resultados que desea imprimir.

- ▶ Elija en los menús:
 - Archivo
 - Imprimir...
- ▶ Elija Selección.

Presentación preliminar en el Visor de borrador

Presentación preliminar muestra el aspecto que presentará cada página impresa de los documentos de borrador. Es recomendable comprobar la presentación preliminar antes de imprimir un documento del Visor porque permite ver los elementos que pudieran no ajustarse al tamaño de la página, entre ellos:

- Tablas de gran longitud.
- Tablas muy anchas producidas por los resultados de tablas pivote convertidas sin especificación del ancho de columna.
- Resultados de texto creados con la opción Ancho de la sección del ancho de página (opciones del Visor de borrador) y con la orientación de impresión definida en Vertical.

Los resultados demasiado anchos para la página no se imprimen en otra página, sino que se truncan. Existen varios métodos para evitar esto:

- Utilice un tamaño de fuente más pequeño (menú Formato, Fuentes).
- Seleccione la opción Horizontal de orientación de página (menú Archivo, Preparar página).
- Para los nuevos resultados, especifique un valor bajo para el ancho de columna máximo (menú Edición, Opciones, pestaña Visor de borrador).

En las tablas de gran longitud, utilice saltos de página (menú Insertar, Salto de página) para controlar la ubicación de los saltos en la tabla.

Para obtener una presentación preliminar de los resultados de borrador

- ▶ En los menús del Visor de borrador, seleccione:
 - Archivo
 - Presentación preliminar

Almacenamiento de los resultados del Visor de borrador

- ▶ En los menús del Visor de borrador, seleccione:

- Archivo
- Guardar

Los resultados del Visor de borrador se guardan con formato RTF (formato de texto enriquecido).

Para guardar los resultados de borrador como texto

- ▶ En los menús del Visor de borrador, seleccione:

- Archivo
- Exportar...

Puede exportar todo el texto o sólo la porción seleccionada. En los archivos exportados sólo se almacenan los resultados de texto (los resultados de las tablas pivote convertidas y los resultados de texto); no se incluyen los gráficos.

Tablas pivote

Muchos de los resultados del Visor se presentan en tablas que se pueden pivotar de forma interactiva. Es decir, puede reorganizar las filas, columnas y capas.

Gestión de una tabla pivote

Entre las opciones para manipular una tabla pivote se incluyen:

- Transposición de filas y columnas
- Desplazamiento de filas y columnas
- Creación de capas multidimensionales
- Agrupación y separación de filas y columnas
- Presentación y ocultación de casillas
- Rotación de etiquetas de fila y de columna
- Búsqueda de definiciones de términos

Para editar una tabla pivote

- ▶ Pulse dos veces en la tabla.

Al realizar esta acción se activará el Editor de tablas pivote.

Para editar dos o más tablas pivote al mismo tiempo

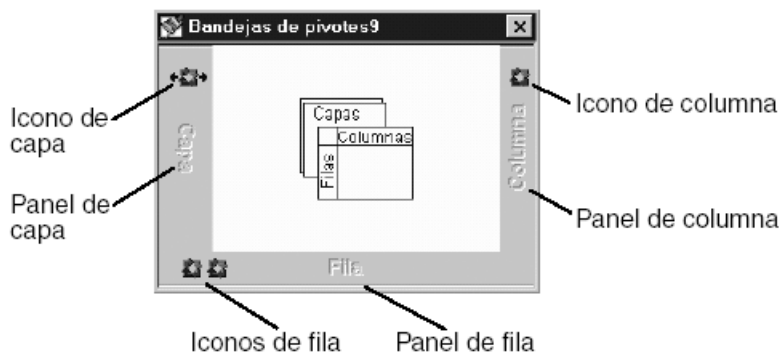
- ▶ Pulse en la tabla pivote con el botón derecho del ratón.

- ▶ Elija en el menú contextual:
Objeto Tabla pivote de SPSS
Abrir
- ▶ Repita los pasos para cada tabla pivote que desee editar.
Cada tabla pivote aparecerá en una ventana independiente lista para ser editada.

Para pivotar una tabla utilizando iconos

- ▶ Active la tabla pivote.
- ▶ Elija en los menús de la tabla pivote:
Pivotar
Paneles de pivotado
- ▶ Sitúe el puntero del ratón sobre cada icono para obtener una pista emergente que le indicará la dimensión de la tabla que representa el icono.
- ▶ Arrastre un icono de un panel a otro.

Figura 11-1
Paneles de pivotado



Este procedimiento permite cambiar la disposición de la tabla. Por ejemplo, suponga que el icono representa una variable con las categorías Sí y No y que lo arrastra desde el panel Fila hasta el panel Columna. Antes de arrastrarlo, Sí y No eran etiquetas de fila; después de moverlo son etiquetas de columna.

Para identificar las dimensiones de una tabla pivote

- ▶ Active la tabla pivote.
- ▶ Si los paneles de pivotado no están activados, seleccione en los menús de la tabla:
Pivotar
 Paneles de pivotado
- ▶ Pulse y mantenga pulsado el botón del ratón en un icono.
De este modo se resaltan las etiquetas de dimensión de la tabla pivote.

Para transponer filas y columnas

- ▶ Elija en los menús de la tabla pivote:
Pivotar
 Transponer filas y columnas
- Esta operación tiene el mismo efecto que si se arrastraran todos los iconos de fila hasta el panel Columna y todos los iconos de columna hasta el panel Fila.

Para cambiar el orden de presentación

El orden de los iconos de pivotado en un panel de dimensión refleja el orden de presentación de los elementos en la tabla pivote. Para cambiar el orden de presentación de los elementos de una dimensión:

- ▶ Active la tabla pivote.
- ▶ Si los paneles de pivotado no está activados, elija en los menús de la tabla:
Pivotar
 Paneles de pivotado
- ▶ Arrastre los iconos de cada panel y colóquelos en el orden que desee (de izquierda a derecha o de arriba abajo).

Para mover filas y columnas en una tabla pivote

- ▶ Active la tabla pivote.

- ▶ Pulse en la etiqueta de la fila o la columna que desee mover.
- ▶ Pulse y arrastre la etiqueta hasta la nueva posición.
- ▶ En el menú contextual, seleccione Insertar antes o Intercambiar.

Nota: Asegúrese de que Arrastrar para copiar en el menú Edición *no* no está activado (marcado). Si está activada, anule su selección.

Para agrupar filas o columnas e insertar etiquetas de grupo

- ▶ Active la tabla pivote.
- ▶ Seleccione las etiquetas de las filas o las columnas que desee agrupar; para ello, pulse y arrastre, o bien mantenga pulsada la tecla Mayús para seleccionar varias etiquetas.
- ▶ Elija en los menús:
 - Edición
 - Grupo

Se insertará automáticamente una etiqueta de grupo. Pulse dos veces en la etiqueta de grupo para editar el texto de la etiqueta.

Figura 11-2

Etiquetas y grupos de filas y columnas

		Column Group Label		
		Female	Male	Total
Row Group Label	Clerical	206	157	363
	Custodial		27	27
	Manager	10	74	84

Nota: Para añadir filas o columnas a un grupo existente, deberá desagrupar los elementos actuales del grupo y crear después un nuevo grupo que incluya los elementos adicionales.

Para desagrupar filas o columnas y eliminar etiquetas de grupo

- ▶ Active la tabla pivote.

- ▶ Seleccione la etiqueta de grupo de las filas o columnas que desea desagrupar; para ello, pulse en cualquier lugar de la etiqueta.
- ▶ Elija en los menús:
Edición
Desagrupar

La desagrupación eliminará automáticamente la etiqueta de grupo.

Para rotar etiquetas de tablas pivote

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
Formato
Rotar etiquetas de columna interior

o

Rotar etiquetas de fila exterior

Figura 11-3
Etiquetas de columna rotadas

	Frequency	Percent	Valid Percent	Cumulative Percent
Clerical	363	76.6	76.6	76.6
Custodial	27	5.7	5.7	82.3
Manager				
Total				

	Frequency	Percent	Valid Percent	Cumulative Percent
Clerical	363	76.6	76.6	76.6
Custodial	27	5.7	5.7	82.3
Manager	84	17.7	17.7	100.0
Total	474	100.0	100.0	

Sólo se pueden rotar las etiquetas de columna más al interior y las etiquetas de fila más al exterior.

Para restablecer las opciones por defecto de las tablas pivote

Después de realizar una o varias operaciones de pivotado, puede recuperar la disposición original de la tabla pivote.

- ▶ En el menú Pivotar, seleccione Restablecer pivotes por defecto.

Esta operación sólo restablece los cambios que son el resultado de pivotar los elementos de fila, columna y capa entre dimensiones. No afecta a cambios tales como la agrupación, la desagrupación o el desplazamiento de filas y columnas.

Para buscar la definición de una etiqueta en una tabla pivote

Puede obtener ayuda sensible al contexto sobre las etiquetas de las casillas de las tablas pivote. Por ejemplo, si *Media* aparece como una etiqueta, puede obtener una definición de la media.

- ▶ Pulse en una casilla de etiqueta con el botón derecho del ratón.
- ▶ En el menú contextual, seleccione ¿Qué es esto?.

Debe pulsar con el botón derecho del ratón en la propia casilla de etiqueta, no en las casillas de datos de la fila o la columna.

La ayuda sensible al contexto no se encuentra disponible para las etiquetas definidas por el usuario, como pueden ser los nombres de variable o las etiquetas de valor.

Trabajo con capas

Es posible mostrar una tabla de dos dimensiones distinta para cada categoría o combinación de categorías. La tabla se puede considerar como una apilación de capas en la que sólo puede verse la capa superior.

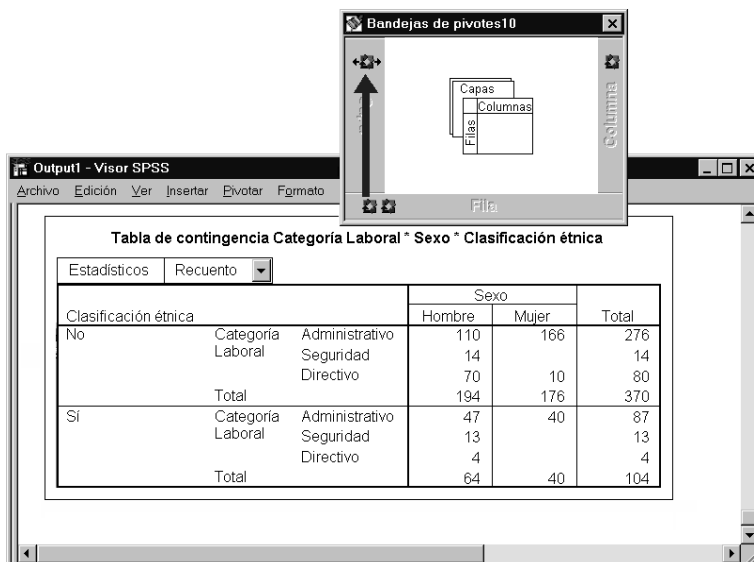
Para crear y mostrar capas

- ▶ Active la tabla pivote y seleccione la opción Paneles de pivotado en el menú Pivotar, si no está seleccionada.

- Arrastre un icono desde el panel Fila o Columna hasta el panel Capa.

Figura 11-4

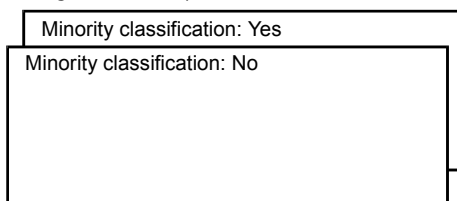
Desplazamiento de categorías a las capas



Cada icono de capa tiene flechas izquierda y derecha. La tabla visible es la tabla de la capa superior.

Figura 11-5

Categorías en capas distintas



Para cambiar las capas

- ▶ Pulse en una de las flechas del icono de capas.
- o*
- ▶ Seleccione una categoría de la lista desplegable de capas.

Figura 11-6
Selección de capas en las listas desplegables

Informes en capas						
Región	Oeste					
	Central					
	Oeste	N	Media	Desv. típ.	% de la suma total	% del total de N
División	Total	109	1,65	,48	30,0%	27,9%

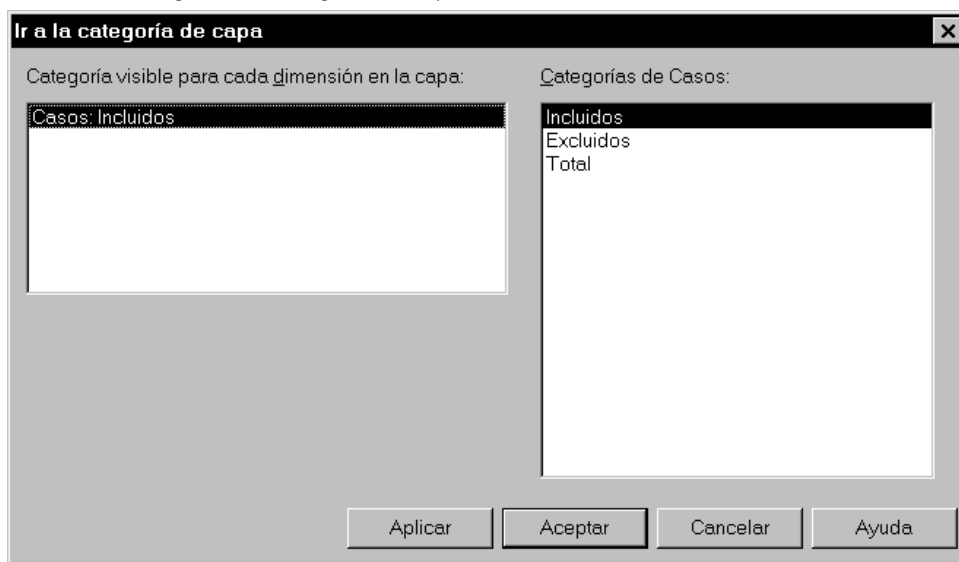
Ir a la categoría de capa

Ir a la categoría de capa permite cambiar las capas de una tabla pivote. Este cuadro de diálogo es particularmente útil cuando hay un gran número de capas o una capa con muchas categorías.

Para ir a una capa de una tabla

- ▶ Elija en los menús de la tabla pivote:
 - Pivotar
 - Ir a capa...

Figura 11-7
Cuadro de diálogo Ir a la categoría de capa



- ▶ Seleccione una dimensión de capa en la lista de categorías visibles. La lista Categorías mostrará todas las categorías de la dimensión seleccionada.
- ▶ Seleccione la categoría que desee en la lista Categorías y pulse en Aceptar. Esta acción cambiará la capa y cerrará el cuadro de diálogo.
Para ver otra capa sin cerrar el cuadro de diálogo:
- ▶ Seleccione la categoría y pulse en Aplicar.

Para mover las capas a las filas o a las columnas

Si la tabla que aparece en la pantalla está apilada en capas y sólo puede ver la superior, podrá mostrar todas las capas a la vez, verticalmente en filas o bien horizontalmente en columnas. Debe haber como mínimo un icono en el panel Capa.

- ▶ En el menú Pivotar, seleccione Mover capas a filas.

o

- ▶ En el menú Pivotar, seleccione Mover capas a columnas.

También puede mover las capas a las filas o a las columnas arrastrando sus iconos entre los paneles de pivotado Capa, Fila y Columna.

Señalizadores

Con los señalizadores puede guardar diferentes presentaciones de una tabla pivote. Los señalizadores guardan:

- Las posiciones de los elementos en las dimensiones de fila, columna y capa
- El orden de presentación de los elementos en cada dimensión
- La capa mostrada actualmente para cada elemento de capa

Para señalar vistas de tablas pivote

- ▶ Active la tabla pivote.
- ▶ Pivote la tabla hasta la vista que desee señalar.
- ▶ Elija en los menús:
 - Pivotar
 - Señalizadores...
- ▶ Escriba un nombre para el señalizador. Los nombres de los señalizadores no distinguen entre mayúsculas y minúsculas.
- ▶ Pulse en Añadir.

Cada tabla pivote tiene su propio conjunto de señalizadores. Los nombres no se pueden repetir en la misma tabla pivote, pero sí en tablas diferentes.

Para mostrar una vista de tabla pivote con señalizador

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
Pivotar
Señalizadores...
- ▶ Pulse en el nombre del señalizador en la lista.
- ▶ Pulse en Ir a.

Para cambiar el nombre de un señalizador de una tabla pivote

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
Pivotar
Señalizadores...
- ▶ Pulse en el nombre del señalizador en la lista.
- ▶ Pulse en Cambiar nombre.
- ▶ Escriba el nuevo nombre.
- ▶ Pulse en Aceptar.

Presentación y ocultación de casillas

Se pueden ocultar muchos tipos de casillas:

- Las etiquetas de dimensión
- Las categorías, incluidas la casilla de la etiqueta y las casillas de datos de una fila o columna
- Las etiquetas de las categorías (sin ocultar las casillas de datos)
- Las notas al pie, los títulos y los pies

Para ocultar las filas y las columnas de una tabla

- ▶ Pulse las teclas Ctrl+Alt y en la etiqueta de categoría de la fila o columna que desee ocultar.
- ▶ Elija en los menús de la tabla pivote:
Ver
Ocult.
o
- ▶ Pulse con el botón derecho del ratón en la fila o columna resaltada para mostrar el menú contextual.
- ▶ En el menú contextual, seleccione Ocultar categoría.

Para mostrar las filas y columnas ocultas de una tabla

- ▶ Seleccione otra etiqueta de la misma dimensión que la fila o columna ocultas.
Por ejemplo, si la categoría *Mujer* de la dimensión Sexo está oculta, pulse en la categoría *Hombre*.
- ▶ Elija en los menús de la tabla pivote:
Ver
Mostrar todas las categorías en nombre de la dimensión
Por ejemplo, seleccione Mostrar todas las categorías en Sexo.
o
- ▶ Elija en los menús de la tabla pivote:
Ver
Mostrar todo
Al hacerlo se mostrarán todas las casillas ocultas de la tabla; Si está seleccionada la opción Ocultar filas y columnas vacías en Propiedades de tabla para esta tabla, las filas o columnas completamente vacías permanecerán ocultas.

Para ocultar o mostrar una etiqueta de dimensión

- ▶ Active la tabla pivote.
- ▶ Seleccione la etiqueta de dimensión o cualquier etiqueta de categoría dentro de la dimensión.
- ▶ Elija en los menús:
 - Ver
 - Ocultar (o Mostrar) etiqueta de dimensión

Para ocultar o mostrar una nota al pie de una tabla

- ▶ Seleccione una nota al pie.
- ▶ Elija en los menús:
 - Ver
 - Ocultar (o Mostrar)

Para ocultar o mostrar un pie o un título de una tabla

- ▶ Seleccione un texto al pie o un título.
- ▶ Elija en los menús:
 - Ver
 - Ocultar (o Mostrar)

Edición de los resultados

Es posible editar el aspecto y el contenido de cada tabla o elemento de resultados de texto. Tiene la posibilidad de:

- Aplicar un aspecto de tabla.
- Modificar las propiedades de la tabla actual.
- Modificar las propiedades de las casillas de la tabla.
- Modificar el texto.
- Añadir notas y textos al pie a las tablas.

- Añadir elementos al Visor.
- Copiar y pegar los resultados en otras aplicaciones.

Cambio del aspecto de las tablas

Puede cambiar el aspecto de una tabla editando sus propiedades o aplicando un aspecto de tabla. Cada aspecto de tabla contiene un conjunto de propiedades de tabla, relativas al aspecto general, los bordes, las propiedades de las notas al pie y las propiedades de las casillas. Puede seleccionar uno de los aspectos de tabla preestablecidos o crear y guardar uno personalizado.

Aspectos de tabla

Un aspecto de tabla es un conjunto de propiedades que definen el aspecto de una tabla. Puede seleccionar un aspecto previamente definido o crear uno propio.

Antes o después de aplicar un aspecto de tabla, se pueden cambiar los formatos de casillas individuales o de grupos de casillas utilizando las propiedades de casilla. Los formatos de casilla editados se conservan aunque se aplique un nuevo aspecto de tabla.

Por ejemplo, puede comenzar aplicando el aspecto de tabla *9POINT*; a continuación, seleccionar una columna de datos y, en el cuadro de diálogo Formatos de casilla, cambiar dicha columna a una fuente en negrita. Después, puede cambiar el aspecto de tabla a *BOXED*. La columna previamente seleccionada conservará la fuente en negrita, mientras que el resto de las características se aplicarán con el aspecto de tabla *BOXED*.

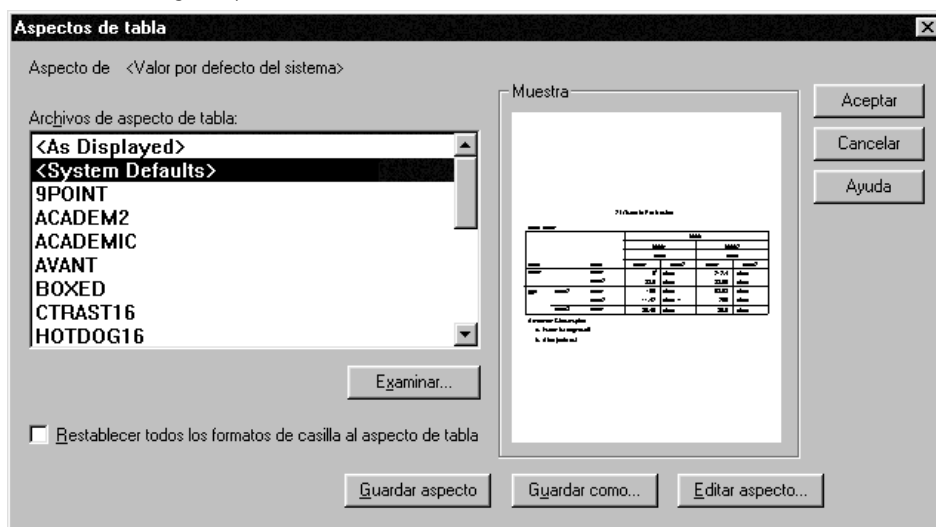
Si lo desea, puede restablecer todas las casillas a los formatos de casilla definidos por el aspecto de tabla actual. De este modo se restablece cualquier casilla que se haya editado. Si la opción Como se visualiza está seleccionada en la lista de archivos de aspecto de tabla, las casillas editadas se restablecerán a las propiedades de tabla actuales.

Para aplicar o guardar un aspecto de tabla

- ▶ Active una tabla pivote.

- ▶ Elija en los menús:
Formato
Aspectos de tabla...

Figura 11-8
Cuadro de diálogo Aspectos de tabla



- ▶ Seleccione un aspecto de tabla en la lista de archivos. Para seleccionar un archivo de otro directorio, pulse en Examinar.
- ▶ Pulse en Aceptar para aplicar el aspecto de tabla a la tabla pivote seleccionada.

Para editar o crear un aspecto de tabla

- ▶ Seleccione un aspecto de tabla en la lista de archivos.
- ▶ Pulse en Editar aspecto.
- ▶ Edite las propiedades de la tabla seleccionando los atributos que desee y pulse en Aceptar.
- ▶ Pulse en Guardar aspecto para guardar el aspecto de tabla editado, o en Guardar como para guardarlo como un nuevo aspecto de tabla.

La edición de un aspecto de tabla sólo afecta a la tabla pivote seleccionada. El aspecto de tabla editado no se aplicará a ninguna otra tabla que utilice ese aspecto, a menos que se seleccione esa tabla y se le vuelva a aplicar el aspecto.

Propiedades de tabla

El cuadro de diálogo Propiedades de tabla permite establecer las propiedades generales de una tabla, definir los estilos de casilla de varias partes de la tabla y guardar un conjunto de esas propiedades como un aspecto de tabla. Utilizando las pestañas de este cuadro de diálogo puede:

- Controlar las propiedades generales, como ocultar filas o columnas vacías y ajustar las propiedades de impresión.
- Controlar el formato y la posición de los marcadores de las notas al pie.
- Determinar formatos específicos para las casillas en el área de datos, para las etiquetas de fila y columna y para otras áreas de la tabla.
- Controlar la anchura y el color de las líneas que forman los bordes de cada área de la tabla.
- Controlar las propiedades de impresión.

Para modificar las propiedades de las tablas pivote

- ▶ Active la tabla pivote (pulsando dos veces en cualquier punto de la tabla).
- ▶ Elija en los menús de la tabla pivote:
 - Formato
 - Propiedades de tabla...
- ▶ Elija una pestaña (General, Notas al pie, Formatos de casilla, Bordes o Impresión).
- ▶ Seleccione las opciones que desee.
- ▶ Pulse en Aceptar o Aplicar.

Las nuevas propiedades se aplicarán a la tabla pivote seleccionada. Para aplicar nuevas propiedades de tabla a un aspecto de tabla en lugar de hacerlo a la tabla seleccionada, edite el aspecto de tabla (menú Formato, Aspectos de tabla).

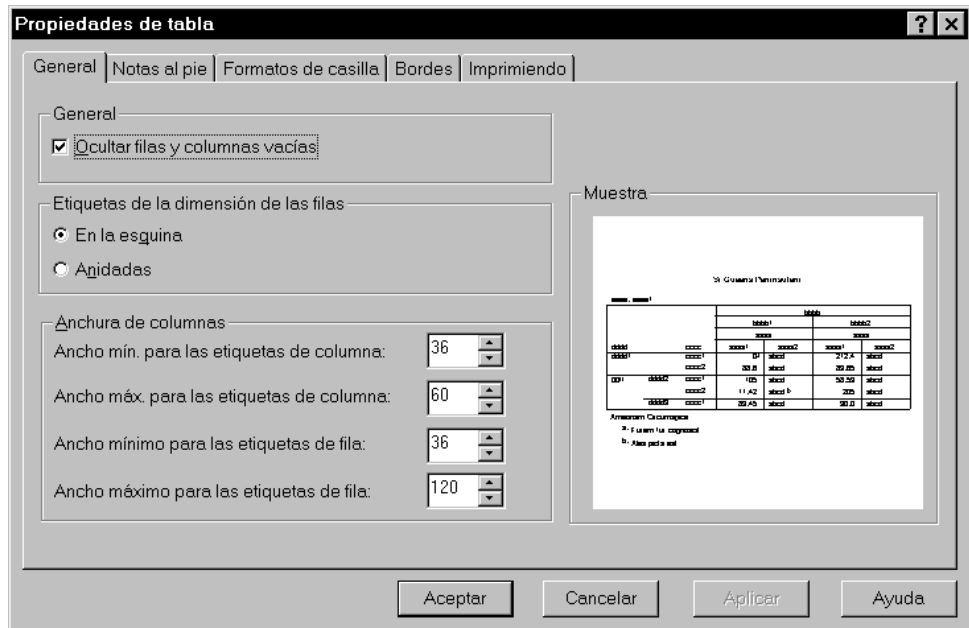
Propiedades de tabla: General

Algunas propiedades son aplicables a la tabla en su conjunto. Tiene la posibilidad de:

- Mostrar u ocultar filas y columnas vacías. (Una fila o una columna vacía no contiene nada en ninguna casilla de datos.)
- Controlar la colocación de las etiquetas de fila. Pueden estar anidadas o en la esquina superior izquierda.
- Controlar la anchura máxima y mínima de la columna (expresada en puntos).

Figura 11-9

Pestaña General del cuadro de diálogo Propiedades de tabla



Para modificar las propiedades generales de las tablas

- ▶ Seleccione la pestaña General.
- ▶ Seleccione las opciones que desee.
- ▶ Pulse en Aceptar o Aplicar.

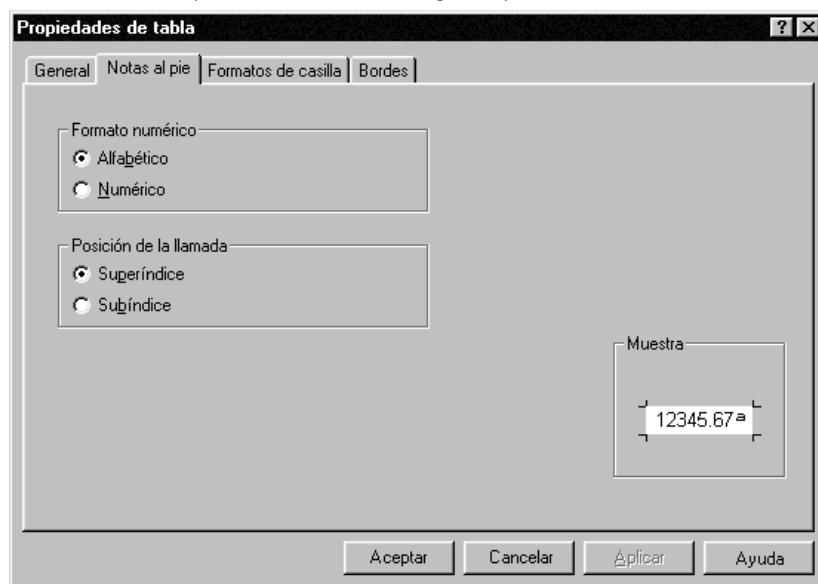
Propiedades de tabla: Notas al pie

Las propiedades de los marcadores de las notas al pie incluyen el estilo y la posición en relación al texto.

- El estilo de los marcadores de las notas al pie puede ser números (1, 2, 3...) o letras (a, b, c...).
- Los marcadores de las notas al pie se pueden anexar al texto como superíndices o como subíndices.

Figura 11-10

Pestaña *Notas al pie* del cuadro de diálogo *Propiedades de tabla*



Para modificar las propiedades de los marcadores de las notas al pie

- ▶ Seleccione la pestaña *Notas al pie*.
- ▶ Elija el formato del marcador.
- ▶ Seleccione la posición del marcador.

- Pulse en Aceptar o Aplicar.

Propiedades de tabla: Formatos de casilla

Para el formato, una tabla se divide en áreas: título, capas, etiquetas de esquina, etiquetas de fila, etiquetas de columna, datos, texto al pie y notas al pie. Para cada área de una tabla se pueden modificar los formatos de casilla asociados. Los formatos de casilla incluyen características del texto (como fuente, tamaño, color y estilo), alineación horizontal y vertical, sombreado de las casillas, colores de la máscara y el fondo, y márgenes internos de las casillas.

Figura 11-11
Áreas de una tabla

Título	
Capas	
Etiqueta de Esquina	Etiqueta de Columnas
Etiqueta de Filas	Datos

Texto al pie

Notas al pie

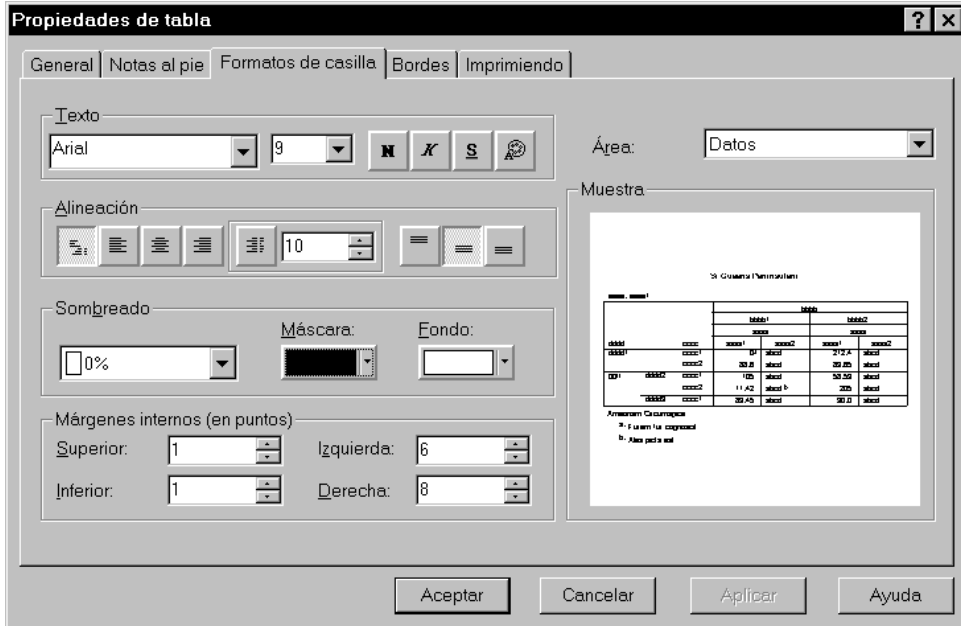
Los formatos de casilla se aplican a las áreas (categorías de información); no son características de las casillas individuales. No son características de casillas individuales. Esta distinción es un aspecto importante a la hora de pivotar una tabla.

Por ejemplo:

- Si especifica una fuente en negrita como formato de casilla de las etiquetas de columna, éstas aparecerán en negrita independientemente de la información que se muestre actualmente en la dimensión de columna; además, si mueve un elemento de la dimensión de columna a otra dimensión, no conservará la característica de negrita de las etiquetas de columna.
- Si pone en negrita las etiquetas de columna simplemente resaltando las casillas de una tabla pivote activada y pulsando en el botón Negrita de la barra de herramientas, el contenido de esas casillas permanecerá en negrita independientemente de la dimensión a las que las mueva, y las etiquetas de

columna no conservarán la característica de negrita para otros elementos movidos a la dimensión de columna.

Figura 11-12
Pestaña *Formatos de casilla* de cuadro de diálogo *Propiedades de tabla*



Para cambiar los formatos de casilla

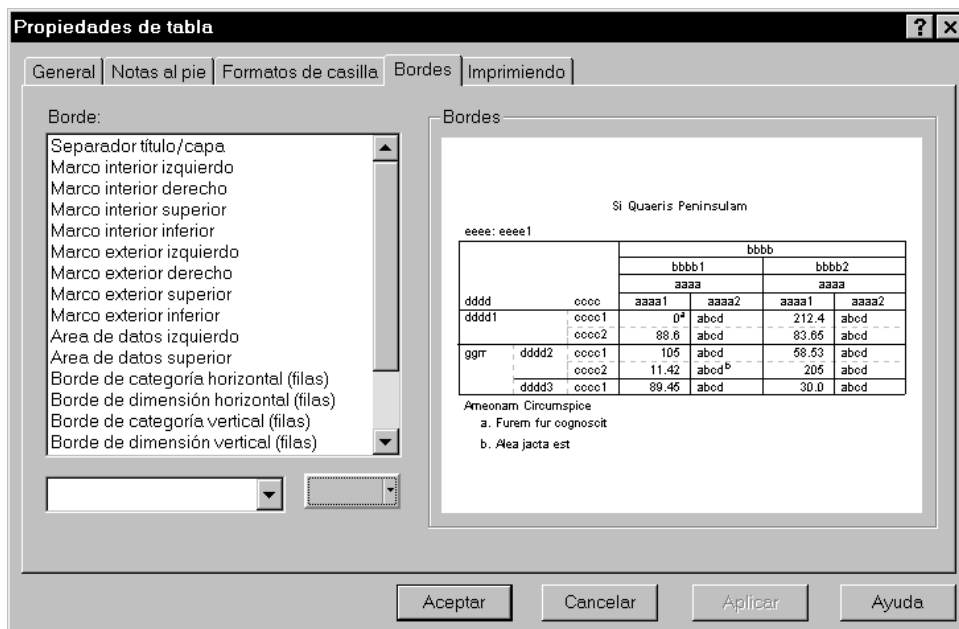
- ▶ Seleccione la pestaña *Formatos de casilla*.
- ▶ Elija un área de la lista desplegable o pulse en un área de la muestra.
- ▶ Seleccione las características para el área. Sus selecciones se reflejarán en la muestra.
- ▶ Pulse en *Aceptar* o *Aplicar*.

Propiedades de tabla: Bordes

Para cada ubicación del borde de una tabla, puede seleccionar un estilo de línea y un color. Si selecciona Ninguno como estilo, no habrá ninguna línea en la ubicación seleccionada.

Figura 11-13

Pestaña Bordes del cuadro de diálogo Propiedades de tabla



Para cambiar los bordes de una tabla

- ▶ Pulse en la pestaña Bordes.
- ▶ Seleccione una ubicación del borde pulsando en el nombre que aparece en la lista o en una línea del área de muestra (mantenga pulsada la tecla Mayús para seleccionar varios nombres o la tecla Ctrl para seleccionar nombres no contiguos). (mantenga pulsada la tecla Mayús para seleccionar varios nombres o la tecla Ctrl para seleccionar nombres no contiguos).
- ▶ Seleccione un estilo de línea o Ninguno.

- ▶ Elija un color.
- ▶ Pulse en Aceptar o Aplicar.

Para mostrar los bordes ocultos de una tabla pivote

En las tablas que tienen pocos bordes visibles es posible mostrar los bordes ocultos. Esto puede facilitar la realización de algunas tareas, como el cambio de los anchos de columna. Los bordes ocultos (líneas de cuadrícula) aparecen en el Visor pero no se imprimen.

- ▶ Active la tabla pivote (pulsando dos veces en cualquier punto de la tabla).
- ▶ Elija en los menús:
 - Ver
 - Cuadrículas

Propiedades de tabla: Imprimiendo

Puede controlar las siguientes propiedades de la impresión de las tablas pivote:

- Imprimir todas las capas o sólo la capa superior de la tabla e imprimir cada capa en una página distinta (esto sólo afecta a la impresión, no a la presentación de las capas en el Visor).
- Reducir una tabla horizontal o verticalmente para que se ajuste a la página a la hora de imprimirla.
- Controlar las líneas viudas y huérfanas; es decir, el número mínimo de filas y columnas que podrá incluir cualquier sección impresa de una tabla si ésta es demasiado ancha o larga para el tamaño de página definido. (*Nota:* Si una tabla es demasiado larga para que quepa en el resto de la página actual porque hay otros resultados en la página por encima de ella pero cabe en la longitud de página definida, la tabla se imprimirá automáticamente en una nueva página, independientemente del valor de líneas viudas/huérfanas establecido).
- Incluir el texto de continuación para las tablas que no quepan en una sola página. Este texto puede mostrarse en la parte inferior o superior de cada página. Si no selecciona ninguna de estas opciones, no aparecerá el texto de continuación.

Para controlar la impresión de las tablas pivote

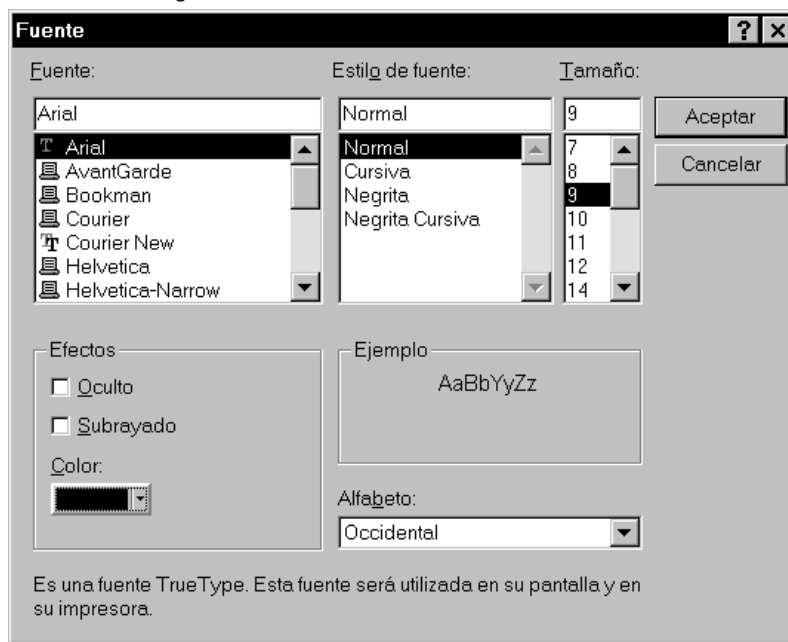
- ▶ Pulse en la pestaña Impresión.
- ▶ Seleccione las opciones de impresión que desee.
- ▶ Pulse en Aceptar o Aplicar.

Fuente

Un aspecto de tabla permite especificar las características de fuente de las distintas áreas de la tabla. Puede cambiar asimismo la fuente de una casilla individual. Las opciones de fuente de una casilla incluyen el aspecto, el estilo, el tamaño y el color de la fuente. También puede ocultar el texto o subrayarlo.

Si especifica las propiedades de fuente de una casilla, éstas se aplicarán a todas las capas de la tabla que tengan la misma casilla.

Figura 11-14
Cuadro de diálogo Fuente



Para cambiar la fuente de una casilla

- ▶ Active la tabla pivote y seleccione el texto que desea cambiar.
- ▶ Elija en los menús de la tabla pivote:
 - Formato
 - Fuente...

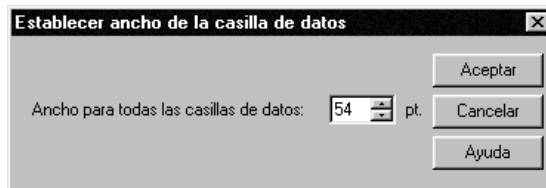
Si lo desea, puede seleccionar una fuente, un estilo de fuente, un tamaño, un color y un estilo de alfabeto, así como determinar que el texto esté oculto o subrayado.

Anchos de casillas de datos

El cuadro de diálogo Establecer ancho de la casilla de datos se utiliza para determinar la misma anchura para todas las casillas de datos.

Figura 11-15

Cuadro de diálogo Establecer ancho de la casilla de datos



Para cambiar el ancho de las casillas de datos

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
 - Formato
 - Ancho de casillas de datos...
- ▶ Introduzca un valor para el ancho de las casillas.

Para cambiar el ancho de una columna de una tabla pivote

- ▶ Active la tabla pivote (pulsando dos veces en cualquier punto de la tabla).

- ▶ Mueva el puntero del ratón por las etiquetas de categoría hasta situarlo en el borde derecho de la columna que desee cambiar. El puntero cambiará a una flecha con dos puntas.
- ▶ Mantenga pulsado el botón del ratón mientras arrastra el borde hasta su nueva posición.

Figura 11-16

Cambio del ancho de una columna

Arastrar el borde de la columna

Recuento		Sexo		Total
		Hombre	Mujer	
Categoría laboral	Administrativo	157	206	363
	Seguridad	27		27
	Directivo	74	10	84
Total		258	216	474

También puede cambiar los bordes de la dimensión y de la categoría vertical en el área de las etiquetas de fila aunque estén ocultos.

- ▶ Mueva el puntero del ratón por las etiquetas de fila hasta que aparezca la flecha con dos puntas.
- ▶ Arrastre hasta la nueva posición con el ancho deseado.

Propiedades de casilla

Las opciones de Propiedades de casilla sólo se aplican a una casilla individual seleccionada. Puede cambiar el formato del valor, la alineación, los márgenes y el sombreado. Las propiedades de casilla anulan las propiedades de tabla; por tanto, si modifica estas últimas, no cambiarán las propiedades de casilla aplicadas individualmente.

Para modificar las propiedades de una casilla

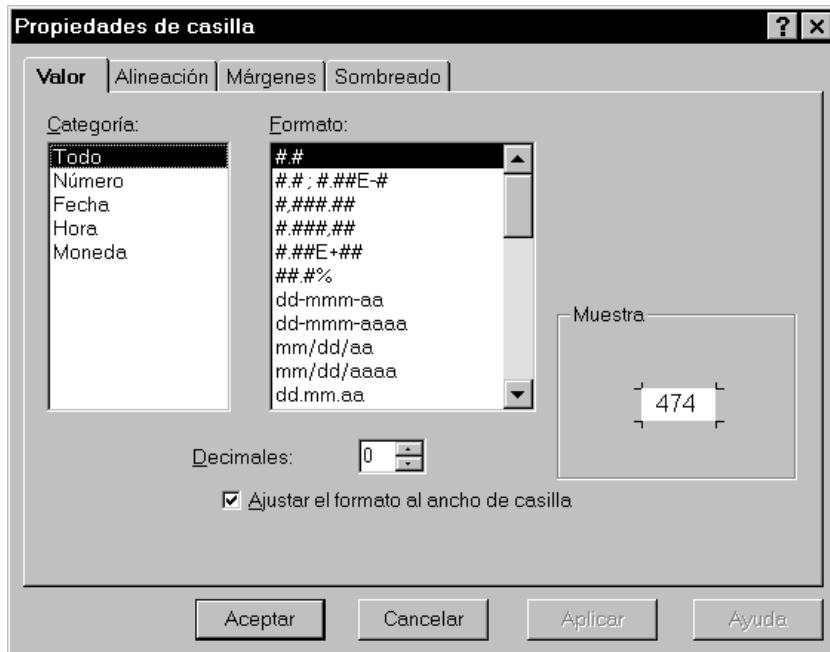
- ▶ Active una tabla y seleccione una de sus casillas.

- ▶ Elija en los menús:
 - Formato
 - Propiedades de casilla...

Propiedades de casilla: Valor

Esta pestaña del cuadro de diálogo controla el formato del valor de una casilla. Puede seleccionar formatos para números, fechas, horas o monedas, así como ajustar el número de dígitos decimales mostrados.

Figura 11-17
Pestaña Valor del cuadro de diálogo Propiedades de casilla



Para cambiar los formatos del valor de una casilla

- ▶ Pulse en la pestaña Valor.
- ▶ Seleccione una categoría y un formato.

- ▶ Seleccione el número de decimales.

Para cambiar los formatos de los valores de una columna

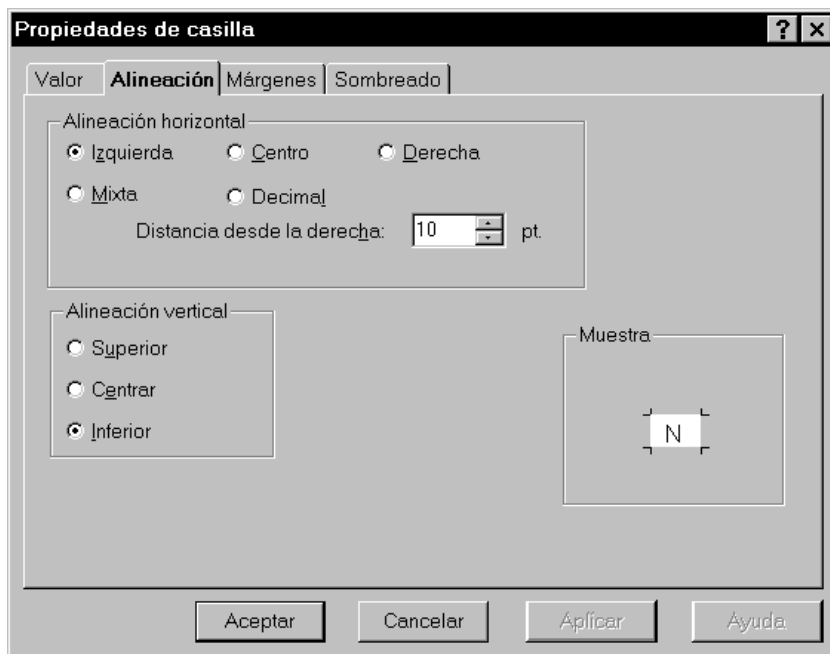
- ▶ Pulse las teclas Ctrl+Alt y en la etiqueta de la columna.
- ▶ Pulse con el botón derecho del ratón en la columna resaltada.
- ▶ En el menú contextual, seleccione Propiedades de casilla.
- ▶ Pulse en la pestaña Valor.
- ▶ Seleccione el formato que desea aplicar a la columna.

Puede utilizar este método para suprimir o añadir signos de porcentaje y de dólar, cambiar el número de decimales mostrados y alternar entre la presentación numérica regular y la notación científica.

Propiedades de casilla: Alineación

Esta pestaña del cuadro de diálogo permite establecer la alineación horizontal y vertical y la dirección del texto de una casilla. Si selecciona Mixta, el contenido de la casilla se alinearán según el tipo de que se trate (número, fecha o texto).

Figura 11-18
Pestaña *Alineación* del cuadro de diálogo *Propiedades de casilla*



Para cambiar la alineación de las casillas

- ▶ Seleccione una casilla de la tabla.
- ▶ Elija en los menús de la tabla pivote:
 - Formato
 - Propiedades de casilla...
- ▶ Pulse en la pestaña *Alineación*.

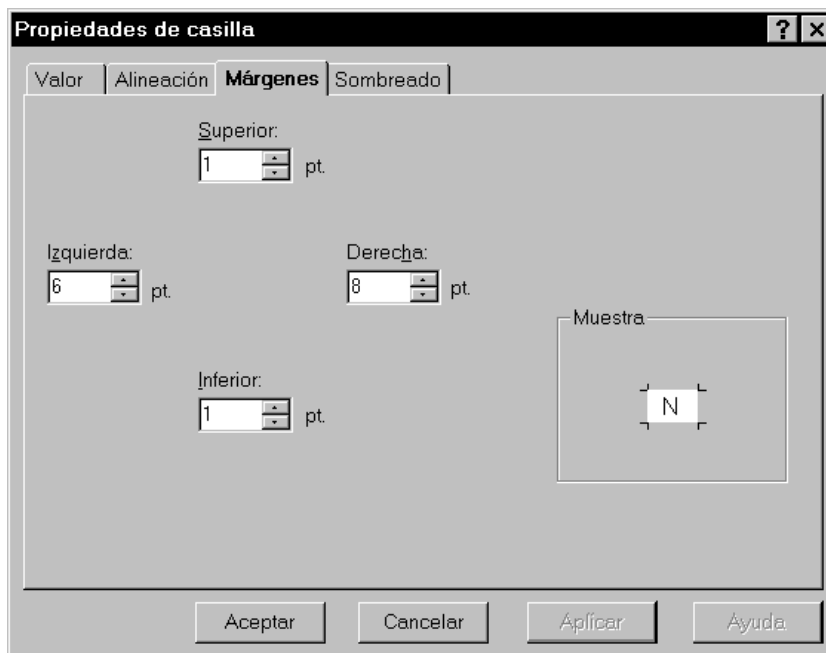
A medida que seleccione las propiedades de alineación para la casilla, éstas se irán reflejando en el área *Muestra*.

Propiedades de casilla: Márgenes

Esta pestaña del cuadro de diálogo permite especificar el margen en cada lateral de una casilla.

Figura 11-19

Pestaña Márgenes del cuadro de diálogo Propiedades de casilla



Para cambiar los márgenes de las casillas

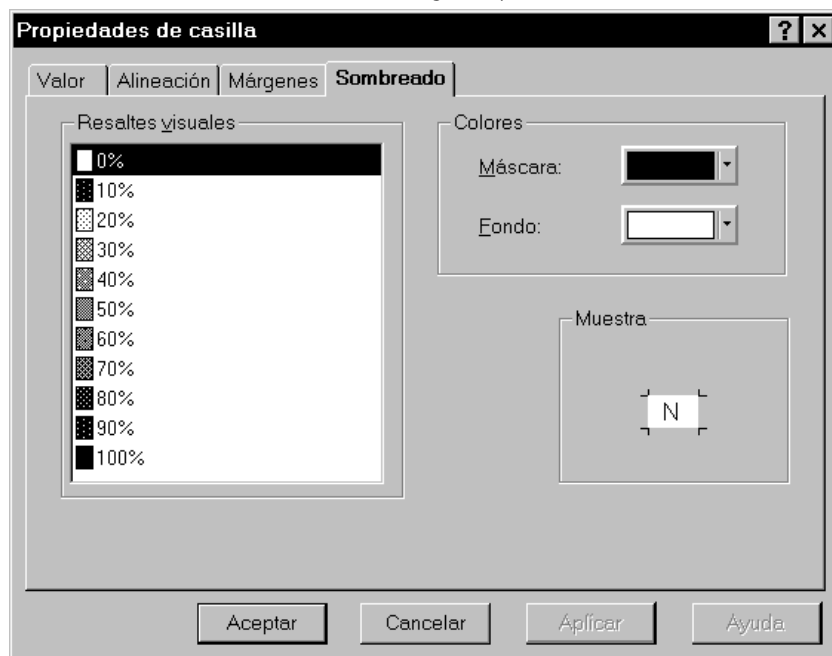
- ▶ Pulse en la pestaña Márgenes.
- ▶ Seleccione el valor para cada uno de los cuatro márgenes.

Propiedades de casilla: Sombreado

Esta pestaña del cuadro de diálogo permite especificar el porcentaje de sombreado, así como los colores de fondo y de máscara para una área de casillas seleccionada. Esto no cambia el color del texto.

Figura 11-20

Pestaña Sombreado del cuadro de diálogo *Propiedades de casilla*



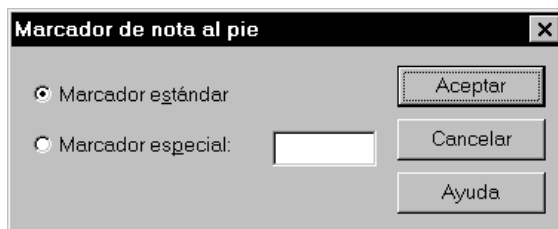
Para cambiar el sombreado de las casillas

- ▶ Pulse en la pestaña Sombreado.
- ▶ Seleccione los resaltos y los colores que desee aplicar a la casilla.

Marcador de nota al pie

Marcador de nota al pie cambia el carácter o los caracteres utilizados para marcar una nota al pie.

Figura 11-21
Cuadro de diálogo Marcador de nota al pie



Para cambiar los caracteres del marcador de notas al pie

- ▶ Seleccione una nota al pie.
- ▶ Elija en los menús de la tabla pivote:
 - Formato
 - Marcador de notas al pie...
- ▶ Escriba uno o dos caracteres.

Para volver a numerar las notas al pie

Cuando se ha pivotado una tabla intercambiando las filas, las columnas y las capas, las notas al pie pueden quedar desordenadas. Para volver a numerarlas:

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
 - Formato
 - Renumerar notas al pie

Selección de filas y columnas en las tablas pivotote

La flexibilidad de las tablas pivotote tiene algunas restricciones en el modo de seleccionar filas y columnas enteras, y el resalte visual que indica la fila o la columna seleccionada puede abarcar áreas no contiguas de la tabla.

Para seleccionar una fila o una columna de una tabla pivotote

- ▶ Active la tabla pivotote (pulsando dos veces en cualquier punto de la tabla).
- ▶ Pulse en una etiqueta de fila o de columna.
- ▶ Elija en los menús:
 - Edición
 - Seleccione
 - Casillas de datos y etiquetas
- o*
- ▶ Pulse las teclas Ctrl+Alt y en una etiqueta de fila o de columna.

Si la tabla contiene más de una dimensión en el área de filas o de columnas, la selección resaltada puede abarcar varias casillas no contiguas.

Modificación de los resultados de las tablas pivotote

Muchos de los elementos que aparecen en el Visor contienen texto. Si lo desea, puede editar el texto o añadir nuevo texto.

Las tablas pivotote se pueden modificar:

- Editando el texto que hay dentro de sus casillas
- Añadiendo texto y notas al pie

Para modificar el texto de una casilla

- ▶ Active la tabla pivotote.

- ▶ Pulse dos veces en la casilla.
- ▶ Edite el texto.
- ▶ Pulse en Intro para registrar los cambios o en Esc para restablecer el contenido previo de la casilla.

Para añadir textos al pie a una tabla

- ▶ Elija en los menús de la tabla pivote:
Insertar
 Texto al pie

Aparecerán las palabras Texto al pie de tabla en la parte inferior de la tabla.
- ▶ Seleccione las palabras Texto al pie de tabla e introduzca el texto al pie.

Para añadir una nota al pie a una tabla

Se puede anexar una nota al pie a cualquier elemento de la tabla.

- ▶ Pulse en un título, casilla o texto al pie dentro de una tabla pivote activada.
- ▶ Elija en los menús de la tabla pivote:
Insertar
 Nota al pie...
- ▶ Seleccione las palabras Nota al pie e introduzca el texto de la nota al pie.

Impresión de tablas pivote

Existen muchos factores que pueden afectar al aspecto que presentan los gráficos pivote impresos. Estos factores pueden controlarse cambiando los atributos de las tablas pivote.

- En las tablas pivote multidimensionales (tablas con capas), puede imprimir todas las capas o sólo la capa superior (la visible).
- En las tablas pivote largas o anchas, puede cambiar automáticamente el tamaño de la tabla para que quepa en la página o controlar la posición de las rupturas de tabla y de los saltos de página.

Utilice Presentación preliminar, en el menú Archivo, para ver cómo aparecerán las tablas pivote en la página impresa.

Para imprimir las capas ocultas de una tabla pivote

- ▶ Active la tabla pivote (pulsando dos veces en cualquier punto de la tabla).
- ▶ Elija en los menús:
 - Formato
 - Propiedades de tabla...
- ▶ En la pestaña Impresión, seleccione Imprimir todas las capas.

También puede imprimir cada capa de una tabla pivote en una hoja diferente.

Control de las rupturas de tabla para tablas anchas y largas

Las tablas pivote que son o demasiado anchas o largas para imprimir las en el tamaño de página definido se dividen automáticamente y se imprimen en varias secciones. (En las tablas anchas, se imprimen varias secciones en la misma página si hay suficiente espacio.) Tiene la posibilidad de:

- Controlar la posición de las filas y las columnas por donde se dividen las tablas de gran tamaño.
- Especificar las filas y las columnas que deben permanecer unidas cuando se dividen las tablas.
- Ajustar las tablas de gran tamaño para que quepan en el tamaño de página definido.

Para especificar las rupturas de filas y columnas para las tablas pivote impresas

- ▶ Active la tabla pivote.
- ▶ Pulse en la etiqueta de la columna situada a la izquierda del lugar donde desee insertar la ruptura, o en la etiqueta de la fila sobre la que desee insertarla.
- ▶ Elija en los menús:
 - Formato
 - Romper aquí

Para especificar las filas o las columnas que se han de mantener unidas

- ▶ Active la tabla pivote.
- ▶ Seleccione las etiquetas de las filas o columnas que desee mantener unidas (Pulse y arrastre, o bien mantenga pulsada la tecla Mayús para seleccionar varias etiquetas de fila o columna.)
- ▶ Elija en los menús:
 - Formato
 - Mantener juntos

Para ajustar la escala de una tabla pivote para que quepa en el tamaño de la página

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
 - Formato
 - Propiedades de tabla
- ▶ Pulse en la pestaña Impresión.

- ▶ Pulse en Reducir tabla ancha para caber en la página.

y/o

- ▶ Pulse en Reducir tabla larga para caber en la página.

Trabajar con sintaxis de comandos

SPSS proporciona un lenguaje de comandos eficaz que permite guardar y automatizar muchas tareas habituales. El lenguaje de comandos también proporciona algunas funcionalidades no incluidas en los menús y cuadros de diálogo.

Puede acceder a la mayoría de los comandos desde los menús y cuadros de diálogo. No obstante, algunos comandos y opciones sólo están disponibles mediante el uso del lenguaje de comandos. El lenguaje de comandos también permite guardar los trabajos en un archivo de sintaxis, con lo que podrá repetir los análisis en otro momento o ejecutarlos en un trabajo automatizado con la Unidad de producción.

Un archivo de sintaxis es simplemente un archivo de texto que contiene comandos. Aunque es posible abrir una ventana de sintaxis y escribir comandos, suele ser más sencillo permitir que el programa le ayude a construir un archivo de sintaxis mediante uno de los siguientes métodos:

- Pegando la sintaxis de comandos desde los cuadros de diálogo
- Copiando la sintaxis desde las anotaciones de los resultados
- Copiando la sintaxis desde el archivo diario

En la ayuda en pantalla de un procedimiento determinado, pulse en el vínculo de sintaxis de comandos que aparece en la lista de Temas relacionados para acceder a un diagrama de sintaxis correspondiente al comando relevante. Si desea información detallada sobre el lenguaje de comandos, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

La documentación completa sobre la sintaxis de comandos se instala automáticamente al mismo tiempo que SPSS. Para acceder a la documentación sobre la sintaxis:

- ▶ Elija en los menús:
 - ?
 - Referencia de sintaxis de comandos

Reglas de la sintaxis

Recuerde las siguientes sencillas reglas al editar y escribir la sintaxis de los comandos:

- Cada comando debe empezar en una línea nueva y terminar con un punto (.).
- La mayoría de los subcomandos están separados por barras inclinadas (/). La barra inclinada que precede al primer subcomando de un comando, generalmente es opcional.
- Los nombres de variable deben escribirse completos.
- El texto incluido entre apóstrofes o comillas debe ir contenido en una sola línea.
- Cada línea de la sintaxis de comando no puede exceder de 80 caracteres.
- Debe utilizarse un punto (.) para indicar decimales, independientemente de la configuración regional de Windows.
- Los nombres de variable que terminen en un punto pueden causar errores en los comandos creados por los cuadros de diálogo. No es posible crear nombres de variable de este tipo en los cuadros de diálogo y en general deben evitarse.

La sintaxis de comandos no distingue las mayúsculas de las minúsculas y permite el uso de abreviaturas de tres o cuatro letras en la mayoría de las especificaciones de los comandos. Puede usar tantas líneas como desee para especificar un único comando. Puede añadir espacios o líneas de separación en casi cualquier punto donde se permita un único espacio en blanco, como alrededor de las barras inclinadas, los paréntesis, los operadores aritméticos o entre los nombres de variable. Por ejemplo:

```
FRECUENCIAS  
VARIABLES=CATLAB SEXO  
/PERCENTILES=25 50 75  
/BARCHART.
```

y

```
freq var=catlab sexo /percent=25 50 75 /bar.
```

son alternativas aceptables que generan los mismos resultados.

Archivos INCLUDE

Para los archivos de comandos ejecutados mediante el comando `INCLUDE`, las reglas de sintaxis son un poco diferentes:

- Cada comando debe empezar en la primera columna de una línea nueva.
- Las líneas de continuación deben estar sangradas al menos un espacio.
- El punto del final del comando es opcional.

A menos que tenga archivos de comandos que ya utilizan el comando `INCLUDE`, debe utilizar el comando `INSERT` en su lugar dado que puede adaptar los archivos de comandos que se ajustan a los dos conjuntos de reglas. Si genera la sintaxis de comandos pegando las selecciones del cuadro de diálogo en una ventana de sintaxis, el formato de los comandos es apto para cualquier modo de operación. Consulte la *referencia de sintaxis de comandos* (disponible en formato PDF en el menú Ayuda) si desea obtener más información.

Pegar sintaxis desde cuadros de diálogo

La manera más fácil de construir un archivo de sintaxis de comandos es realizando las selecciones en los cuadros de diálogo y pegar la sintaxis de las selecciones en una ventana de sintaxis. Si pega la sintaxis en cada paso de un análisis largo, podrá generar un archivo de trabajo que le permitirá repetir el análisis con posterioridad o ejecutar un trabajo automatizado con la Unidad de producción.

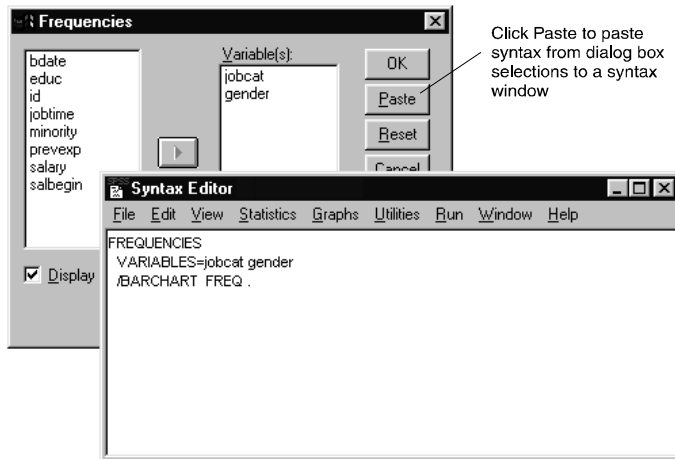
En la ventana de sintaxis, puede ejecutar la sintaxis pegada, editarla y guardarla en un archivo de sintaxis.

Para pegar sintaxis desde cuadros de diálogo

- ▶ Abra el cuadro de diálogo y realice las selecciones que desee.
- ▶ Pulse en Pegar.

La sintaxis de comandos se pegará en la ventana de sintaxis designada. Si no tiene abierta una ventana de sintaxis, se abrirá automáticamente una nueva y se pegará la sintaxis en ella.

Figura 12-1
Sintaxis de comandos pegada desde un cuadro de diálogo



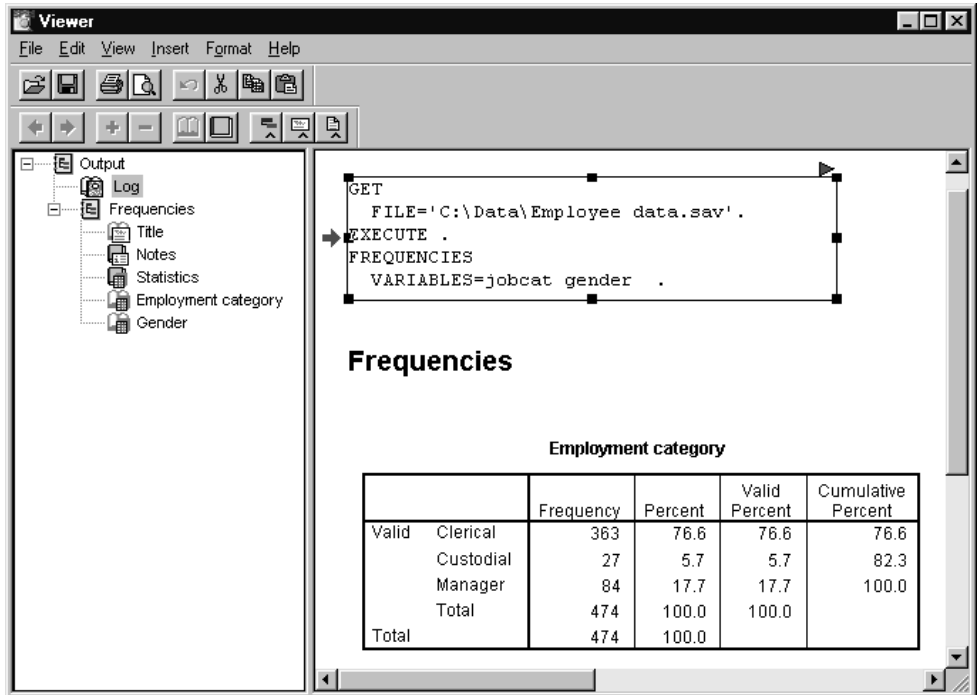
Nota: Si abre un cuadro de diálogo a partir de los menús de una ventana de proceso, el código para ejecutar la sintaxis desde un proceso se pegará en la ventana de proceso.

Copia de la sintaxis desde las anotaciones de los resultados

Puede construir un archivo de sintaxis copiando la sintaxis de comandos de la anotación que aparece en el Visor. Para usar este método debe seleccionar Mostrar comandos en anotaciones en la configuración del Visor (menú Edición, Opciones, pestaña Visor) antes de ejecutar el análisis. Todos los comandos aparecerán en el Visor junto con los resultados del análisis.

En la ventana de sintaxis, puede ejecutar la sintaxis pegada, editarla y guardarla en un archivo de sintaxis.

Figura 12-2
Sintaxis de comando en la anotación



Para copiar la sintaxis desde las anotaciones de los resultados

- ▶ Antes de ejecutar el análisis, elija en los menús:
Edición
Opciones...

- ▶ En la pestaña Visor, seleccione Mostrar comandos en anotaciones.
Mientras ejecuta los análisis, los comandos de las selecciones del cuadro de diálogo se graban en la anotación.
- ▶ Abra un archivo de sintaxis previamente guardado o cree uno nuevo. Para crear un archivo de sintaxis, elija en los menús:
Archivo
 Nuevo
 Sintaxis
- ▶ En el Visor, pulse dos veces en un elemento de anotación para activarlo.
- ▶ Pulse y arrastre el ratón para resaltar la sintaxis que desee copiar.
- ▶ Seleccione en los menús del Visor:
Edición
 Copiar
- ▶ En una ventana de sintaxis, elija en los menús:
Edición
 Pegar

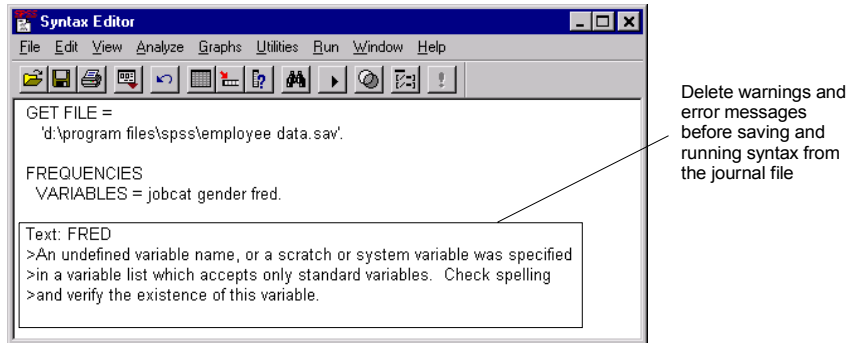
Edición de la sintaxis en un archivo diario

Por defecto, todos los comandos que se han ejecutado durante una sesión se graban en un archivo de diario denominado *spss.jnl* (definido en Opciones en el menú Edición). Podrá editar el archivo de diario y guardarlo como archivo de sintaxis, que podrá utilizar para repetir un análisis ejecutado anteriormente, o bien ejecutarlo en un trabajo automatizado con la Unidad de producción.

El archivo diario es un archivo de texto que puede editarse como cualquier otro archivo de texto. Dado que los mensajes de error y las advertencias también se registran en el archivo diario, junto con la sintaxis de comandos, deberá editar y eliminar los mensajes de error y de advertencia que aparezcan antes de guardar el archivo de sintaxis. No obstante, tenga en cuenta que los errores se deben solucionar o el trabajo no se podrá ejecutar correctamente.

Guarde el archivo diario editado con un nombre diferente. Puesto que el archivo de diario se amplía o sobrescribe de forma automática en cada sesión, utilizar el mismo nombre para un archivo de sintaxis y para un archivo de diario puede generar resultados inesperados.

Figura 12-3
Edición del archivo de diario



Para editar la sintaxis en un archivo diario

- ▶ Para abrir el archivo de diario, elija en los menús:

Archivo
Abrir
Otro...

- ▶ Localice y abra el archivo de diario (por defecto es *spss.jnl*, situado en el directorio *temp*).

En la lista desplegable Archivo de tipo seleccione Todos los archivos (*.*) o escriba *.jnl en el cuadro de texto Nombre de archivo para mostrar los archivos de diario en la lista. Si tiene dificultades para localizar el archivo, consulte las Opciones en el menú Edición para ver dónde se guarda el diario en su sistema.

- ▶ Edite el archivo para eliminar los mensajes de error o de advertencia que contenga, indicados por el signo >.
- ▶ Guarde el archivo diario editado con un nombre diferente. (Se recomienda usar un nombre de archivo con la extensión *.sps*, la extensión por defecto de los archivos de sintaxis.)

Para ejecutar la sintaxis de comandos

- ▶ Resalte los comandos que desee ejecutar en la ventana de sintaxis.
 - ▶ Pulse en el botón Ejecutar (el triángulo que apunta hacia la derecha) en la barra de herramientas del Editor de sintaxis.
- o*
- ▶ Seleccione una de las opciones del menú Ejecutar.
 - **Todo.** Ejecuta todos los comandos de la ventana de sintaxis.
 - **Selección.** Ejecuta los comandos seleccionados. Esto incluye los comandos parcialmente resaltados.
 - **Actual.** Ejecuta el comando donde se encuentra el cursor.
 - **Hasta el final.** Ejecuta todos los comandos incluidos desde la posición actual del cursor hasta el final del archivo de sintaxis de comandos.

El botón Ejecutar de la barra de herramientas del Editor de sintaxis ejecuta los comandos seleccionados o el comando donde se encuentra el cursor si no hay nada seleccionado.

Figura 12-4
Barra de herramientas del Editor de sintaxis.



El botón Ejecutar ejecuta los comandos seleccionados en la ubicación del cursor

Varios comandos Ejecutar

La sintaxis pegada desde cuadros de diálogo o copiada desde el registro o el diario puede contener comandos EXECUTE. Al ejecutar varios comandos desde una ventana de sintaxis, no es necesario que haya varios comandos EXECUTE y es posible que afecte negativamente al rendimiento ya que este comando lee todo el archivo de datos.

- Si el último comando del archivo de sintaxis es un comando que lee el archivo de datos (como un procedimiento estadístico o gráfico), no es necesario ningún comando EXECUTE y se pueden eliminar.
- Si no está seguro de si el último comando lee el archivo de datos, en la mayoría de los casos puede eliminarlos todos menos el último comando EXECUTE del archivo de sintaxis.

Funciones de retardo

Una excepción importante son los comandos de transformación que contienen funciones de retardo. En una serie de comandos de transformación sin intervención de comandos EXECUTE, ni ningún otro comando que lea datos, las funciones de retardo se calculan después de las restantes transformaciones, con independencia del orden de los comandos. Por ejemplo:

```
COMPUTE lagvar=LAG(var1)
COMPUTE var1=var1*2
```

y

```
COMPUTE lagvar=LAG(var1)
EXECUTE
COMPUTE var1=var1*2
```

ofrece resultados muy diferentes para el valor de *lagvar* dado que el anterior utiliza el valor transformado de *var1* mientras que el último utiliza el valor original.

Frecuencias

El procedimiento Frecuencias proporciona estadísticos y representaciones gráficas que resultan útiles para describir muchos tipos de variables. Es un buen procedimiento para una inspección inicial de los datos.

Para los informes de frecuencias y los gráficos de barras, puede organizar los diferentes valores en orden ascendente o descendente u ordenar las categorías por sus frecuencias. Es posible suprimir el informe de frecuencias cuando una variable posee muchos valores diferentes. Puede etiquetar los gráficos con las frecuencias (la opción por defecto) o con los porcentajes.

Ejemplo. ¿Cuál es la distribución de los clientes de una empresa por tipo de industria? En los resultados podría observar que el 37,5% de sus clientes pertenece a agencias gubernamentales, el 24,9% a corporaciones, el 28,1% a instituciones académicas, y el 9,4% a la industria sanitaria. Con respecto a los datos continuos, cuantitativos, como los ingresos por ventas, podría comprobar que el promedio de ventas de productos es de 3.576 dólares con una desviación típica de 1.078 dólares.

Estadísticos y gráficos. Frecuencias, porcentajes, porcentajes acumulados, media, mediana, moda, suma, desviación típica, varianza, amplitud, valores mínimo y máximo, error típico de la media, asimetría y curtosis (ambos con sus errores típicos), cuartiles, percentiles especificados por el usuario, gráficos de barras, gráficos de sectores e histogramas.

Datos. Utilice códigos numéricos o cadenas cortas para codificar las variables categóricas (medidas de nivel nominal u ordinal).

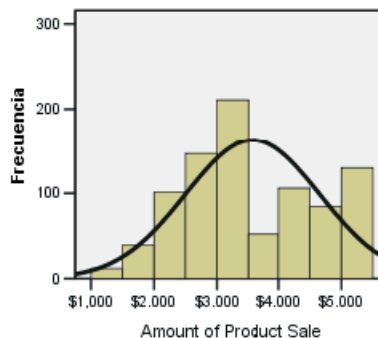
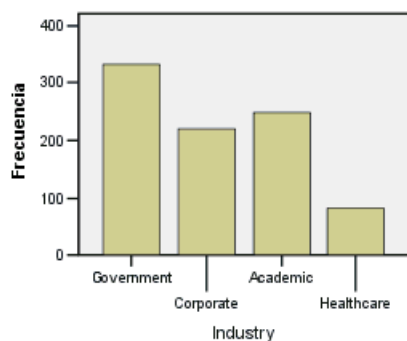
Supuestos. Las tabulaciones y los porcentajes proporcionan una descripción útil para los datos de cualquier distribución, especialmente para las variables con categorías ordenadas o desordenadas. Muchos de los estadísticos de resumen optativos, tales como la media y la desviación típica, se basan en la teoría normal y son apropiados para las variables cuantitativas con distribuciones simétricas. Los estadísticos

robustos, tales como la mediana, los cuartiles y los percentiles son apropiados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad.

Figura 13-1
Resultados de Frecuencias

		Industry			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Government	331	37.5	37,5	37,5
	Corporate	220	24.9	24,9	62,5
	Academic	248	28.1	28,1	90,6
	Healthcare	83	9.4	9,4	100,0
	Total	882	100.0	100,0	

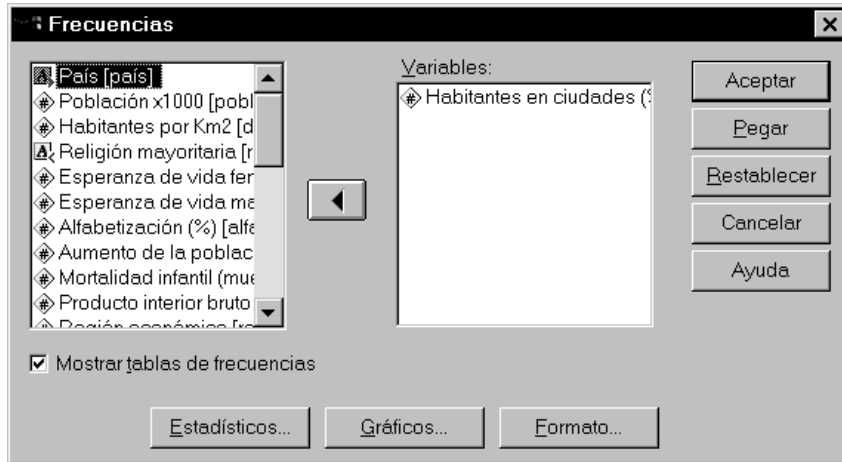
Estadísticos			
	Media	Mediana	Desv. típ.
Amount of Product Sale	\$3.576,52	\$3.417,50	\$1.077,836



Para obtener tablas de frecuencias

- ▶ Elija en los menús:
Analizar
Estadísticos descriptivos
Frecuencias...

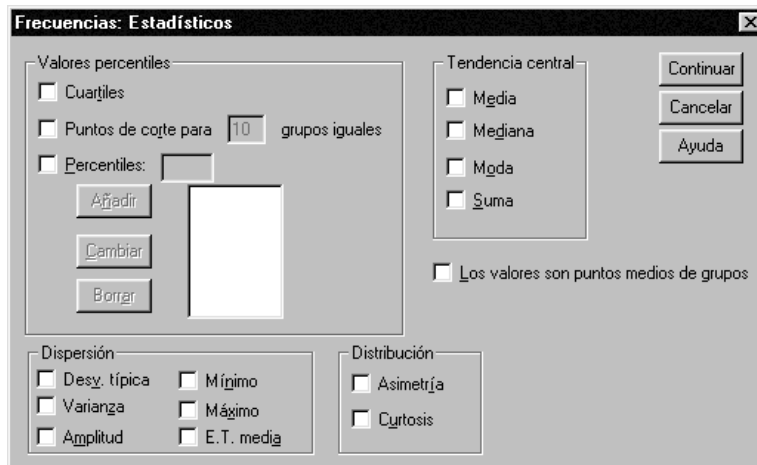
Figura 13-2
Cuadro de diálogo Frecuencias



- ▶ Seleccione una o más variables categóricas o cuantitativas.
Si lo desea, tiene la posibilidad de:
 - Pulsar en Estadísticos para obtener estadísticos descriptivos para las variables cuantitativas.
 - Pulsar en Gráficos para obtener gráficos de barras, gráficos de sectores e histogramas.
 - Pulsar en Formato para determinar el orden en el que se muestran los resultados.

Frecuencias: Estadísticos

Figura 13-3
Cuadro de diálogo Frecuencias: Estadísticos



Valores percentiles. Los valores de una variable cuantitativa que dividen los datos ordenados en grupos, de forma que un porcentaje de los casos se encuentre por encima y otro porcentaje se encuentre por debajo. Los cuartiles (los percentiles 25, 50 y 75) dividen las observaciones en cuatro grupos de igual tamaño. Si desea un número igual de grupos que no sea cuatro, seleccione Puntos de corte para n grupos iguales. También puede especificar percentiles individuales (por ejemplo, el percentil 95, el valor por debajo del cual se encuentran el 95% de las observaciones).

Tendencia central. Los estadísticos que describen la localización de la distribución, incluyen: Media, Mediana, Moda y Suma de todos los valores.

- **Media.** Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.
- **Mediana.** Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Cuando el número de observaciones es par, la mediana es el promedio de las dos observaciones centrales, una vez que han sido ordenadas de manera ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

- **Moda.** El valor que ocurre con mayor frecuencia. Si varios valores comparten la mayor frecuencia de aparición, cada una de ellas es una moda. El procedimiento de frecuencias devuelve sólo la más pequeña de esas modas múltiples.
- **Suma.** La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

Dispersión. Los estadísticos que miden la cantidad de variación o de dispersión en los datos, incluyen: Desviación típica, Varianza, Rango, Mínimo, Máximo y Error típico de la media.

- **Desviación típica.** Medida de dispersión en torno a la media. En una distribución normal, el 68% de los casos se encuentra dentro de una DT respecto a la media y el 95% de los casos se encuentra dentro de 2 DT respecto a la media. Por ejemplo, si la media de edad es 45, con una desviación típica de 10, el 95% de los casos estaría entre 25 y 65 en una distribución normal.
- **Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.
- **Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.
- **Mínimo.** Valor más pequeño de una variable numérica.
- **Máximo.** El mayor valor de una variable numérica.
- **E. T. media.** Medida de cuánto puede variar el valor de la media de una muestra a otra, extraídas éstas de la misma distribución. Puede utilizarse para comparar de forma aproximada la media observada con un valor hipotetizado (es decir, podremos concluir que dos valores son distintos si la razón de la diferencia respecto al error típico es menor que -2 o mayor que +2).

Distribución. Asimetría y curtosis son estadísticos que describen la forma y la simetría de la distribución. Estos estadísticos se muestran con sus errores típicos.

- **Asimetría.** Medida de la asimetría de una distribución. La distribución normal es simétrica por lo que tiene un valor de asimetría 0. Una distribución que tenga una asimetría positiva significativa tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Un valor

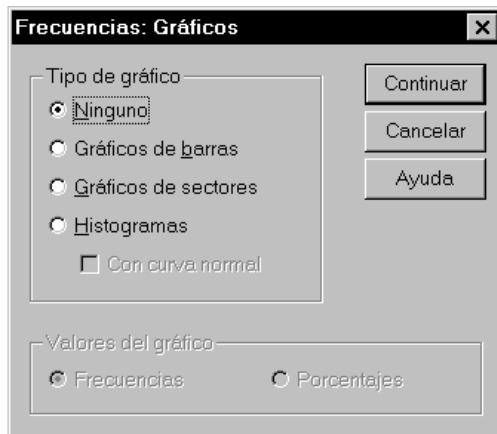
de asimetría mayor que 1, en valor absoluto, indica generalmente una distribución que difiere de manera significativa de la distribución normal.

- **Curtosis.** Medida del grado en que las observaciones están agrupadas en torno al punto central. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

Los valores son puntos medios de grupos. Si los valores de los datos son puntos medios de grupos (por ejemplo, si las edades de todas las personas entre treinta y cuarenta años se codifican como 35), seleccione esta opción para estimar la mediana y los percentiles para los datos originales no agrupados.

Frecuencias: Gráficos

Figura 13-4
Cuadro de diálogo Frecuencias: Gráficos



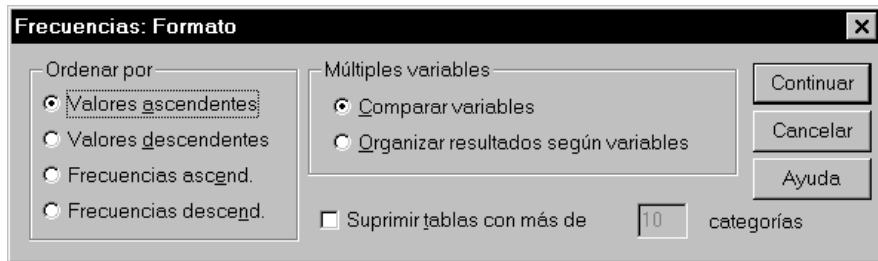
Tipo de gráfico. Los gráficos de sectores muestran la contribución de las partes a un todo. Cada sector de un gráfico de este tipo corresponde a un grupo, definido por una única variable de agrupación. Los gráficos de barras muestran la frecuencia de cada valor o categoría distinta como una barra diferente, permitiendo comparar las categorías de forma visual. Los histogramas también cuentan con barras, pero se representan a lo largo de una escala de intervalos iguales. La altura de cada barra es el

recuento de los valores que están dentro del intervalo para una variable cuantitativa. Los histogramas muestran la forma, el centro y la dispersión de la distribución. Una curva normal superpuesta en un histograma ayuda a juzgar si los datos están normalmente distribuidos.

Valores del gráfico. Para los gráficos de barras, puede etiquetar el eje de escala con las frecuencias o los porcentajes.

Frecuencias: Formato

Figura 13-5
Cuadro de diálogo Frecuencias: Formato



Ordenar por. La tabla de frecuencias se puede organizar respecto a los valores actuales de los datos o respecto al recuento (frecuencia de aparición) de esos valores, y en orden ascendente o descendente. Sin embargo, si solicita un histograma o percentiles, Frecuencias asumirá que la variable es cuantitativa y mostrará sus valores en orden ascendente.

Múltiples variables. Si desea generar tablas de estadísticos para múltiples variables, podrá mostrar todas las variables en una sola tabla (Comparar variables), o bien mostrar una tabla de estadísticos independiente para cada variable (Organizar resultados según variables).

Suprimir tablas con más de n categorías. Esta opción impide que se muestren tablas que contengan más valores que el número especificado.

Descriptivos

El procedimiento Descriptivos muestra estadísticos de resumen univariados para varias variables en una única tabla y calcula valores tipificados (puntuaciones z). Las variables se pueden ordenar por el tamaño de sus medias (en orden ascendente o descendente), alfabéticamente o por el orden en el que se seleccionen las variables (el valor por defecto).

Cuando se guardan las puntuaciones z , éstas se añaden a los datos del Editor de datos y quedan disponibles para los gráficos, el listado de los datos y los análisis. Cuando las variables se registran en unidades diferentes (por ejemplo, producto interior bruto per cápita y porcentaje de alfabetización), una transformación de puntuación z pondrá las variables en una escala común para poder compararlas visualmente con más facilidad.

Ejemplo. Si cada caso de los datos contiene los totales de ventas diarias de cada vendedor (por ejemplo, una entrada para Bob, una para Kim, una para Brian, etc.) recogidas cada día durante varios meses, el procedimiento Descriptivos puede calcular la media diaria de ventas para cada vendedor y ordenar los resultados del promedio de ventas de mayor a menor.

Estadísticos. Tamaño de muestra, media, mínimo, máximo, desviación típica, varianza, rango, suma, error típico de la media, curtosis y asimetría con sus errores típicos.

Datos. Utilice variables numéricas después de haberlas inspeccionado gráficamente para registrar errores, valores atípicos y anomalías de distribución. El procedimiento Descriptivos es muy eficaz para archivos grandes (de miles de casos).

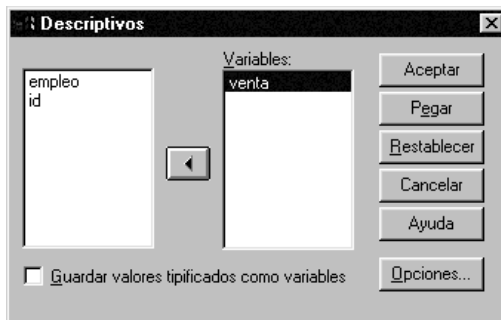
Supuestos. La mayoría de los estadísticos disponibles (incluyendo las puntuaciones z) se basan en la teoría normal y son adecuados para variables cuantitativas (medidas a nivel de razón o de intervalo) con distribuciones simétricas (se deben evitar variables con categorías no ordenadas o distribuciones asimétricas). La distribución

de puntuaciones z tiene la misma forma que la de los datos originales; por tanto, el cálculo de puntuaciones z no es una solución para los datos con problemas.

Para obtener estadísticos descriptivos

- ▶ Elija en los menús:
Analizar
 Estadísticos descriptivos
 Descriptivos...

Figura 14-1
Cuadro de diálogo Descriptivos



- ▶ Seleccione una o más variables.

Si lo desea, tiene la posibilidad de:

- Seleccionar Guardar valores tipificados como variables para guardar las puntuaciones z como nuevas variables.
- Pulsar en Opciones para seleccionar estadísticos opcionales y el orden de presentación.

Descriptivos: Opciones

Figura 14-2
Cuadro de diálogo Descriptivos: Opciones



Media y suma. Se muestra por defecto la media o promedio aritmético.

Dispersión. Los estadísticos que miden la dispersión o variación en los datos incluyen la desviación típica, la varianza, el rango, el mínimo, el máximo y el error típico de la media.

- **Desviación típica.** Medida de dispersión en torno a la media. En una distribución normal, el 68% de los casos se encuentra dentro de una DT respecto a la media y el 95% de los casos se encuentra dentro de 2 DT respecto a la media. Por ejemplo, si la media de edad es 45, con una desviación típica de 10, el 95% de los casos estaría entre 25 y 65 en una distribución normal.
- **Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.
- **Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.
- **Mínimo.** Valor más pequeño de una variable numérica.

- **Máximo.** El mayor valor de una variable numérica.
- **Error típico de la media.** Medida de cuánto puede variar el valor de la media de una muestra a otra, extraídas éstas de la misma distribución. Puede utilizarse para comparar de forma aproximada la media observada con un valor hipotetizado (es decir, podremos concluir que dos valores son distintos si la razón de la diferencia respecto al error típico es menor que -2 o mayor que $+2$).

Distribución. La curtosis y la asimetría son los estadísticos que caracterizan la forma y simetría de la distribución. Estos se muestran con sus errores típicos.

- **Curtosis.** Medida del grado en que las observaciones están agrupadas en torno al punto central. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.
- **Asimetría.** Medida de la asimetría de una distribución. La distribución normal es simétrica por lo que tiene un valor de asimetría 0. Una distribución que tenga una asimetría positiva significativa tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Un valor de asimetría mayor que 1, en valor absoluto, indica generalmente una distribución que difiere de manera significativa de la distribución normal.

Orden de presentación. Por defecto, las variables se muestran en el orden en que se hayan seleccionado. Si lo desea, se pueden mostrar las variables alfabéticamente, por medias ascendentes o por medias descendentes.

Explorar

El procedimiento Explorar genera estadísticos de resumen y representaciones gráficas, bien para todos los casos o bien de forma separada para grupos de casos. Existen numerosas razones para utilizar este procedimiento: para inspeccionar los datos, identificar valores atípicos, obtener descripciones, comprobar supuestos y caracterizar diferencias entre subpoblaciones (grupos de casos). La inspección de los datos puede mostrar que existen valores inusuales, valores extremos, discontinuidades en los datos u otras peculiaridades. La exploración de los datos puede ayudar a determinar si son adecuadas las técnicas estadísticas que está teniendo en consideración para el análisis de los datos. La exploración puede indicar que necesita transformar los datos si la técnica necesita una distribución normal. O bien, el usuario puede decidir que necesita utilizar pruebas no paramétricas.

Ejemplo. Observe la distribución de los tiempos de aprendizaje de laberintos de una serie de ratas sometidas a cuatro programas de refuerzo diferentes. Para cada uno de los cuatro grupos, se puede observar si la distribución de tiempos es aproximadamente normal y si las cuatro varianzas son iguales. También se pueden identificar los casos con los cinco valores de tiempo mayores y los cinco menores. Los diagramas de caja y los gráficos de tallo y hojas resumen gráficamente la distribución del tiempo de aprendizaje de cada uno de los grupos.

Estadísticos y gráficos. Media, mediana, media recortada al 5%, error típico, varianza, desviación típica, mínimo, máximo, amplitud, amplitud intercuartil, asimetría y curtosis y sus errores típicos, intervalo de confianza para la media (y el nivel de confianza especificado), percentiles, estimador-M de Huber, estimador en onda de Andrews, estimador-M redescendente de Hampel, estimador bponderado de Tukey, cinco valores mayores y cinco menores, estadístico de Kolmogorov-Smirnov con el nivel de significación de Lilliefors para contrastar la normalidad y estadístico de Shapiro-Wilk. Diagramas de caja, gráficos de tallo y hojas, histogramas, diagramas

de normalidad y diagramas de dispersión por nivel con pruebas de Levene y transformaciones.

Datos. El procedimiento Explorar se puede utilizar para las variables cuantitativas (nivel de medida de razón o de intervalo). Una variable de factor (utilizada para dividir los datos en grupos de casos) debe tener un número razonable de valores distintivos (categorías). Estos valores pueden ser de cadena corta o numéricos. La variable de etiquetas de caso, utilizada para etiquetar valores atípicos en los diagramas de caja, puede ser de cadena corta, de cadena larga (los 15 primeros caracteres) o numérica.

Supuestos. La distribución de los datos no tiene que ser simétrica ni normal.

Figura 15-1
Resultados de Explorar

			Descriptivos			
			Tiempo			
			Horario			
			1	2	3	4
Media	Estadístico		2,760	4,850	6,900	9,010
	Error típ.		,165	,422	,445	,289
Intervalo de confianza para la media al 95%	Límite inferior	Estadístico	2,387	3,895	5,893	8,357
	Límite superior	Estadístico	3,133	5,805	7,907	9,663
Media recortada al 5%		Estadístico	2,761	4,889	6,911	8,994
Mediana		Estadístico	2,850	4,900	7,050	9,000
Varianza		Estadístico	,272	1,783	1,982	,834
Desv. típ.		Estadístico	,521	1,335	1,408	,913
Mínimo		Estadístico	2,0	2,3	4,5	7,8
Máximo		Estadístico	3,5	6,7	9,1	10,5
Rango		Estadístico	1,5	4,4	4,6	2,7
Amplitud intercuartil		Estadístico	,925	2,250	2,400	1,650
Asimetría		Estadístico	-,116	-,559	-,197	,219
		Error típ.	,687	,687	,687	,687
Curtosis		Estadístico	-1,210	-,104	-,606	-1,350
		Error típ.	1,334	1,334	1,334	1,334

Valores extremos

			Número de caso	Horario	Valor
Hora	Mayores	1	31	4	10,5
		2	33	4	9,9
		3	39	4	9,8
		4	32	4	9,5
		5	36	4	9,3
	Menores	1	2	1	2,0
		2	7	1	2,1
		3	1	1	2,3
		4	11	2	2,3
		5	3	1	2,5

Frecuencia Tallo y hojas

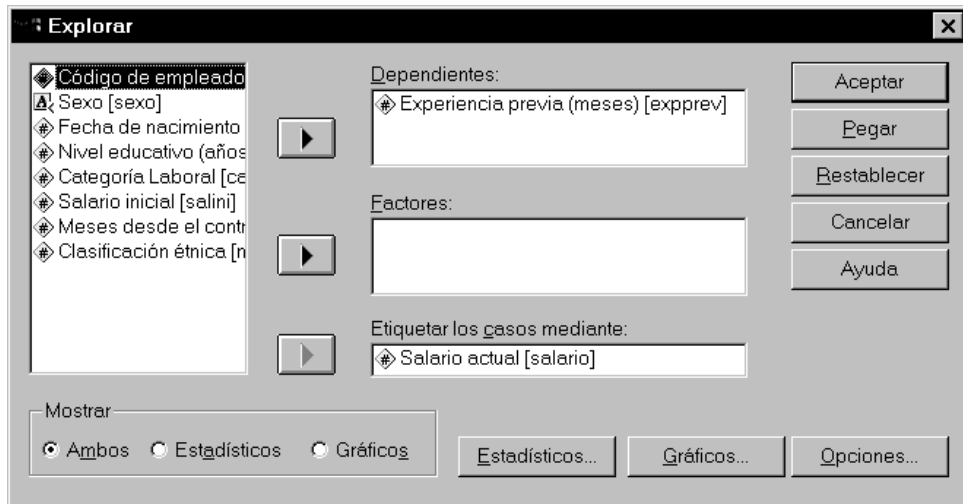
7,00 2 , 0133589
6,00 3 , 014577
3,00 4 , 568
5,00 5 , 05779
4,00 6 , 1379
3,00 7 , 268
6,00 8 , 012237
5,00 9 , 13589
1,00 10 , 5

Anchura del tallo: 1,0
Cada hoja: 1 caso(s)

Para explorar los datos

- Elija en los menús:
 - Analizar
 - Estadísticos descriptivos
 - Explorar...

Figura 15-2
Cuadro de diálogo Explorar



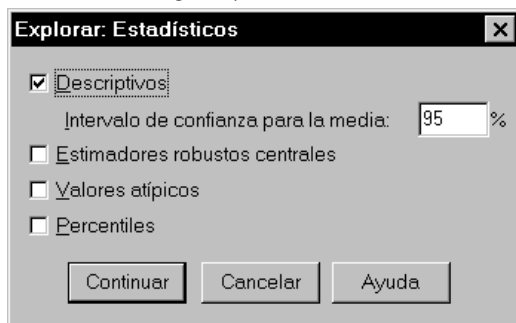
- ▶ Seleccione una o más variables dependientes.

Si lo desea, tiene la posibilidad de:

- Seleccionar una o más variables de factor, cuyos valores definirán grupos de casos.
- Seleccionar una variable de identificación para etiquetar los casos.
- Pulse en Estadísticos para obtener estimadores robustos, valores atípicos, percentiles y tablas de frecuencias.
- Pulse en Gráficos para obtener histogramas, pruebas y gráficos de probabilidad normal y diagramas de dispersión por nivel con estadísticos de Levene.
- Pulse en Opciones para manipular los valores perdidos.

Explorar: Estadísticos

Figura 15-3
Cuadro de diálogo Explorar: Estadísticos



Descriptivos. Por defecto se muestran estas medidas de dispersión y de tendencia central. Éstas últimas indican la localización de la distribución, e incluyen la media, la mediana y la media recortada al 5%. Las medidas de dispersión muestran la disimilaridad de los valores, incluyen: los errores típicos, la varianza, la desviación típica, el mínimo, el máximo, la amplitud y la amplitud intercuartil. Los estadísticos descriptivos también incluyen medidas de la forma de la distribución: la asimetría y la curtosis se muestran con sus errores típicos. También se muestra el intervalo de confianza a un nivel del 95%; aunque se puede especificar otro nivel.

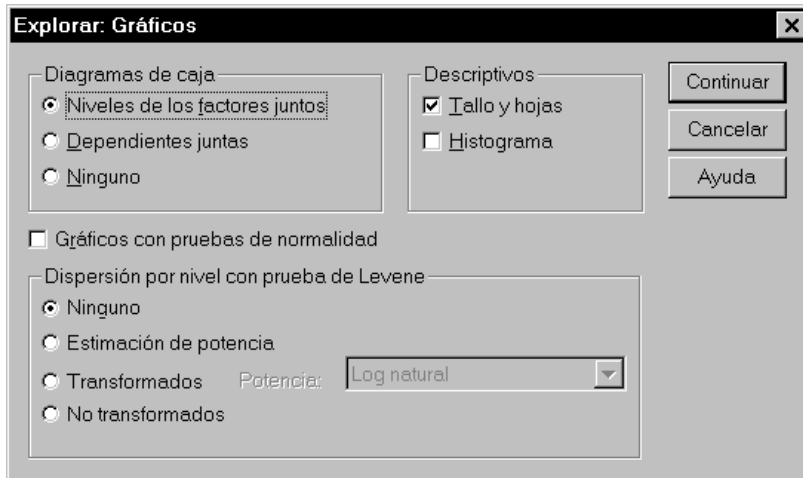
Estimadores robustos centrales. Alternativas robustas a la mediana y a la media muestral para estimar el centro de la localización. Los estimadores calculados se diferencian por las ponderaciones que aplican a los casos. Se muestran los siguientes: el estimador-M de Huber, el estimador en onda de Andrew, el estimador-M redescendente de Hampel y el estimador bponderado de Tukey.

Valores atípicos. Muestra los cinco valores mayores y los cinco menores, con las etiquetas de caso.

Percentiles. Muestra los valores de los percentiles 5, 10, 25, 50, 75, 90 y 95.

Explorar: Gráficos

Figura 15-4
Cuadro de diálogo Explorar: Gráficos



Diagramas de caja. Estas alternativas controlan la presentación de los diagramas de caja cuando existe más de una variable dependiente. Niveles de los factores juntos genera una presentación para cada variable dependiente. En cada una se muestran diagramas de caja para cada uno de los grupos definidos por una variable de factor. Dependientes juntas genera una presentación para cada grupo definido por una variable de factor. En cada una se muestran juntos los diagramas de caja de cada variable dependiente. Esta disposición es particularmente útil cuando las variables representan una misma característica medida en momentos distintos.

Descriptivos. La sección Descriptivos permite seleccionar gráficos de tallo y hojas e histogramas.

Gráficos con pruebas de normalidad. Muestra los diagramas de probabilidad normal y de probabilidad sin tendencia. Se muestra el estadístico de Kolmogorov-Smirnov con un nivel de significación de Lilliefors para contrastar la normalidad. Si se especifican ponderaciones no enteras, se calculará el estadístico de Shapiro-Wilk cuando el tamaño de la muestra ponderada esté entre 3 y 50. Si no hay ponderaciones o éstas son enteras, se calculará el estadístico cuando el tamaño muestral esté entre 3 y 5000.

Dispersión por nivel con prueba de Levene. Controla la transformación de los datos para los diagramas de dispersión por nivel. Para todos los diagramas de dispersión por nivel se muestra la pendiente de la línea de regresión y las pruebas robustas de Levene sobre la homogeneidad de varianza. Si selecciona una transformación, las pruebas de Levene se basarán en los datos transformados. Si no selecciona ninguna variable de factor, no se generará ningún diagrama de dispersión por nivel. Estimación de potencia produce un gráfico de los logaritmos naturales de las amplitudes intercuartiles respecto a los logaritmos naturales de las medianas de todas las casillas, así como una estimación de la transformación de potencia necesaria para conseguir varianzas iguales en las casillas. Un diagrama de dispersión por nivel ayuda a determinar la potencia que precisa una transformación para estabilizar (igualar) las varianzas de los grupos. Transformados permite seleccionar una de las alternativas de potencia, quizás siguiendo las recomendaciones de la estimación de potencia, y genera gráficos de los datos transformados. Se trazan la amplitud intercuartil y la mediana de los datos transformados. No transformados genera gráficos de los datos brutos. Es equivalente a una transformación con una potencia de 1.

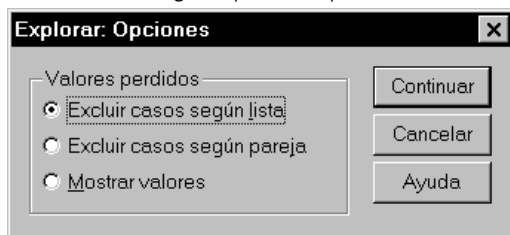
Explorar: Transformaciones de potencia

A continuación aparecen las transformaciones de potencia para los diagramas de dispersión por nivel. Para transformar los datos, deberá seleccionar una potencia para la transformación. Puede elegir una de las siguientes alternativas:

- **Log natural.** Transformación de logaritmo natural. Este es el método por defecto.
- **1/raíz cuadrada.** Para cada valor de los datos se calcula el inverso de la raíz cuadrada.
- **Recíproco.** Se calcula el inverso de cada valor de los datos.
- **Raíz cuadrada.** Se calcula la raíz cuadrada de cada valor de los datos.
- **Cuadrado.** Se calcula el cuadrado de cada valor de los datos.
- **Cubo.** Se calcula el cubo de cada valor de los datos.

Explorar: Opciones

Figura 15-5
Cuadro de diálogo Explorar: Opciones



Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable de factor o variable dependiente se excluyen de todos los análisis. Este es el método por defecto.
- **Excluir casos según pareja.** Los casos que no tengan valores perdidos para las variables de un grupo (casilla) se incluyen en el análisis de ese grupo. El caso puede tener valores perdidos para las variables utilizadas en otros grupos.
- **Mostrar los valores.** Los valores perdidos para las variables de factor se tratan como una categoría diferente. Todos los resultados se generan para esta categoría adicional. Las tablas de frecuencias incluyen categorías para los valores perdidos. Los valores perdidos para una variable de factor se incluyen pero se etiquetan como perdidos.

Tablas de contingencia

El procedimiento Tablas de contingencia crea tablas de clasificación doble y múltiple y, además, proporciona una serie de pruebas y medidas de asociación para las tablas de doble clasificación. La estructura de la tabla y el hecho de que las categorías estén ordenadas o no determinan las pruebas o medidas que se utilizaban.

Los estadísticos de tablas de contingencia y las medidas de asociación sólo se calculan para las tablas de doble clasificación. Si especifica una fila, una columna y un factor de capa (variable de control), el procedimiento Tablas de contingencia crea un panel de medidas y estadísticos asociados para cada valor del factor de capa (o una combinación de valores para dos o más variables de control). Por ejemplo, si *sexo* es un factor de capa para una tabla de *casado* (sí, no) en función de *vida* (vida emocionante, rutinaria o aburrida), los resultados para una tabla de doble clasificación para las mujeres se calculan de forma independiente de los resultados de los hombres y se imprimen en paneles uno detrás del otro.

Ejemplo. ¿Es más probable que los clientes de las empresas pequeñas sean más rentables en la venta de servicios (por ejemplo, formación y asesoramiento) que los clientes de las empresas grandes? A partir de una tabla de contingencia podría deducir que la prestación de servicios a la mayoría de las empresas pequeñas (con menos de 500 empleados) produce considerables beneficios, mientras que con la mayoría de las empresas de gran tamaño (con más de 2.500 empleados), los beneficios obtenidos son mucho menores.

Estadísticos y medidas de asociación. Chi-cuadrado de Pearson, chi-cuadrado de la razón de verosimilitud, prueba de asociación lineal por lineal, prueba exacta de Fisher, chi-cuadrado corregido de Yates, r de Pearson, rho de Spearman, coeficiente de contingencia, phi, V de Cram, lambdas simétricas y asimétricas, tau de Kruskal y Goodman, coeficiente de incertidumbre, gamma, d de Somers, tau- b de Kendall, tau- c de Kendall, coeficiente eta, kappa de Cohen, estimación de riesgo relativo, razón de ventajas, prueba de McNemar y estadísticos de Cochran y Mantel-Haenszel.

Datos. Para definir las categorías de cada variable, utilice valores de una variable numérica o de cadena corta (ocho caracteres o menos). Por ejemplo, para *sexo*, podría codificar los datos como 1 y 2 o como *varón* y *mujer*.

Supuestos. En algunos estadísticos y medidas se asume que hay unas categorías ordenadas (datos ordinales) o unos valores cuantitativos (datos de intervalos o de proporciones), como se explica en la sección sobre los estadísticos. Otros estadísticos son válidos cuando las variables de la tabla tienen categorías no ordenadas (datos nominales). Para los estadísticos basados en chi-cuadrado (ϕ , V de Cramér y coeficiente de contingencia), los datos deben ser una muestra aleatoria de una distribución multinomial.

Nota: Las variables ordinales pueden ser códigos numéricos que representen categorías (por ejemplo, 1 = *bajo*, 2 = *medio*, 3 = *alto*) o valores de cadena. Sin embargo, se supone que el orden alfabético de los valores de cadena indica el orden correcto de las categorías. Por ejemplo, en una variable de cadena cuyos valores sean *bajo*, *medio*, *alto*, se interpreta el orden de las categorías como *alto*, *bajo*, *medio* (orden que no es el correcto). Por norma general, se puede indicar que es más fiable utilizar códigos numéricos para representar datos ordinales.

Figura 16-1
Resultados de las tablas de contingencia

Service Profitability * Company Size Crosstabulation

Service Profitability	Company Size			Total
	1-500	501-2,500	> 2,500	
Low	200	85	135	420
High	251	106	105	462
Total	451	191	240	882

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9.848	2	.007
Likelihood Ratio	9.852	2	.007
Linear-by-Linear Association	7.869	1	.005
N of Valid Cases	882		

Para obtener tablas de contingencia

- ▶ Elija en los menús:
 - Analizar
 - Estadísticos descriptivos
 - Tablas de contingencia...

Figura 16-2
Cuadro de diálogo Tablas de contingencia



- ▶ Seleccione una o más variables de fila y una o más variables de columna. Si lo desea, tiene la posibilidad de:
 - Seleccionar una o más variables de control.
 - Pulsar en Estadísticos para obtener pruebas y medidas de asociación para tablas o subtablas de doble clasificación.
 - Pulsar en Casillas para obtener porcentajes, residuos y valores esperados y observados.
 - Pulsar en Formato para controlar el orden de las categorías.

Capas de las tablas de contingencia

Si se seleccionan una o más variables de capas, se generará una tabla de contingencia por cada categoría de cada variable de capas (variable de control). Por ejemplo, si emplea una variable de fila, una variable de columna y una variable de capas con dos categorías, obtendrá una tabla de doble clasificación por cada categoría de la variable de capas. Para crear otra capa de variables de control, pulse en Siguiente. Se crean subtablas para cada combinación de categorías para cada variable de la 1ª capa con cada variable de la 2ª capa, y así sucesivamente. Si se solicitan estadísticos y medidas de asociación, se aplicarán sólo a las tablas de doble clasificación.

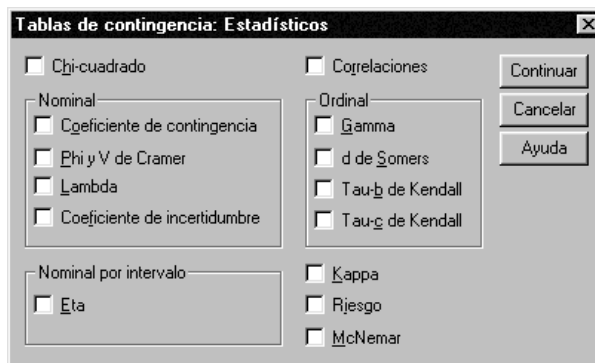
Gráficos de barras agrupadas de las tablas de contingencia

Mostrar los gráficos de barras agrupadas. Los gráficos de barras agrupadas ayudan a resumir los datos por grupos de casos. Hay una agrupación de barras por cada valor de la variable especificada en el cuadro Filas. La variable que define las barras dentro de cada agrupación es la variable especificada en el cuadro Columnas. Por cada valor de esta variable hay un conjunto de barras de distinto color o trama. Si especifica más de una variable en Columnas o en Filas, se generará un gráfico de barras agrupadas por cada combinación de dos variables.

Tablas de contingencia: Estadísticos

Figura 16-3

Cuadro de diálogo Tablas de contingencia: Estadísticos



Chi-cuadrado. Para las tablas con dos filas y dos columnas, seleccione Chi-cuadrado para calcular el chi-cuadrado de Pearson, el chi-cuadrado de la razón de verosimilitud, la prueba exacta de Fisher y el chi-cuadrado corregido de Yates (corrección por continuidad). Para las tablas 2×2 , se calcula la prueba exacta de Fisher cuando una tabla (que no resulte de perder columnas o filas en una tabla mayor) presente una casilla con una frecuencia esperada menor que 5. Para las restantes tablas 2×2 se calcula el chi-cuadrado corregido de Yates. Para las tablas con cualquier número de filas y columnas, seleccione Chi-cuadrado para calcular el chi-cuadrado de Pearson y el chi-cuadrado de la razón de verosimilitud. Cuando ambas variables de tabla son cuantitativas, Chi-cuadrado da como resultado la prueba de asociación lineal por lineal.

Correlaciones. Para las tablas en las que tanto las columnas como las filas contienen valores ordenados, Correlaciones da como resultado rho, el coeficiente de correlación de Spearman (sólo datos numéricos). La rho de Spearman es una medida de asociación entre órdenes de rangos. Cuando ambas variables de tabla (factores) son cuantitativas, Correlaciones da como resultado r , el coeficiente de correlación de Pearson, una medida de asociación lineal entre las variables.

Nominal. Para los datos nominales (sin orden intrínseco, como católico, protestante o judío), puede seleccionar el coeficiente Phi y V de Cramér, el Coeficiente de contingencia, Lambda (lambdas simétricas y asimétricas y tau de Kruskal y Goodman) y el Coeficiente de incertidumbre.

- **Coeficiente de contingencia.** Medida de asociación basada en chi-cuadrado. El valor siempre está comprendido entre 0 y 1. El valor 0 indica que no hay asociación entre la fila y la columna. Los valores cercanos a 1 indican que hay gran relación entre las variables. El valor máximo posible depende del número de filas y columnas de la tabla.
- **Desviación típica (Función de resumen de los gráficos).** Medida de cuánto varían las observaciones respecto a la media, expresada en las mismas unidades que los datos.
- **Lambda.** Medida de asociación que refleja la reducción proporcional en el error cuando se utilizan los valores de la variable independiente para pronosticar los valores de la variable dependiente. Un valor igual a 1 significa que la variable independiente pronostica perfectamente la variable dependiente. Un valor igual a

0 significa que la variable independiente no ayuda en absoluto a pronosticar la variable dependiente.

- **Coefficiente de incertidumbre.** Medida de asociación que indica la reducción proporcional del error cuando los valores de una variable se emplean para pronosticar valores de la otra variable. Por ejemplo, un valor de 0.83 indica que el conocimiento de una variable reduce en un 83% el error al pronosticar los valores de la otra variable. SPSS calcula tanto la versión simétrica como la asimétrica del coeficiente de incertidumbre.

Ordinal. Para las tablas en las que tanto las filas como las columnas contienen valores ordenados, seleccione Gamma (orden cero para tablas de doble clasificación y condicional para tablas cuyo factor de clasificación va de 3 a 10), Tau-b de Kendall y Tau-c de Kendall. Para pronosticar las categorías de columna de las categorías de fila, seleccione d de Somers.

- **Gamma.** Medida de asociación simétrica entre dos variables ordinales cuyo valor siempre está comprendido entre menos 1 y 1. Los valores próximos a 1, en valor absoluto, indican una fuerte relación entre las dos variables. Los valores próximos a cero indican que hay poca o ninguna relación entre las dos variables. Para las tablas de doble clasificación, se muestran las gammas de orden cero. Para las tablas de tres o más factores de clasificación, se muestran las gammas condicionales.
- **d de Somers.** Medida de asociación entre dos variables ordinales que toma un valor comprendido entre -1 y 1. Los valores próximos a 1, en valor absoluto, indican una fuerte relación entre las dos variables. Los valores próximos a cero indican que hay poca o ninguna relación entre las dos variables. La d de Somers es una extensión asimétrica de gamma que difiere sólo en la inclusión del número de pares no empatados en la variable independiente. También se calcula una versión no simétrica de este estadístico.
- **Tau-b de Kendall.** Medida no paramétrica de la correlación para variables ordinales o de rangos que tiene en consideración los empates. El signo del coeficiente indica la dirección de la relación y su valor absoluto indica la magnitud de la misma, de tal modo que los mayores valores absolutos indican relaciones más fuertes. Los valores posibles van de -1 a 1, pero un valor de -1 o +1 sólo se puede obtener a partir de tablas cuadradas.
- **Tau-c de Kendall.** Medida no paramétrica de asociación para variables ordinales que ignora los empates. El signo del coeficiente indica la dirección de la relación y su valor absoluto indica la magnitud de la misma, de tal modo que los mayores

valores absolutos indican relaciones más fuertes. Los valores posibles van de -1 a 1, pero un valor de -1 o +1 sólo se puede obtener a partir de tablas cuadradas.

Nominal por intervalo. Cuando una variable es categórica y la otra es cuantitativa, seleccione Eta. La variable categórica debe codificarse numéricamente.

- **Eta.** Medida de asociación cuyo valor siempre está comprendido entre 0 y 1. El valor 0 indica que no hay asociación entre las variables de fila y de columna. Los valores cercanos a 1 indican que hay gran relación entre las variables. Eta resulta apropiada para una variable dependiente medida en una escala de intervalo (por ejemplo, ingresos) y una variable independiente con un número limitado de categorías (por ejemplo, género). Se calculan dos valores de eta: uno trata la variable de las filas como una variable de intervalo; el otro trata la variable de las columnas como una variable de intervalo.

Kappa. La kappa de Cohen mide el acuerdo entre las evaluaciones de dos jueces cuando ambos están valorando el mismo objeto. Un valor igual a 1 indica un acuerdo perfecto. Un valor igual a 0 indica que el acuerdo no es mejor que el que se obtendría por azar. Kappa sólo está disponible para las tablas cuadradas (tablas en las que ambas variables tienen el mismo número de categorías).

Riesgo. Para tablas 2x2, una medida del grado de asociación entre la presencia de un factor y la ocurrencia de un evento. Si el intervalo de confianza para el estadístico incluye un valor de 1, no se podrá asumir que el factor está asociado con el evento. Cuando la ocurrencia del factor es rara, se puede utilizar la razón de las ventajas (odds ratio) como estimación del riesgo relativo.

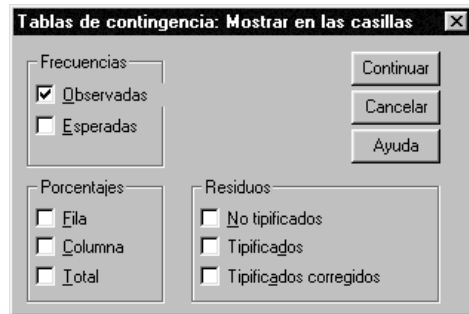
McNemar. Prueba no paramétrica para dos variables dicotómicas relacionadas. Contrasta los cambios en las respuestas utilizando la distribución de chi-cuadrado. Es útil para detectar cambios en las respuestas debidas a la intervención experimental en los diseños del tipo “antes-después”. Para las tablas cuadradas de mayor orden se informa de la prueba de simetría de McNemar-Bowker.

Estadísticos de Cochran y de Mantel-Haenszel. Los estadísticos de Cochran y Mantel-Haenszel pueden utilizarse para contrastar la independencia entre una variable de factor dicotómica y una variable de respuesta dicotómica, condicionada por los patrones en las covariables, los cuales vienen definidos por la variable o variables de las capas (variables de control). Tenga en cuenta que mientras que otros estadísticos se calculan capa por capa, los estadísticos de Cochran y Mantel-Haenszel se calculan una sola vez para todas las capas.

Tablas de contingencia: Mostrar en las casillas

Figura 16-4

Cuadro de diálogo Tablas de contingencia: Mostrar en las casillas



Para ayudarle a descubrir las tramas en los datos que contribuyen a una prueba de chi-cuadrado significativa, el procedimiento Tablas de contingencia muestra las frecuencias esperadas y tres tipos de residuos (desviaciones) que miden la diferencia entre las frecuencias observadas y las esperadas. Cada casilla de la tabla puede contener cualquier combinación de recuentos, porcentajes y residuos seleccionados.

Recuentos. El número de casos realmente observados y el número de casos esperados si las variables de fila y columna son independientes entre sí.

Porcentajes. Los porcentajes se pueden sumar a través de las filas o a lo largo de las columnas. También se encuentran disponibles los porcentajes del número total de casos representados en la tabla (una capa).

Residuos. Los residuos brutos no tipificados presentan la diferencia entre los valores observados y los esperados. También se encuentran disponibles los residuos tipificados y tipificados corregidos.

- **No tipificados.** La diferencia entre un valor observado y el valor pronosticado por el modelo. El valor pronosticado es el número de casos que se esperarían en la casilla si no hubiera relación entre las dos variables. Un residuo positivo indica que hay más casos en la casilla de los que habría en ella si las variables de fila y columna fueran independientes.

- **Tipificados.** El residuo dividido por una estimación de su error típico. Los residuos tipificados, que son conocidos también como los residuos de Pearson o residuos estandarizados, tienen una media de 0 y una desviación típica de 1.
- **Tipificados corregidos.** El residuo de una casilla (valor observado menos valor pronosticado) dividido por una estimación de su error típico. El residuo tipificado resultante viene expresado en unidades de desviación típica, por encima o por debajo de la media.

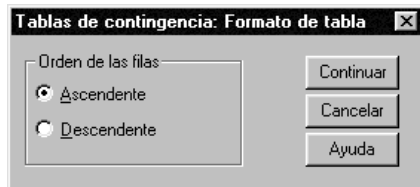
Ponderaciones no enteras. Los recuentos de las casillas suelen ser valores enteros, ya que representan el número de casos de cada casilla. Sin embargo, si el archivo de datos está ponderado en un momento determinado por una variable de ponderación con valores fraccionarios (por ejemplo, 1,25), los recuentos de las casillas pueden que también sean valores fraccionarios. Puede truncar o redondear estos valores antes o después de calcular los recuentos de las casillas o bien utilizar recuentos de casillas fraccionarios en la presentación de las tablas y los cálculos de los estadísticos.

- **Redondear las frecuencias de casilla.** Las ponderaciones de los casos se utilizan tal cual, pero las ponderaciones acumuladas en las casillas se redondean antes de calcular cualquiera de los estadísticos.
- **Truncar las frecuencias de casilla.** Las ponderaciones de los casos se utilizan tal cual, pero las ponderaciones acumuladas en las casillas se truncan antes de calcular cualquiera de los estadísticos.
- **Redondear las ponderaciones de los casos.** Se redondean las ponderaciones de los casos antes de utilizarlas.
- **Truncar las ponderaciones de los casos.** Se truncan las ponderaciones de los casos antes de utilizarlas.
- **Sin ajustes.** Las ponderaciones de los casos se utilizan tal cual y se utilizan las frecuencias de casilla fraccionales. Sin embargo, cuando se solicitan Estadísticos exactos (disponibles sólo con la opción Pruebas exactas), los pesos acumulados en las casillas se truncan o redondean antes de calcular los estadísticos de las Pruebas exactas.

Tablas de contingencia: Formato de tabla

Figura 16-5

Cuadro de diálogo Tablas de contingencia: Formato de tabla



Puede ordenar las filas en orden ascendente o descendente de los valores de la variable de fila.

Resumir

El procedimiento Resumir calcula estadísticos de subgrupo para las variables dentro de las categorías de una o más variables de agrupación. Se cruzan todos los niveles de las variables de agrupación. Puede elegir el orden en el que se mostrarán los estadísticos. También se muestran estadísticos de resumen para cada variable a través de todas las categorías. Los valores de los datos en cada categoría pueden mostrarse en una lista o suprimirse. Con grandes conjuntos de datos, tiene la opción de listar sólo los primeros n casos.

Ejemplo. ¿Cuál es la media de las ventas por regiones o por tipo de cliente? Podrá descubrir que el importe medio de las ventas es ligeramente superior en la región occidental respecto a las demás regiones, y que la media más alta se da entre los clientes de empresas privadas de la zona occidental .

Estadísticos. Suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del N total, porcentaje de la suma en, porcentaje de N en, media geométrica y media armónica.

Datos. Las variables de agrupación son variables categóricas cuyos valores pueden ser numéricos o de cadena corta. El número de categorías debe ser razonablemente pequeño. Las otras variables deben poder ordenarse mediante rangos.

Supuestos. Algunos de los estadísticos opcionales de subgrupo, como la media y la desviación típica, se basan en la teoría normal y son adecuados para variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana y el rango, son adecuados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad.

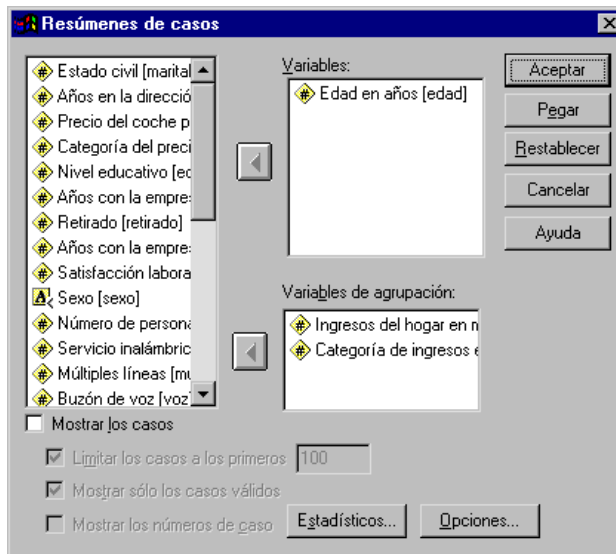
Figura 17-1
Resultados del procedimiento Resumen

Resúmenes de casos					
Venta promedio del producto por región e industria					
Región	Región				Total
	Gobierno	Privada	Educación	Sanidad	
Este	\$3.594,65	\$3.953,76	\$3.764,91	\$3.722,32	\$3,735.45
Central	\$3.370,12	\$3.268,47	\$3.317,81	\$3,165.11	\$3,305.03
Oeste	\$3.552,50	\$4.649,00	\$4.276,25	\$4,027.00	\$4,079.46
Total	\$3.503,75	\$3.727,50	\$3.579,76	\$3,456.93	\$3,576.52

Para obtener resúmenes de casos

- Elija en los menús:
Analizar
Informes
Resúmenes de casos...

Figura 17-2
Cuadro de diálogo Resúmenes de casos



- Seleccione una o más variables.

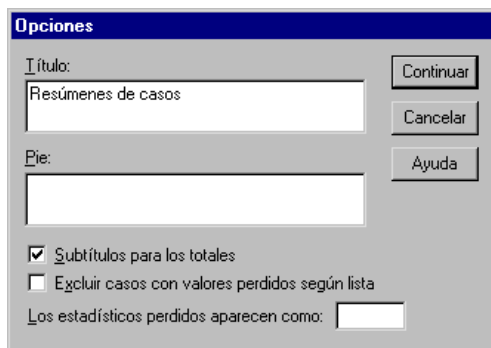
Si lo desea, tiene la posibilidad de:

- Seleccionar una o más variables de agrupación para dividir los datos en subgrupos.
- Pulsar en Opciones para cambiar el título de los resultados, añadir un texto al pie debajo de los resultados o excluir los casos con valores perdidos.
- Pulsar en Estadísticos para acceder a estadísticos adicionales.
- Seleccionar Mostrar los casos para listar los casos en cada subgrupo. Por defecto, el sistema enumera sólo los 100 primeros casos del archivo. Puede aumentar o disminuir el valor de Limitar los casos a los primeros *n* o desactivar ese elemento para enumerar todos los casos.

Resumir: Opciones

Figura 17-3

Cuadro de diálogo Resúmenes de casos: Opciones



The image shows a dialog box titled "Opciones" with a blue header. It contains the following elements:

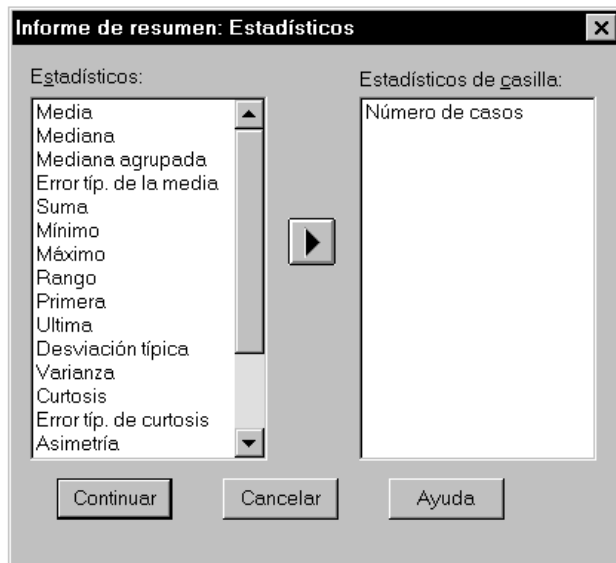
- A "Título:" label followed by a text input field containing "Resúmenes de casos".
- A "Pie:" label followed by an empty text input field.
- A checked checkbox labeled "Subtítulos para los totales".
- An unchecked checkbox labeled "Excluir casos con valores perdidos según lista".
- A label "Los estadísticos perdidos aparecen como:" followed by a small empty text input field.
- Three buttons on the right: "Continuar", "Cancelar", and "Ayuda".

Resumir permite cambiar el título de los resultados o añadir un texto que aparecerá debajo de la tabla de resultados. Puede controlar el ajuste de las líneas en los títulos y textos escribiendo `\n` en el lugar donde desee insertar una línea de separación.

Además, puede elegir entre mostrar o suprimir los subtítulos para los totales e incluir o excluir los casos con valores perdidos para cualquiera de las variables utilizadas en cualquiera de los análisis. A menudo es aconsejable representar los casos perdidos en los resultados con un punto o un asterisco. Introduzca un carácter, frase o código que desee que aparezca cuando haya un valor perdido; de lo contrario, no se aplicará ningún tratamiento especial a los casos perdidos en los resultados.

Resumir: Estadísticos

Figura 17-4
Cuadro de diálogo Resumir los casos: Estadísticos



Puede elegir uno o más de los siguientes estadísticos de subgrupo para las variables dentro de cada categoría de cada variable de agrupación: suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del N total, porcentaje de la suma en, porcentaje de N en, media geométrica y media armónica. El orden en el que aparecen los estadísticos en la lista Estadísticos de casilla es el orden en el que se mostrarán en los resultados. También se muestran estadísticos de resumen para cada variable a través de todas las categorías.

Primero. Muestra el primer valor de los datos encontrado en el archivo de datos.

Media geométrica. La raíz n -ésima del producto de los valores de los datos, donde n representa el número de casos.

Mediana agrupada. La mediana calculada para los datos que se codifican en grupos. Por ejemplo, para los datos de edades, si cada valor de la treintena se codifica como 35, cada valor de los 40 se codifica como 45, etc. la mediana agrupada es la mediana calculada a partir de los datos codificados.

Media armónica. Se utiliza para estimar el tamaño promedio de un grupo cuando los tamaños de las muestras de los grupos no son iguales. La media armónica es el número total de muestras dividido por la suma de los inversos de los tamaños de las muestras.

Curtosis. Medida del grado en que las observaciones están agrupadas en torno al punto central. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

Último. Muestra el último valor de los datos encontrado en el archivo de datos.

Máximo. El mayor valor de una variable numérica.

Media. Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.

Mediana. Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Cuando el número de observaciones es par, la mediana es el promedio de las dos observaciones centrales, una vez que han sido ordenadas de manera ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

Mínimo. Valor más pequeño de una variable numérica.

N. Número de casos (observaciones o registros).

Porcentaje del N total. Porcentaje del número total de casos en cada categoría.

Porcentaje de la suma total. Porcentaje de la suma total en cada categoría.

Amplitud. Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.

Asimetría. Medida de la asimetría de una distribución. La distribución normal es simétrica por lo que tiene un valor de asimetría 0. Una distribución que tenga una asimetría positiva significativa tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Un valor de asimetría mayor que 1, en valor absoluto, indica generalmente una distribución que difiere de manera significativa de la distribución normal.

Error típico de la curtosis. La razón de la curtosis sobre su error típico puede utilizarse como contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la curtosis indica que las colas son más largas que las de una distribución normal; por el contrario, un valor extremo y negativo indica que las colas son más cortas (llegando a tener forma de caja como en la distribución uniforme).

Error típico de la asimetría. La razón de la asimetría sobre su error típico puede utilizarse como un contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la asimetría indica una cola larga a la derecha; un valor extremo y negativo indica una cola larga por la izquierda.

Suma. La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

Varianza. Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.

Medias

El procedimiento Medias calcula medias de subgrupo y estadísticos univariados relacionados para variables dependientes dentro de las categorías de una o más variables independientes. Si lo desea, puede obtener el análisis de varianza de un factor, la η^2 y pruebas de linealidad.

Ejemplo. Mida la cantidad media de grasa absorbida en función de tres tipos distintos de aceite comestible y realice un análisis de varianza de un factor para comprobar si difieren las medias.

Estadísticos. Suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del N total, porcentaje de la suma en, porcentaje de N en, media geométrica y media armónica. Las opciones incluyen: análisis de varianza, η^2 , η^2 cuadrado y pruebas de linealidad de R y R^2 .

Datos. Las variables dependientes son cuantitativas y las independientes son categóricas. Los valores de las variables categóricas pueden ser numéricos o de cadena corta.

Supuestos. Algunos de los estadísticos opcionales de subgrupo, como la media y la desviación típica, se basan en la teoría normal y son adecuados para variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana son adecuados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad. El análisis de varianza es robusto a las desviaciones de la normalidad, aunque los datos de cada casilla deberían ser simétricos. El análisis de varianza también supone que los grupos proceden de poblaciones con la misma varianza. Para comprobar este supuesto, utilice la prueba de homogeneidad de las varianzas de Levene, disponible en el procedimiento ANOVA de un factor.

Figura 18-1
Resultados del procedimiento Medias

Informe

Gramos de grasa absorbida

De Cacahuete	Media	72.00
	N	6
	Desv. típ.	13.34
Manteca	Media	85.00
	N	6
	Desv. típ.	7.77
De Maíz	Media	62.00
	N	6
	Desv. típ.	8.22
Total	Media	73.00
	N	18
	Desv. típ.	13.56

Tabla de ANOVA

			Suma de cuadrados	gl	Media cuadrática	F	Sig.
Gramos de grasa absorbida * Tipo de aceite	Inter-grupos	(Combinadas)	1596.000	2	798.000	7.824	.005
	Intra-grupos		1530.000	15	102.000		
	Total		3126.000	17			

Para obtener medias de subgrupo

- ▶ Elija en los menús:
 - Analizar
 - Comparar medias
 - Medias...

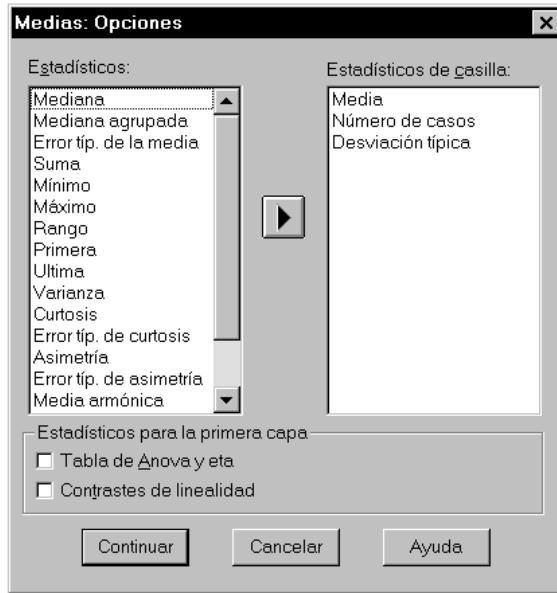
Figura 18-2
Cuadro de diálogo Medias



- ▶ Seleccione una o más variables dependientes.
- ▶ Hay dos formas de seleccionar las variables independientes categóricas:
 - Seleccione una o más variables independientes. Se mostrarán resultados individuales para cada variable independiente.
 - Seleccione una o más capas de variables independientes. Cada capa subdivide consecutivamente la muestra. Si tiene una variable independiente en Capa 1 y otra en Capa 2, los resultados se mostrarán en una tabla cruzada en contraposición a tablas individuales para cada variable independiente.
- ▶ Si lo desea, tiene la posibilidad de:
 - Pulsar en Opciones si desea obtener estadísticos opcionales, tabla de análisis de varianza, eta, eta cuadrado, R , y R^2 .

Medias: Opciones

Figura 18-3
Cuadro de diálogo Medias: Opciones



Puede elegir uno o más de los siguientes estadísticos de subgrupo para las variables dentro de cada categoría de cada variable de agrupación: suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del N total, porcentaje de la suma en, porcentaje de N en, media geométrica, media armónica. Se puede cambiar el orden de aparición de los estadísticos de subgrupo. El orden en el que aparecen en la lista Estadísticos de casilla es el mismo orden que presentarán en los resultados. También se muestran estadísticos de resumen para cada variable a través de todas las categorías.

Primero. Muestra el primer valor de los datos encontrado en el archivo de datos.

Media geométrica. La raíz n -ésima del producto de los valores de los datos, donde n representa el número de casos.

Mediana agrupada. La mediana calculada para los datos que se codifican en grupos. Por ejemplo, para los datos de edades, si cada valor de la treintena se codifica como 35, cada valor de los 40 se codifica como 45, etc. la mediana agrupada es la mediana calculada a partir de los datos codificados.

Media armónica. Se utiliza para estimar el tamaño promedio de un grupo cuando los tamaños de las muestras de los grupos no son iguales. La media armónica es el número total de muestras dividido por la suma de los inversos de los tamaños de las muestras.

Curtosis. Medida del grado en que las observaciones están agrupadas en torno al punto central. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

Último. Muestra el último valor de los datos encontrado en el archivo de datos.

Máximo. El mayor valor de una variable numérica.

Media. Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.

Mediana. Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Cuando el número de observaciones es par, la mediana es el promedio de las dos observaciones centrales, una vez que han sido ordenadas de manera ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

Mínimo. Valor más pequeño de una variable numérica.

N. Número de casos (observaciones o registros).

Porcentaje del N total. Porcentaje del número total de casos en cada categoría.

Porcentaje de la suma total. Porcentaje de la suma total en cada categoría.

Amplitud. Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.

Asimetría. Medida de la asimetría de una distribución. La distribución normal es simétrica por lo que tiene un valor de asimetría 0. Una distribución que tenga una asimetría positiva significativa tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Un valor de asimetría mayor que 1, en valor absoluto, indica generalmente una distribución que difiere de manera significativa de la distribución normal.

Error típico de la curtosis. La razón de la curtosis sobre su error típico puede utilizarse como contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la curtosis indica que las colas son más largas que las de una distribución normal; por el contrario, un valor extremo y negativo indica que las colas son más cortas (llegando a tener forma de caja como en la distribución uniforme).

Error típico de la asimetría. La razón de la asimetría sobre su error típico puede utilizarse como un contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la asimetría indica una cola larga a la derecha; un valor extremo y negativo indica una cola larga por la izquierda.

Suma. La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

Varianza. Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.

Estadísticos para la primera capa

Tabla de ANOVA y eta. Muestra una tabla de análisis de varianza de un factor y calcula la eta y la eta cuadrado (son medidas de asociación) para cada variable independiente de la primera capa.

Contrastes de linealidad. Calcula la suma de cuadrados, los grados de libertad y la media cuadrática asociados con los componentes lineal y no lineal, así como la razón F, la R y la R cuadrado. Si la variable independiente es una cadena corta entonces la linealidad no se calcula.

Cubos OLAP

El procedimiento Cubos OLAP (siglas del inglés On-Line Analytic Processing, «Procesamiento analítico interactivo») calcula totales, medias y otros estadísticos univariantes para variables de resumen continuas dentro de las categorías de una o más variables categóricas de agrupación. En la tabla se creará una nueva capa para cada categoría de cada variable de agrupación.

Ejemplo. El total y el promedio de ventas para diversas regiones y líneas de producto, dentro de las regiones.

Estadísticos. Suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de casos totales, porcentaje de la suma total, porcentaje de casos totales dentro de las variables agrupadas, porcentaje de la suma total dentro de las variables agrupadas, media geométrica y media armónica.

Datos. Las variables de resumen son cuantitativas (variables continuas medidas en una escala de intervalo o de razón) y las variables de agrupación son categóricas. Los valores de las variables categóricas pueden ser numéricos o de cadena corta.

Supuestos. Algunos de los estadísticos opcionales de subgrupo, como la media y la desviación típica, se basan en la teoría normal y son adecuados para variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana y el rango, son adecuados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad.

Figura 19-1
Resultado de cubos OLAP

**Ventas realizadas
 por departamento
 y región durante 1996**

División: Total

Región: Total

Suma	\$145.038.250
Media	\$371.893
Mediana	\$307.500
Desv. típica	\$171.311

**Ventas realizadas
 por departamento
 y región durante 1996**

División: Productos del cliente

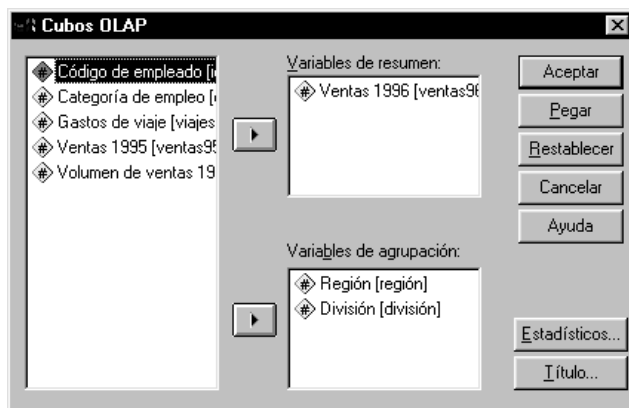
Región: Este

Suma	\$18.548.100
Media	\$289.814,06
Mediana	\$273.600.00
Desv. típica	\$80.674,66

Para obtener cubos OLAP

- ▶ Elija en los menús:
 Analizar
 Informes
 Cubos OLAP...

Figura 19-2
Cuadro de diálogo Cubos OLAP



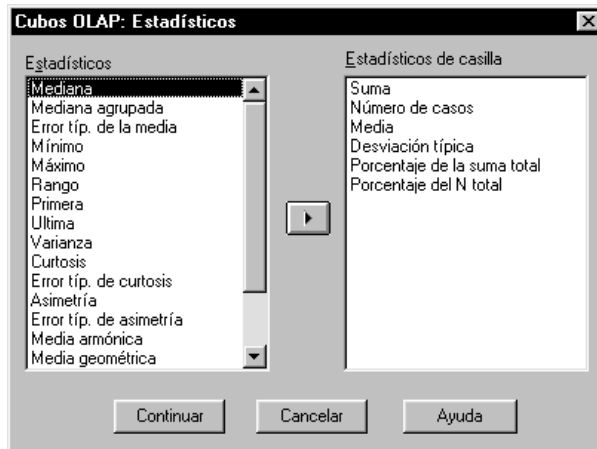
- ▶ Seleccione una o más variables de resumen continuas.
- ▶ Seleccione una o más variables de agrupación categóricas.

Si lo desea, tiene la posibilidad de:

- Seleccionar diferentes estadísticos de resumen (pulse en Estadísticos). Debe seleccionar una o más variables de agrupación para poder seleccionar estadísticos de resumen.
- Calcule las diferencias existentes entre los pares de variables y los pares de grupos definidos por una variable de agrupación (pulse en Diferencias).
- Crear títulos de tabla personalizados (pulse en Título).

Cubos OLAP: Estadísticos

Figura 19-3
Cuadro de diálogo Cubos OLAP: Estadísticos



Puede elegir uno o varios de los siguientes estadísticos de subgrupo para las variables de resumen dentro de cada categoría de cada variable de agrupación: Suma, Número de casos, Media, Mediana, Mediana agrupada, Error típico de la media, Mínimo, Máximo, Rango, Valor de la variable para la primera categoría de la variable de agrupación, Valor de la variable para la última categoría de la variable de agrupación, Desviación típica, Varianza, Curtosis, Error típico de curtosis, Asimetría, Error típico de asimetría, Porcentaje de casos totales, Porcentaje de la suma total, Porcentaje de casos totales dentro de las variables de agrupación, Porcentaje de la suma total dentro de las variables de agrupación, Media geométrica y Media armónica.

Se puede cambiar el orden de aparición de los estadísticos de subgrupo. El orden en el que aparecen en la lista Estadísticos de casilla es el mismo orden que presentarán en los resultados. También se muestran estadísticos de resumen para cada variable a través de todas las categorías.

Primero. Muestra el primer valor de los datos encontrado en el archivo de datos.

Media geométrica. La raíz enésima del producto de los valores de los datos, donde n representa el número de casos.

Mediana agrupada. La mediana calculada para los datos que se codifican en grupos. Por ejemplo, para los datos de edades, si cada valor de la treintena se codifica como 35, cada valor de los 40 se codifica como 45, etc. la mediana agrupada es la mediana calculada a partir de los datos codificados.

Media armónica. Se utiliza para estimar el tamaño promedio de un grupo cuando los tamaños de las muestras de los grupos no son iguales. La media armónica es el número total de muestras dividido por la suma de los inversos de los tamaños de las muestras.

Curtosis. Medida del grado en que las observaciones están agrupadas en torno al punto central. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

Último. Muestra el último valor de los datos encontrado en el archivo de datos.

Máximo. El mayor valor de una variable numérica.

Media. Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.

Mediana. Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Cuando el número de observaciones es par, la mediana es el promedio de las dos observaciones centrales, una vez que han sido ordenadas de manera ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

Mínimo. Valor más pequeño de una variable numérica.

N. Número de casos (observaciones o registros).

Porcentaje del N en. Porcentaje del número de casos para la variable de agrupación especificada dentro de las categorías de otras variables de agrupación. Si sólo tiene una variable de agrupación, este valor es idéntico al porcentaje del número de casos total.

Porcentaje de la suma en. Porcentaje de la suma para la variable de agrupación especificada dentro de las categorías de otras variables de agrupación. Si sólo tiene una variable de agrupación, este valor es idéntico al porcentaje de la suma total.

Porcentaje del N total. Porcentaje del número total de casos en cada categoría.

Porcentaje de la suma total. Porcentaje de la suma total en cada categoría.

Amplitud. Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.

Asimetría. Medida de la asimetría de una distribución. La distribución normal es simétrica por lo que tiene un valor de asimetría 0. Una distribución que tenga una asimetría positiva significativa tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Un valor de asimetría mayor que 1, en valor absoluto, indica generalmente una distribución que difiere de manera significativa de la distribución normal.

Error típico de la curtosis. La razón de la curtosis sobre su error típico puede utilizarse como contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la curtosis indica que las colas son más largas que las de una distribución normal; por el contrario, un valor extremo y negativo indica que las colas son más cortas (llegando a tener forma de caja como en la distribución uniforme).

Error típico de la asimetría. La razón de la asimetría sobre su error típico puede utilizarse como un contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la asimetría indica una cola larga a la derecha; un valor extremo y negativo indica una cola larga por la izquierda.

Suma. La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

Varianza. Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.

Cubos OLAP: Diferencias

Figura 19-4
Cuadro de diálogo Cubos OLAP: Diferencias

Cubos OLAP: Diferencias

Diferencias para estadísticos de resumen:

Ninguno

Diferencias entre variables

Diferencias entre grupos

Tipo de diferencia:

Diferencia de porcentaje

Diferencia aritmética

Continuar

Cancelar

Ayuda

Diferencias entre variables:

Variable:

Menos variable:

Etiqueta de porcentaje:

Etiqueta aritmética:

Pares:

Eliminar par

Diferencias entre grupos de casos:

Variable de agrupación: género

Categoría:

Menos categoría:

Etiqueta de porcentaje:

Etiqueta aritmética:

Pares:

Eliminar par

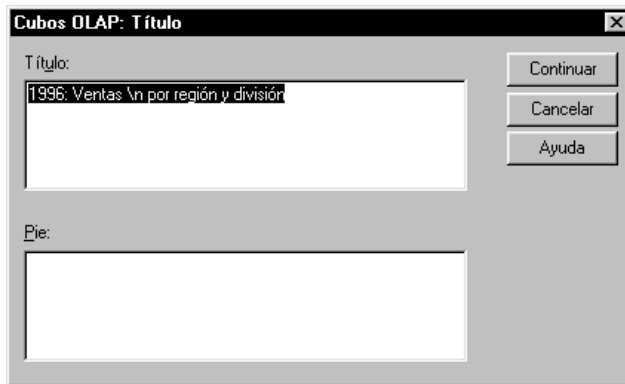
Este cuadro de diálogo le permite calcular el porcentaje y las diferencias aritméticas entre las variables de resumen o entre los grupos definidos por una variable de agrupación. Las diferencias se calculan para todas las medidas seleccionadas en el cuadro de diálogo Cubos OLAP: Estadísticos.

Diferencias entre variables. Calcula las diferencias entre pares de variables. Los valores de los estadísticos de resumen para la segunda variable de cada par (la variable Menos) se restan de los valores de los estadísticos de resumen correspondientes a la primera variable del par. En cuanto a las diferencias de porcentaje, el valor de la variable de resumen para la variable Menos es el que se usa como denominador. Debe seleccionar al menos dos variables de resumen en el cuadro de diálogo principal para poder especificar las diferencias entre las variables.

Diferencias entre grupos de casos. Calcula las diferencias entre pares de grupos definidos por una variable de agrupación. Los valores de los estadísticos de resumen para la segunda categoría de cada par (la variable Menos) se restan de los valores de los estadísticos de resumen correspondientes a la primera categoría del par. Las diferencias de porcentaje utilizan el valor del estadístico de resumen de la categoría Menos como denominador. Debe seleccionar una o más variables de agrupación en el cuadro de diálogo principal para poder especificar las diferencias entre los grupos.

Cubos OLAP: Título

Figura 19-5
Cuadro de diálogo Cubos OLAP: Título



Puede cambiar el título de los resultados o añadir un texto al pie que aparecerá debajo de la tabla de resultados. También puede controlar el ajuste de las líneas de los títulos y de los textos al pie escribiendo `\n` en el lugar del texto donde desea insertar una línea de separación.

Pruebas T

Hay tres tipos de pruebas *t*:

Prueba T para muestras independientes (prueba T para dos muestras). Compara las medias de una variable para dos grupos de casos. Se ofrecen estadísticos descriptivos para cada grupo y la prueba de Levene sobre la igualdad de las varianzas, así como valores *t* de igualdad de varianzas y varianzas desiguales y un intervalo de confianza al 95% para la diferencia entre las medias.

Prueba T para muestras relacionadas (prueba T dependiente). Compara las medias de dos variables en un solo grupo. Esta prueba también se utiliza para pares relacionados o diseños de estudio de control de casos. El resultado incluye estadísticos descriptivos de las variables que se van a contrastar, la correlación entre ellas, estadísticos descriptivos de las diferencias emparejadas, la prueba *t* y un intervalo de confianza al 95%.

Prueba *t* para una muestra. Compara la media de una variable con un valor conocido o hipotetizado. Se muestran estadísticos descriptivos para las variables de contraste junto con la prueba *t*. Por defecto, en los resultados se incluye un intervalo de confianza al 95% para la diferencia entre la media de la variable de contraste y el valor hipotetizado de la prueba.

Prueba T para muestras independientes

El procedimiento Prueba T para muestras independientes compara las medias de dos grupos de casos. Lo ideal es que para esta prueba los sujetos se asignen aleatoriamente a dos grupos, de forma que cualquier diferencia en la respuesta sea debida al tratamiento (o falta de tratamiento) y no a otros factores. Este caso no ocurre si se comparan los ingresos medios para hombres y mujeres. El sexo de una persona no se asigna aleatoriamente. En estas situaciones, debe asegurarse de que las

diferencias en otros factores no enmascaren o resalten una diferencia significativa entre las medias. Las diferencias de ingresos medios pueden estar sometidas a la influencia de factores como los estudios y no solamente el sexo.

Ejemplo. Se asigna aleatoriamente un grupo de pacientes con hipertensión arterial a un grupo con placebo y otro con tratamiento. Los sujetos con placebo reciben una pastilla inactiva y los sujetos con tratamiento reciben un nuevo medicamento del cual se espera que reduzca la tensión arterial. Después de tratar a los sujetos durante dos meses, se utiliza la prueba t para dos muestras para comparar la tensión arterial media del grupo con placebo y del grupo con tratamiento. Cada paciente se mide una sola vez y pertenece a un solo grupo.

Estadísticos. Para cada variable: tamaño muestral, media, desviación típica y error típico de la media. Para la diferencia entre las medias: media, error típico e intervalo de confianza (puede especificar el nivel de confianza). Pruebas: prueba de Levene sobre la igualdad de varianzas y pruebas t de varianzas combinadas y separadas sobre la igualdad de las medias.

Datos. Los valores de la variable cuantitativa de interés se hallan en una única columna del archivo de datos. El procedimiento utiliza una variable de agrupación con dos valores para separar los casos en dos grupos. La variable de agrupación puede ser numérica (valores como 1 y 2, o 6,25 y 12,5) o de cadena corta (como *sí* y *no*). También puede usar una variable cuantitativa, como la *edad*, para dividir los casos en dos grupos especificando un punto de corte (el punto de corte 21 divide la *edad* en un grupo de menos de 21 años y otro de más de 21).

Supuestos. Para la prueba t de igualdad de varianzas, las observaciones deben ser muestras aleatorias independientes de distribuciones normales con la misma varianza de población. Para la prueba t de varianzas desiguales, las observaciones deben ser muestras aleatorias independientes de distribuciones normales. La prueba t para dos muestras es bastante robusta a las desviaciones de la normalidad. Al contrastar las distribuciones gráficamente, compruebe que son simétricas y que no contienen valores atípicos.

Figura 20-1
Resultados de la prueba T para muestras independientes

			N	Mean	Std. Deviation	Std. Error Mean			
Blood pressure	Treatment	placebo	10	142.50	17.04	5.39			
		new_drug	10	116.40	13.62	4.31			

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Significance	t	df	Significance (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Mean	
									Lower	Upper
Blood pressure	Equal variances assumed	.134	.719	3.783	18	.001	26.10	6.90	11.61	40.59
	Equal variances not assumed			3.783	17.163	.001	26.10	6.90	11.56	40.64

Para obtener una prueba T para muestras independientes

- Elija en los menús:
 - Analizar
 - Comparar medias
 - Prueba T para muestras independientes...

Figura 20-2
Cuadro de diálogo Prueba T para muestras independientes



- ▶ Seleccione una o más variables de contraste cuantitativas. Se calcula una prueba t para cada variable.
- ▶ Seleccione una sola variable de agrupación y pulse en Definir grupos para especificar dos códigos para los grupos que desee comparar.

Si lo desea, puede pulsar en Opciones para controlar el tratamiento de los datos perdidos y el nivel del intervalo de confianza.

Definición de grupos en la prueba T para muestras independientes

Figura 20-3
Cuadro de diálogo Definir grupos para variables numéricas

Para las variables de agrupación numéricas, defina los dos grupos de la prueba t especificando dos valores o un punto de corte:

- **Usar valores especificados.** Escriba un valor para el Grupo 1 y otro para el Grupo 2. Los casos con otros valores quedarán excluidos del análisis. Los números no tienen que ser enteros (por ejemplo, 6,25 y 12,5 son válidos).
- **Punto de corte.** Opcionalmente, puede escribir un número que divida los valores de la variable de agrupación en dos conjuntos. Todos los casos con valores menores que el punto de corte forman un grupo, y los casos con valores mayores o iguales que el punto de corte forman el otro grupo.

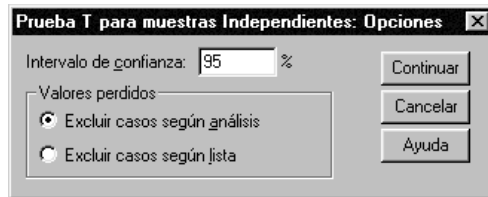
Figura 20-4
Cuadro de diálogo Definir grupos para variables de cadena

Para las variables de agrupación de cadena corta, escriba una cadena para el Grupo 1 y otra para el Grupo 2; por ejemplo *sí* y *no*. Los casos con otras cadenas se excluyen del análisis.

Prueba T para muestras independientes: Opciones

Figura 20-5

Cuadro de diálogo Prueba T para muestras independientes: Opciones



Intervalo de confianza. Por defecto se muestra un intervalo de confianza al 95% para la diferencia entre las medias. Introduzca un valor entre 1 y 99 para solicitar otro nivel de confianza.

Valores perdidos. Si ha probado varias variables y se han perdido los datos de una o más de ellas, puede indicar al procedimiento qué casos desea incluir (o excluir):

- **Excluir datos perdidos según análisis.** Cada prueba t utiliza todos los casos que tienen datos válidos para las variables contrastadas. Los tamaños muestrales pueden variar de una prueba a otra.
- **Excluir casos según lista.** Cada prueba t utiliza sólo aquellos casos que contienen datos válidos para todas las variables utilizadas en las pruebas t solicitadas. El tamaño muestral es constante en todas las pruebas.

Prueba T para muestras relacionadas

El procedimiento Prueba T para muestras relacionadas compara las medias de dos variables de un solo grupo. Calcula las diferencias entre los valores de las dos variables de cada caso y contrasta si la media difiere de 0.

Ejemplo. En un estudio sobre la hipertensión sanguínea, se toma la tensión a todos los pacientes al comienzo del estudio, se les aplica un tratamiento y se les toma la tensión otra vez. De esta manera, a cada paciente le corresponden dos medidas, normalmente

denominadas medidas *pre* y *post*. Un diseño alternativo para el que se utiliza esta prueba consiste en un estudio de pares relacionados o un estudio de control de casos. En estos, cada registro en el archivo de datos contiene la respuesta del paciente y de su sujeto de control correspondiente. En un estudio sobre la tensión sanguínea, pueden emparejarse pacientes y controles por edad (un paciente de 75 años con un miembro del grupo de control de 75 años).

Estadísticos. Para cada variable: media, tamaño muestral, desviación típica y error típico de la media. Para cada pareja de variables: correlación, diferencia promedio entre las medias, prueba *t* de intervalo de confianza para la diferencia entre las medias (puede especificarse el nivel de confianza). Desviación típica y error típico de la diferencia entre las medias.

Datos. Especifique dos variables cuantitativas (nivel de medida de intervalo o de razón) para cada prueba de pares. En un estudio de pares relacionados o de control de casos, la respuesta de cada sujeto de la prueba y su sujeto de control correspondiente deberán hallarse en el mismo caso en el archivo de datos.

Supuestos. Las observaciones de cada par deben hacerse en las mismas condiciones. Las diferencias entre las medias deben estar normalmente distribuidas. Las varianzas de cada variable pueden ser iguales o desiguales.

Figura 20-6

Resultados de la prueba *T* para muestras relacionadas

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	After treatment	116.40	10	13.62	4.31
	Before treatment	142.50	10	17.04	5.39

Prueba de muestras relacionadas

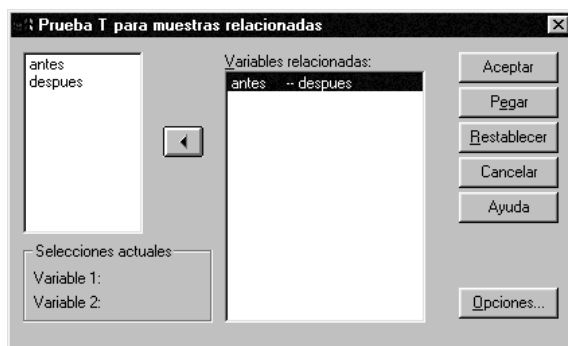
		Diferencias relacionadas				t	gl	Sig. (bilateral)
		Media	Desviación típ.	Error típ. de la media	Intervalo de confianza para la diferencia			
					Inferior			
Par 1	Antes del tratamiento - Después del tratamiento	26.10	19.59	6.19	12.09	4.214	9	.002

Para obtener una prueba T para muestras relacionadas

- ▶ Elija en los menús:
 - Analizar
 - Comparar medias
 - Prueba T para muestras relacionadas...

Figura 20-7

Cuadro de diálogo Prueba T para muestras relacionadas



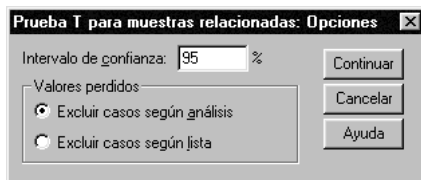
- ▶ Seleccione un par de variables como se indica a continuación:
 - Pulse en cada una de las dos variables. La primera variable aparecerá en la sección Selecciones actuales como *Variable 1* y la segunda aparecerá como *Variable 2*.
 - Una vez seleccionado el par de variables, pulse en el botón de flecha para moverlo a la lista Variables relacionadas. Si lo desea, pueden seleccionar más parejas de variables. Para eliminar una pareja de variables del análisis, selecciónela en la lista Variables relacionadas y pulse en el botón de flecha.

Si lo desea, puede pulsar en Opciones para controlar el tratamiento de los datos perdidos y el nivel del intervalo de confianza.

Prueba T para muestras relacionadas: Opciones

Figura 20-8

Cuadro de diálogo Prueba T para muestras relacionadas: Opciones



Intervalo de confianza. Por defecto se muestra un intervalo de confianza al 95% para la diferencia entre las medias. Introduzca un valor entre 1 y 99 para solicitar otro nivel de confianza.

Valores perdidos. Si ha probado varias variables y se han perdido los datos de una o más de ellas, puede indicar al procedimiento qué casos desea incluir (o excluir):

- **Excluir casos según análisis.** Cada prueba t utilizará todos los casos que contienen datos válidos para la pareja de variables contrastadas. Los tamaños muestrales pueden variar de una prueba a otra.
- **Excluir casos según lista.** Cada prueba t utilizará únicamente los casos que contengan datos válidos para todas las parejas de variables contrastadas. El tamaño muestral es constante en todas las pruebas.

Prueba T para una muestra

El procedimiento Prueba T para una muestra contrasta si la media de una sola variable difiere de una constante especificada.

Ejemplos. Un investigador desea comprobar si la puntuación media del coeficiente intelectual de un grupo de alumnos difiere de 100. O bien, un fabricante de copos de cereales puede tomar una muestra de envases de la línea de producción y comprobar si el peso medio de las muestras difiere de 1 Kg con un nivel de confianza al 95%.

Estadísticos. Para cada variable a contrastar: media, desviación típica y error típico de la media. La diferencia promedio entre cada valor de los datos y el valor de contraste hipotetizado, una prueba t que contrasta que esta diferencia es 0 y un intervalo de confianza para la diferencia promedio (para el que puede especificarse el nivel de confianza).

Datos. Para contrastar los valores de una variable cuantitativa con un valor de contraste hipotetizado, elija una variable cuantitativa e introduzca un valor de contraste hipotetizado.

Supuestos. Esta prueba asume que los datos están normalmente distribuidos; sin embargo, esta prueba es bastante robusto frente a las desviaciones de la normalidad.

Figura 20-9
Resultados de Prueba T para una muestra

One-Sample Statistics

	IQ
N	15
Mean	109.33
Std. Deviation	12.03
Std. Error Mean	3.11

Rows and columns have been transposed.

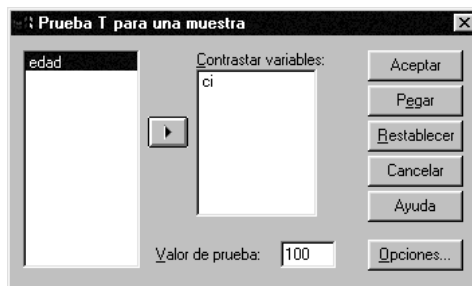
One-Sample Test

Test Value = 100						
	t	df	Significance (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
IQ	3.005	14	.009	9.33	2.67	15.99

Para obtener una prueba T para una muestra

- ▶ Elija en los menús:
Analizar
Comparar medias
Prueba T para una muestra...

Figura 20-10
Cuadro de diálogo Prueba T para una muestra



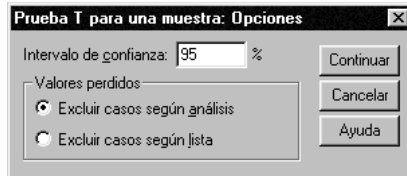
- ▶ Seleccione una o más variables para contrastarlas con el mismo valor hipotetizado.
- ▶ Introduzca un valor de contraste numérico para compararlo con cada media muestral.

Si lo desea, puede pulsar en Opciones para controlar el tratamiento de los datos perdidos y el nivel del intervalo de confianza.

Prueba T para una muestra: Opciones

Figura 20-11

Cuadro de diálogo Prueba T para una muestra: Opciones



Intervalo de confianza. Por defecto se muestra un intervalo de confianza al 95% para la diferencia entre la media y el valor de contraste hipotetizado. Introduzca un valor entre 1 y 99 para solicitar otro nivel de confianza.

Valores perdidos. Si desea contrastar varias variables y existen datos perdidos en una o en varias de las variables, podrá indicar al procedimiento qué casos deberán ser incluidos (o excluidos).

- **Excluir casos según análisis.** Cada prueba t utiliza todos los casos que tienen datos válidos para la variable contrastada. Los tamaños muestrales pueden variar de una prueba a otra.
- **Excluir casos según lista.** Cada prueba t utiliza sólo aquellos casos que contienen datos válidos para todas las variables utilizadas en cualquiera de las pruebas t solicitadas. El tamaño muestral es constante en todas las pruebas.

ANOVA de un factor

El procedimiento ANOVA de un factor genera un análisis de varianza de un factor para una variable dependiente cuantitativa respecto a una única variable de factor (la variable independiente). El análisis de varianza se utiliza para contrastar la hipótesis de que varias medias son iguales. Esta técnica es una extensión de la prueba t para dos muestras.

Además de determinar que existen diferencias entre las medias, es posible que desee saber qué medias difieren. Existen dos tipos de contrastes para comparar medias: a priori y post hoc. Los contrastes a priori se plantean *antes* de ejecutar el experimento y los contrastes post hoc se realizan *después* de haber llevado a cabo el experimento. También puede contrastar las tendencias existentes a través de las categorías.

Ejemplo. Las rosquillas absorben diferentes cantidades de grasa cuando se fríen. Se plantea un experimento utilizando tres tipos de grasas: aceite de cacahuete, aceite de maíz y manteca de cerdo. El aceite de cacahuete y el aceite de maíz son grasas no saturadas y la manteca es una grasa saturada. Además de determinar si la cantidad de grasa absorbida depende del tipo de grasa utilizada, también se podría preparar un contraste a priori para determinar si la cantidad de absorción de la grasa difiere para las grasas saturadas y las no saturadas.

Estadísticos. Para cada grupo: número de casos, media, desviación típica, error típico de la media, mínimo, máximo, intervalo de confianza al 95% para la media. Prueba de Levene sobre la homogeneidad de varianzas, tabla de análisis de varianza y contrastes robustos de igualdad de medias para cada variable dependiente, contrastes a priori especificados por el usuario y las pruebas de rango y de comparaciones múltiples post hoc: Bonferroni, Sidak, diferencia honestamente significativa de Tukey, GT2 de Hochberg, Gabriel, Dunnett, prueba F de Ryan-Einot-Gabriel-Welsch, (R-E-G-W F), prueba de rango de Ryan-Einot-Gabriel-Welsch (R-E-G-W Q), T2 de Tamhane, T3 de Dunnett, Games-Howell, C , de Dunnett, prueba de rango múltiple de Duncan,

Student-Newman-Keuls (S-N-K), *b* de Tukey, Waller-Duncan, Scheffé y diferencia menos significativa.

Datos. Los valores de la variable de factor deben ser enteros y la variable dependiente debe ser cuantitativa (nivel de medida de intervalo).

Supuestos. Cada grupo es una muestra aleatoria independiente procedente de una población normal. El análisis de varianza es robusto a las desviaciones de la normalidad, aunque los datos deberán ser simétricos. Los grupos deben proceder de poblaciones con varianzas iguales. Para contrastar este supuesto, utilice la prueba de Levene de homogeneidad de varianzas.

Figura 21-1
Resultados del procedimiento ANOVA de un factor

ANOVA

		Sumas de cuadrados	Gl	Media cuadrática	F	Sig.
Gramos de grasa absorbida	Entre Grupos	1596.000	2	798.000	7.824	.005
	Dentro de Grupos	1530.000	15	102.000		
	Total	3126.000	17			

Descriptivos

			N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95%		Mínimo	Máximo
							Límite inferior	Límite superior		
Gramos de grasa absorbida	Tipo de aceite	De Cacahuete	6	72.00	13.34	5.45	58.00	86.00	56	95
		Manteca	6	85.00	7.77	3.17	76.84	93.16	77	97
		De Maíz	6	62.00	8.22	3.36	53.37	70.63	49	70
		Total	18	73.00	13.56	3.20	66.26	79.74	49	97

Coefficientes de contraste

Contraste	Tipo de aceite		
	De Cacahuete	Manteca	De Maíz
1	-.5	1	-.5

Pruebas de contraste

		Contraste	Valor de contraste	Error típico	t	gl	Sig. (bilateral)
Gramos de grasa absorbida	Suponer igualdad de varianzas	1	18.00	5.05	3.565	15	.003
	No asume igualdad de varianzas	1	18.00	4.51	3.995	12.542	.002

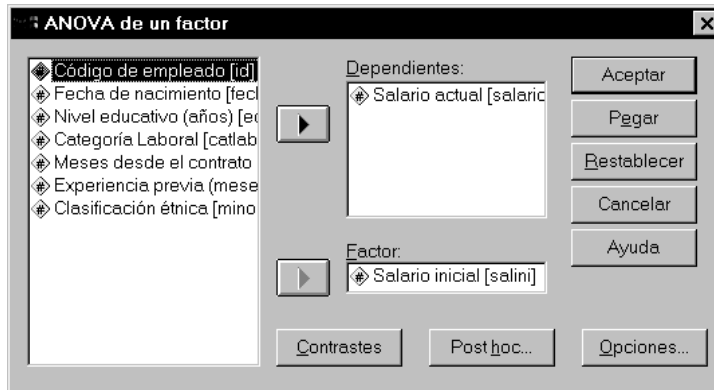
Prueba de homogeneidad de varianzas

	Estadístico de Levene	gl1	gl2	Sig.
Gramos de grasa absorbida	.534	2	15	.597

Para obtener un análisis de varianza de un factor

- ▶ Elija en los menús:
 Analizar
 Comparar medias
 ANOVA de un factor...

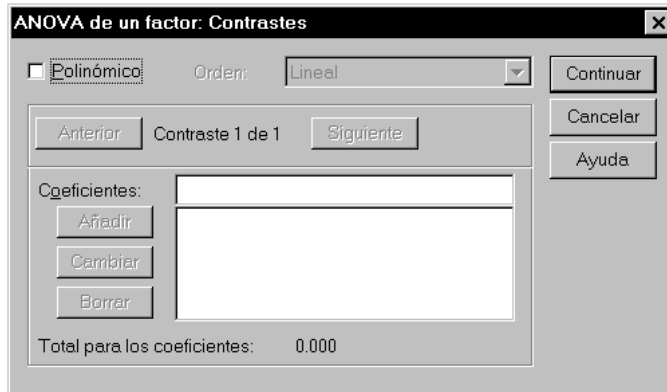
Figura 21-2
 Cuadro de diálogo ANOVA de un factor



- ▶ Seleccione una o más variables dependientes.
- ▶ Seleccione una sola variable de factor independiente.

ANOVA de un factor: Contrastes

Figura 21-3
Cuadro de diálogo ANOVA de un factor: Contrastes



Puede dividir las sumas de cuadrados inter-grupos en componentes de tendencia o especificar contrastes a priori.

Polinómico. Divide las sumas de cuadrados inter-grupos en componentes de tendencia. Puede contrastar la existencia de tendencia en la variable dependiente a través de los niveles ordenados de la variable de factor. Por ejemplo, podría contrastar si existe una tendencia lineal (creciente o decreciente) en el salario, a través de los niveles ordenados de la titulación mayor obtenida.

■ **Orden.** Se puede elegir un orden polinómico 1°, 2°, 3°, 4° o 5°.

Coeficientes. Contrastes a priori especificados por el usuario que serán contrastados mediante el estadístico t . Introduzca un coeficiente para cada grupo (categoría) de la variable factor y pulse en Añadir después de cada entrada. Cada nuevo valor se añade al final de la lista de coeficientes. Para especificar conjuntos de contrastes adicionales, pulse en Siguiente. Utilice Siguiente y Anterior para desplazarse por los conjuntos de contrastes.

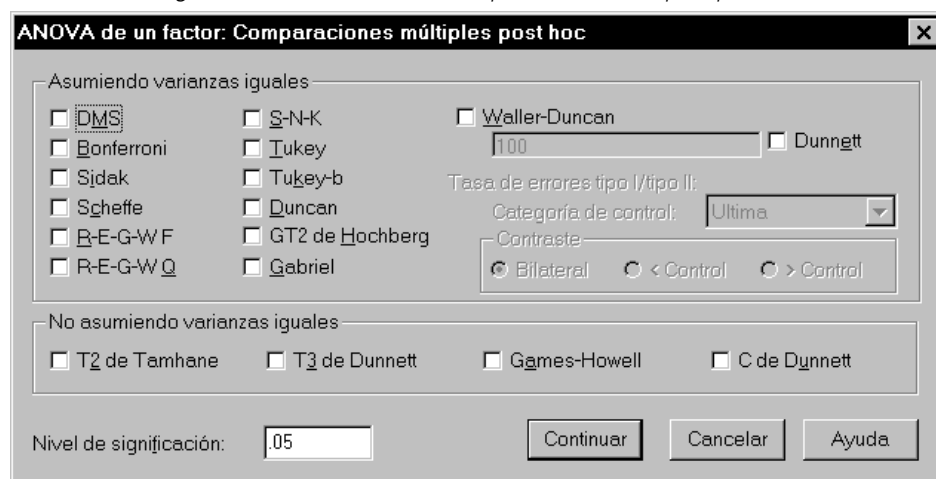
El orden de los coeficientes es importante porque se corresponde con el orden ascendente de los valores de las categorías de la variable de factor. El primer coeficiente en la lista se corresponde con el menor de los valores de grupo en la variable de factor y el último coeficiente se corresponde con el valor más alto. Por ejemplo, si existen seis categorías en la variable factor, los coeficientes $-1, 0, 0, 0, 0,5$ y $0,5$ contrastan el primer grupo con los grupos quinto y sexto. Para la mayoría de las

aplicaciones, la suma de los coeficientes debería ser 0. Los conjuntos que no sumen 0 también se pueden utilizar, pero aparecerá un mensaje de advertencia.

ANOVA de un factor: Contrastes post hoc

Figura 21-4

Cuadro de diálogo ANOVA de un factor: Comparaciones múltiples post hoc



Una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc y las comparaciones múltiples por parejas permiten determinar qué medias difieren. Las pruebas de rango identifican subconjuntos homogéneos de medias que no se diferencian entre sí. Las comparaciones múltiples por parejas contrastan la diferencia entre cada pareja de medias y dan lugar a una matriz donde los asteriscos indican las medias de grupo significativamente diferentes a un nivel alfa de 0,05.

Asumiendo varianzas iguales

La prueba de la diferencia honestamente significativa de Tukey, la GT2 de Hochberg, la prueba de Gabriel y la prueba de Scheffé son pruebas de comparaciones múltiples y pruebas de rango. Otras pruebas de rango disponibles son la *b* de Tukey, S-N-K (Student-Newman-Keuls), Duncan, R-E-G-W *F* (prueba *F* de Ryan-Einot-Gabriel-Welsch), R-E-G-W *Q* (prueba de rango de

Ryan-Einot-Gabriel-Welsch) y Waller-Duncan. Las pruebas de comparaciones múltiples disponibles son Bonferroni, Diferencia honestamente significativa de Tukey, Sidak, Gabriel, Hochberg, Dunnett, Scheffé y DMS (diferencia menos significativa).

- **LSD.** Utiliza pruebas t para llevar a cabo todas las comparaciones por pares entre las medias de los grupos. No se efectúa ninguna corrección de la tasa de error para el hecho de realizar múltiples comparaciones.
- **Bonferroni.** Utiliza las pruebas de t para realizar comparaciones por pares entre las medias de los grupos, pero controla la tasa de error global estableciendo que la tasa de error de cada prueba sea igual a la tasa de error por experimento dividida entre el número total de contrastes. Así, se corrige el nivel crítico por el hecho de que se están realizando múltiples comparaciones.
- **Sidak.** Prueba de comparaciones múltiples por parejas basada en un estadístico t. La prueba de Sidak corrige el nivel de significación para las comparaciones múltiples y da lugar a límites más estrechos que los de Bonferroni.
- **Scheffé.** Realiza comparaciones múltiples conjuntas por parejas para todas las parejas de combinaciones de las medias posibles. Utiliza la distribución muestral F. Puede utilizarse para examinar todas las combinaciones lineales de grupos de medias posibles, no sólo las comparaciones por parejas.
- **R-E-G-W F.** Procedimiento múltiple por pasos (por tamaño de las distancias) de Ryan-Einot-Gabriel-Welsch que se basa en una prueba F.
- **R-E-G-W Q.** Procedimiento múltiple por pasos (por tamaño de las distancias) de Ryan-Einot-Gabriel-Welsch que se basa en el rango estudentizado.
- **S-N-K.** Realiza todas las comparaciones por parejas entre las medias utilizando la distribución del rango de Student. Con tamaños de grupo iguales, también compara pares de medias dentro de subconjuntos homogéneos utilizando un procedimiento por pasos. Las medias se ordenan de mayor a menor y se comparan primero las diferencias más extremas.
- **Tukey.** Utiliza el estadístico del rango estudentizado para realizar todas las comparaciones por pares entre los grupos. Establece la tasa de error por experimento como la tasa de error para el conjunto de todas las comparaciones por pares.
- **Tukey-b.** Prueba que emplea la distribución del rango estudentizado para realizar comparaciones por pares entre los grupos. El valor crítico es el promedio de los valores correspondientes a la diferencia honestamente significativa de Tukey y al método de Student-Newman-Keuls.

- **Duncan.** Realiza comparaciones por pares utilizando un orden por pasos idéntico al orden usado por la prueba de Student-Newman-Keuls, pero establece un nivel de protección en la tasa de error para la colección de contrastes, en lugar de usar una tasa de error para los contrastes individuales. Utiliza el estadístico del rango estudentizado.
- **GT2 de Hochberg.** Prueba de comparaciones múltiples y de rango que utiliza el módulo máximo estudentizado. Es similar a la prueba de la diferencia honestamente significativa de Tukey.
- **Gabriel.** Prueba de comparación por parejas que utiliza el módulo máximo estudentizado y que es generalmente más potente que la GT2 de Hochberg, cuando los tamaños de las casillas son desiguales. La prueba de Gabriel puede volverse liberal cuando los tamaños de las casillas varían mucho.
- **Waller-Duncan.** Prueba de comparaciones múltiples basada en un estadístico t . Utiliza la aproximación Bayesiana.
- **Dunnnett.** Prueba de comparaciones múltiples por parejas que compara un conjunto de tratamientos respecto a una única media de control. La última categoría es la categoría de control por defecto. Si lo desea, puede seleccionar la primera categoría. Para comprobar que la media de cualquier nivel del factor (excepto la categoría de control) no es igual a la de la categoría de control, utilice una prueba bilateral. Para contrastar si la media en cualquier nivel del factor es menor que la de la categoría de control, seleccione $<Control$. Para contrastar si la media en cualquier nivel del factor es mayor que la de la categoría de control, seleccione $>Control$.

No asumiendo varianzas iguales

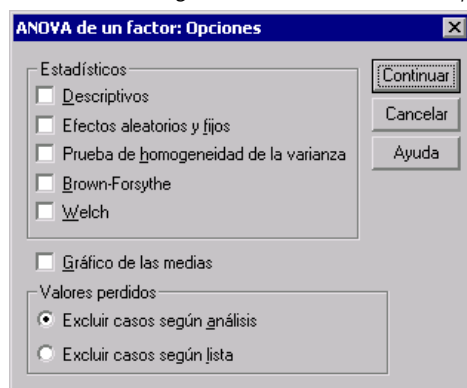
Las pruebas de comparaciones múltiples que no suponen varianzas iguales son T2 de Tamhane, T3 de Dunnnett, Games-Howell y C de Dunnnett.

- **T2 de Tamhane.** Prueba conservadora de comparación por parejas basada en la prueba t . Esta prueba es adecuada cuando las varianzas son desiguales.
- **T3 de Dunnnett.** Prueba de comparación por parejas basada en el módulo máximo estudentizado. Esta prueba es adecuada cuando las varianzas son desiguales.
- **Games-Howell.** Prueba de comparación por parejas que es en ocasiones liberal. Esta prueba es adecuada cuando las varianzas son desiguales.
- **C de Dunnnett.** Prueba de comparación por parejas basada en el rango estudentizado. Esta prueba es adecuada cuando las varianzas son desiguales.

Nota: Posiblemente le resulte más fácil interpretar el resultado de los contrastes post hoc si desactiva Ocultar filas y columnas vacías en el cuadro de diálogo Propiedades de tabla (en una tabla pivote activada, seleccione Propiedades de tabla en el menú Formato).

ANOVA de un factor: Opciones

Figura 21-5
Cuadro de diálogo ANOVA de un factor: Opciones



Estadísticos. Elija uno o más entre los siguientes:

- **Descriptivos.** Calcula los siguientes estadísticos: Número de casos, Media, Desviación típica, Error típico de la media, Mínimo, Máximo y los Intervalos de confianza al 95% de cada variable dependiente para cada grupo.
- **Efectos fijos y aleatorios.** Muestra la desviación típica, el error típico y un intervalo de confianza del 95% para el modelo de efectos fijos, y el error típico, un intervalo de confianza del 95% y una estimación de la varianza entre componentes para el modelo de efectos aleatorios.
- **Prueba de homogeneidad de las varianzas.** Calcula el estadístico de Levene para contrastar la igualdad de las varianzas de grupo. Esta prueba no depende del supuesto de normalidad.

- **Brown-Forsythe.** Calcula el estadístico de Brown-Forsythe para contrastar la igualdad de las medias de grupo. Este estadístico es preferible al estadístico F si no se supone la igualdad de las varianzas.
- **Welch.** Calcula el estadístico de Welch para contrastar la igualdad de las medias de grupo. Este estadístico es preferible al estadístico F si no se supone la igualdad de las varianzas.

Gráfico de las medias. Muestra un gráfico que representa las medias de los subgrupos (las medias para cada grupo definido por los valores de la variable factor).

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según análisis.** Un caso que tenga un valor perdido para la variable dependiente o la variable de factor en un análisis determinado, no se utiliza en ese análisis. Además, los casos fuera del rango especificado para la variable de factor no se utilizan.
- **Excluir casos según lista.** Se excluyen de todos los análisis los casos con valores perdidos para la variable de factor o para cualquier variable dependiente incluida en la lista de variables dependientes en el cuadro de diálogo principal. Si no se han especificado varias variables dependientes, esta opción no surte efecto.

MLG Análisis univariante

El procedimiento MLG Univariante proporciona un análisis de regresión y un análisis de varianza para una variable dependiente mediante uno o más factores o variables. Las variables de factor dividen la población en grupos. Con el procedimiento Modelo lineal general se pueden contrastar hipótesis nulas sobre los efectos de otras variables en las medias de varias agrupaciones de una única variable dependiente. Se pueden investigar las interacciones entre los factores así como los efectos de los factores individuales, algunos de los cuales pueden ser aleatorios. Además, se pueden incluir los efectos de las covariables y las interacciones de covariables con los factores. Para el análisis de regresión, las variables (predictoras) independientes se especifican como covariables.

Se pueden contrastar tanto los modelos equilibrados como los no equilibrados. Se considera que un diseño está equilibrado si cada casilla del modelo contiene el mismo número de casos. Además de contrastar hipótesis, MLG Univariante genera estimaciones de los parámetros.

También se encuentran disponibles los contrastes a priori de uso más habitual para contrastar las hipótesis. Además, si una prueba F global ha mostrado cierta significación, pueden emplearse las pruebas post hoc para evaluar las diferencias entre las medias específicas. Las medias marginales estimadas ofrecen estimaciones de valores de las medias pronosticados para las casillas del modelo; los gráficos de perfil (gráficos de interacciones) de estas medias permiten observar fácilmente algunas de estas relaciones.

En su archivo de datos puede guardar residuos, valores pronosticados, distancia de Cook y valores de influencia como variables nuevas para comprobar los supuestos.

Ponderación MCP permite especificar una variable usada para aplicar a las observaciones una ponderación diferente en un análisis de mínimos cuadrados ponderados (MCP), por ejemplo para compensar la distinta precisión de las medidas.

Ejemplo. Se recogen datos de los corredores individuales en el maratón de Chicago durante varios años. El tiempo final de cada corredor es la variable dependiente. Influyen otros factores como el clima (frío, calor o temperatura agradable), los meses de entrenamiento, el número de maratones anteriores y el sexo. La edad se considera una covariable. Observará que el sexo es un efecto significativo y que la interacción del sexo con el clima es significativa.

Métodos. Las sumas de cuadrados de Tipo I, Tipo II, Tipo III y Tipo IV pueden emplearse para evaluar las diferentes hipótesis. Tipo III es el valor por defecto.

Estadísticos. Las pruebas de rango post hoc y las comparaciones múltiples: Diferencia menos significativa (DMS), Bonferroni, Sidak, Scheffé, Múltiples F de Ryan-Einot-Gabriel-Welsch (R-E-G-W-F), Rango múltiple de Ryan-Einot-Gabriel-Welsch, Student-Newman-Keuls (S-N-K), Diferencia honestamente significativa de Tukey, b de Tukey, Duncan, GT2 de Hochberg, Gabriel, Pruebas t de Waller Duncan, Dunnett (unilateral y bilateral), T2 de Tamhane, T3 de Dunnett, Games-Howell y C de Dunnett. Estadísticos descriptivos: medias observadas, desviaciones típicas y frecuencias de todas las variables dependientes en todas las casillas. Prueba de Levene para la homogeneidad de varianzas.

Diagramas. Diagramas de dispersión por nivel, gráficos de residuos, gráficos de perfil (interacción).

Datos. La variable dependiente es cuantitativa. Los factores son categóricos; pueden tener valores numéricos o valores de cadena de hasta ocho caracteres. Pueden tener valores numéricos o valores de cadena de hasta ocho caracteres. Las covariables son variables cuantitativas que están relacionadas con la variable dependiente.

Supuestos. Los datos son una muestra aleatoria de una población normal; en la población, todas las varianzas de las casillas son iguales. El análisis de varianza es robusto a las desviaciones de la normalidad, aunque los datos deberán ser simétricos. Para comprobar los supuestos, puede utilizar la prueba de homogeneidad de varianzas y los gráficos de dispersión por nivel. También puede examinar los residuos y los gráficos de residuos.

Figura 22-1
Resultados de MLG Univariante
Pruebas de Efectos inter-sujetos

Variable dependiente: SPVOL

Origen	Sumas de cuadrados tipo III	Gl	Media cuadrática	F	Sig.
Modelo correcto	22.520	11	2.047	12.376	.000
Intersección	1016.981	1	1016.981	6147.938	.000
Harina	8.691	3	2.897	17.513	.000
Grasa	10.118	2	5.059	30.583	.000
Surfactante	.997	2	.499	3.014	.082
Grasa* surfactante	5.639	4	1.410	8.522	.001
Error	2.316	14	.165		
Total	1112.960	26			
Total corregido	24.835	25			

1. R-cuadrado = .907 (R-cuadrado ajustado= .833)

Grasa* surfactante

Variable dependiente: SPVOL

Grasa	Surfactante	Media	Std. Error	Intervalo de confianza al 95%	
				Límite inferior	Límite superior
1	1	5.536	.240	5.021	6.052
	2	5.891	.239	5.378	6.404
	3	6.123	.241	5.605	6.641
2	1	7.023	.241	6.505	7.541
	2	6.708	.301	6.064	7.353
	3	6.000	.203	5.564	6.436
3	1	6.629	.301	5.984	7.274
	2	7.200	.203	6.764	7.636
	3	8.589	.300	7.945	9.233

Para obtener un análisis MLG Univariante

- ▶ Elija en los menús:
 Analizar
 Modelo lineal general
 Univariante...

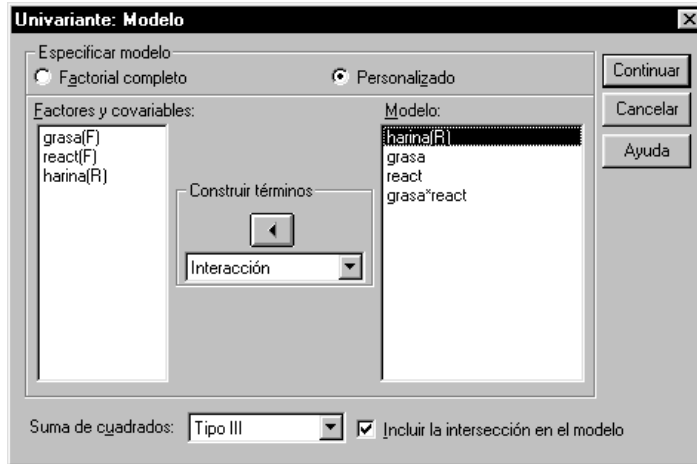
Figura 22-2
Cuadro de diálogo Univariante



- ▶ Seleccione una variable dependiente.
- ▶ Seleccione variables para Factores fijos, Factores aleatorios y Covariables, en función de los datos.
- ▶ Si lo desea, puede utilizar la Ponderación MCP para especificar una variable de ponderación para el análisis de mínimos cuadrados ponderados. Si el valor de la variable de ponderación es cero, negativo o perdido, el caso queda excluido del análisis. Una variable que ya se haya utilizado en el modelo no puede usarse como variable de ponderación.

MLG: Modelo

Figura 22-3
Cuadro de diálogo Univariante: Modelo



Especificar modelo. Un modelo factorial completo contiene todos los efectos principales del factor, todos los efectos principales de las covariables y todas las interacciones factor por factor. No contiene interacciones de covariable. Seleccione Personalizado para especificar sólo un subconjunto de interacciones o para especificar interacciones factor por covariable. Indique todos los términos que desee incluir en el modelo.

Factores y Covariables. Muestra una lista de los factores y las covariables, etiquetando con (F) los factores fijos y con (C) las covariables. En un análisis univariante, (R) indica un factor aleatorio.

Modelo. El modelo depende de la naturaleza de los datos. Después de seleccionar Personalizado, puede elegir los efectos principales y las interacciones que sean de interés para el análisis.

Suma de cuadrados Determina el método para calcular las sumas de cuadrados. Para los modelos equilibrados y no equilibrados sin casillas perdidas, el método más utilizado para la suma de cuadrados es el Tipo III.

Incluir la intersección en el modelo. La intersección se incluye normalmente en el modelo. Si supone que los datos pasan por el origen, puede excluir la intersección.

Construir términos

Para las covariables y los factores seleccionados:

Interacción. Crea el término de interacción de mayor nivel con todas las variables seleccionadas. Este es el método por defecto.

Efectos principales. Crea un término de efectos principales para cada variable seleccionada.

Todas de 2. Crea todas las interacciones dobles posibles de las variables seleccionadas.

Todas de 3. Crea todas las interacciones triples posibles de las variables seleccionadas.

Todas de 4. Crea todas las interacciones cuádruples posibles de las variables seleccionadas.

Todas de 5. Crea todas las interacciones quíntuples posibles de las variables seleccionadas.

Suma de cuadrados

Para el modelo, puede elegir un tipo de suma de cuadrados. El Tipo III es el más utilizado y es el tipo por defecto.

Tipo I. Este método también se conoce como el método de descomposición jerárquica de la suma de cuadrados. Cada término se corrige sólo respecto al término que le precede en el modelo. El método Tipo I para la obtención de sumas de cuadrados se utiliza normalmente para:

- Un modelo ANOVA equilibrado en el que se especifica cualquier efecto principal antes de cualquier efecto de interacción de primer orden, cualquier efecto de interacción de primer orden se especifica antes de cualquier efecto de interacción de segundo orden, y así sucesivamente.
- Un modelo de regresión polinómica en el que se especifica cualquier término de orden inferior antes que cualquier término de orden superior.
- Un modelo puramente anidado en el que el primer efecto especificado está anidado dentro del segundo efecto especificado, el segundo efecto especificado está anidado dentro del tercero, y así sucesivamente. Esta forma de anidamiento solamente puede especificarse utilizando la sintaxis.

Tipo II. Este método calcula cada suma de cuadrados del modelo considerando sólo los efectos pertinentes. Un efecto pertinente es el que corresponde a todos los efectos que no contienen el que se está examinando. El método Tipo II para la obtención de sumas de cuadrados se utiliza normalmente para:

- Un modelo ANOVA equilibrado.
- Cualquier modelo que sólo tenga efectos de factor principal.
- Cualquier modelo de regresión.
- Un diseño puramente anidado (esta forma de anidamiento solamente puede especificarse utilizando la sintaxis).

Tipo III. Es el método por defecto. Este método calcula las sumas de cuadrados de un efecto del diseño como las sumas de cuadrados corregidas respecto a cualquier otro efecto que no lo contenga y ortogonales a cualquier efecto (si existe) que lo contenga. Las sumas de cuadrados de Tipo III tienen una gran ventaja por ser invariables respecto a las frecuencias de casilla, siempre que la forma general de estimabilidad permanezca constante. Así, este tipo de sumas de cuadrados se suele considerar de gran utilidad para un modelo no equilibrado sin casillas perdidas. En un diseño factorial sin casillas perdidas, este método equivale a la técnica de cuadrados ponderados de las medias de Yates. El método Tipo III para la obtención de sumas de cuadrados se utiliza normalmente para:

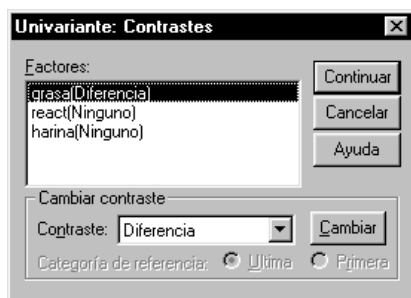
- Cualquiera de los modelos que aparecen en los tipos I y II.
- Cualquier modelo equilibrado o desequilibrado sin casillas vacías.

Tipo IV. Este método está diseñado para una situación en la que hay casillas perdidas. Para cualquier efecto F en el diseño, si F no está contenida en cualquier otro efecto, entonces Tipo IV = Tipo III = Tipo II. Cuando F está contenida en otros efectos, el Tipo IV distribuye equitativamente los contrastes que se realizan entre los parámetros en F a todos los efectos de nivel superior. El método Tipo IV para la obtención de sumas de cuadrados se utiliza normalmente para:

- Cualquiera de los modelos que aparecen en los tipos I y II.
- Cualquier modelo equilibrado o no equilibrado con casillas vacías.

MLG: Contrastes

Figura 22-4
Cuadro de diálogo Univariante: Contrastes



Los contrastes se utilizan para contrastar las diferencias entre los niveles de un factor. Puede especificar un contraste para cada factor en el modelo (en un modelo de medidas repetidas, para cada factor inter-sujetos). Los contrastes representan las combinaciones lineales de los parámetros.

El contraste de hipótesis se basa en la hipótesis nula $\mathbf{LB} = 0$, donde \mathbf{L} es la matriz de coeficientes de contraste y \mathbf{B} es el vector de parámetros. Cuando se especifica un contraste, SPSS crea una matriz \mathbf{L} en la que las columnas correspondientes al factor coinciden con el contraste. El resto de las columnas se corrigen para que la matriz \mathbf{L} sea estimable.

Los resultados incluyen un estadístico F para cada conjunto de contrastes. Para el contraste de diferencias también se muestran los intervalos de confianza simultáneos de tipo Bonferroni basados en la distribución t de Student.

Contrastes disponibles

Los contrastes disponibles son de desviación, simples, de diferencias, de Helmert, repetidos y polinómicos. En los contrastes de desviación y los contrastes simples, es posible determinar que la categoría de referencia sea la primera o la última categoría.

Tipos de contrastes

Desviación. Compara la media de cada nivel (excepto una categoría de referencia) con la media de todos los niveles (media global). Los niveles del factor pueden colocarse en cualquier orden.

Simple. Compara la media de cada nivel con la media de un nivel especificado. Este tipo de contraste resulta útil cuando existe un grupo de control. Puede seleccionar la primera o la última categoría como referencia.

Diferencia. Compara la media de cada nivel (excepto el primero) con la media de los niveles anteriores (a veces también se denominan contrastes de Helmert inversos). (a veces también se denominan contrastes de Helmert inversos).

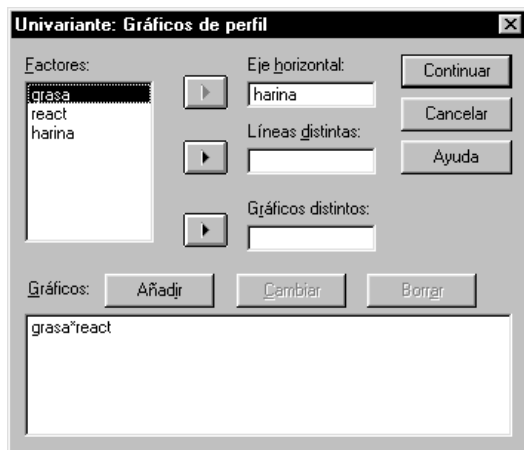
Helmert. Compara la media de cada nivel del factor (excepto el último) con la media de los niveles siguientes.

Repetidas. Compara la media de cada nivel (excepto el último) con la media del nivel siguiente.

Polinómico. Compara el efecto lineal, cuadrático, cúbico, etc. El primer grado de libertad contiene el efecto lineal a través de todas las categorías; el segundo grado de libertad, el efecto cuadrático, y así sucesivamente. Estos contrastes se utilizan a menudo para estimar las tendencias polinómicas.

MLG: Gráficos de perfil

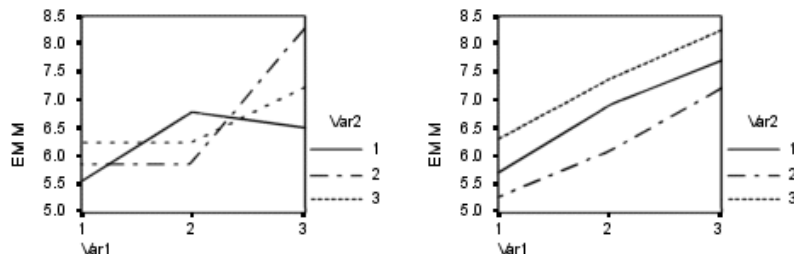
Figura 22-5
Cuadro de diálogo Univariante: Gráficos de perfil



Los gráficos de perfil (gráficos de interacción) sirven para comparar las medias marginales en el modelo. Un gráfico de perfil es un gráfico de líneas en el que cada punto indica la media marginal estimada de una variable dependiente (corregida respecto a las covariables) en un nivel de un factor. Los niveles de un segundo factor se pueden utilizar para generar líneas diferentes. Cada nivel en un tercer factor se puede utilizar para crear un gráfico diferente. Todos los factores fijos y aleatorios, si existen, están disponibles para los gráficos. Para los análisis multivariantes, los gráficos de perfil se crean para cada variable dependiente. En un análisis de medidas repetidas, es posible utilizar tanto los factores inter-sujetos como los intra-sujetos en los gráficos de perfil. Las opciones MLG - Multivariante y MLG - Medidas repetidas sólo estarán disponibles si tiene instalada la opción Modelos avanzados.

Un gráfico de perfil de un factor muestra si las medias marginales estimadas aumentan o disminuyen a través de los niveles. Para dos o más factores, las líneas paralelas indican que no existe interacción entre los factores, lo que significa que puede investigar los niveles de un único factor. Las líneas no paralelas indican una interacción.

Figura 22-6
Gráfico no paralelo (izquierda) y gráfico paralelo (derecha)



Después de especificar un gráfico mediante la selección de los factores del eje horizontal y, de manera opcional, los factores para distintas líneas y gráficos, el gráfico deberá añadirse a la lista de gráficos.

MLG: Comparaciones post hoc

Figura 22-7
Cuadro de diálogo Univariante: Comparaciones múltiples post hoc para las medias observadas

Univariante: Comparaciones múltiples post hoc para las medias observadas

Factores:
grasa
react

Contrastes post hoc para:
grasa

Continuar
Cancelar
Ayuda

Asumiendo varianzas iguales:

- DMS
- Bonferroni
- Sidak
- Scheffe
- R-E-G-W F
- R-E-G-W Q
- S-N-K
- Tukey
- Tukey-b
- Duncan
- GT2 de Hochberg
- Gabriel
- Waller-Duncan
- Dunnett

Tasa de errores tipo I/tipo II: 100

Categoría de control: Ultima

Contraste:
 Bijateral < Control > Control

No asumiendo varianzas iguales:

- T₂ de Tamhane
- T₃ de Dunnett
- Games-Howell
- C de Dunnett

Pruebas de comparaciones múltiples post hoc Una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc y las comparaciones múltiples por parejas permiten determinar qué medias difieren. Las comparaciones se

realizan sobre valores sin corregir. Estas pruebas se utilizan únicamente para factores inter-sujetos fijos. En MLG Medidas repetidas, estas pruebas no están disponibles si no existen factores inter-sujetos y las pruebas de comparación múltiple post hoc se realizan para la media a través de los niveles de los factores intra-sujetos. Para MLG - Multivariante, las pruebas post hoc se realizan por separado para cada variable dependiente. Las opciones MLG - Multivariante y MLG - Medidas repetidas sólo estarán disponibles si tiene instalada la opción Modelos avanzados.

Las pruebas de diferencia honestamente significativa de Tukey y de Bonferroni son pruebas de comparación múltiple muy utilizadas. La **prueba de Bonferroni**, basada en el estadístico t de Student, corrige el nivel de significación observado por el hecho de que se realizan comparaciones múltiples. La **prueba t de Sidak** también corrige el nivel de significación y da lugar a límites más estrechos que los de Bonferroni. La **prueba de diferencia honestamente significativa de Tukey** utiliza el estadístico del rango estudentizado para realizar todas las comparaciones por pares entre los grupos y establece la tasa de error por experimento como la tasa de error para el conjunto de todas las comparaciones por pares. Cuando se contrasta un gran número de pares de medias, la prueba de la diferencia honestamente significativa de Tukey es más potente que la prueba de Bonferroni. Para un número reducido de pares, Bonferroni es más potente.

GT2 de Hochberg es similar a la prueba de la diferencia honestamente significativa de Tukey, pero se utiliza el módulo máximo estudentizado. La prueba de Tukey suele ser más potente. La **prueba de comparación por parejas de Gabriel** también utiliza el módulo máximo estudentizado y es generalmente más potente que la GT2 de Hochberg cuando los tamaños de las casillas son desiguales. La prueba de Gabriel se puede convertir en liberal cuando los tamaños de las casillas varían mucho.

La **prueba t de comparación múltiple por parejas de Dunnett** compara un conjunto de tratamientos con una media de control simple. La última categoría es la categoría de control por defecto. Si lo desea, puede seleccionar la primera categoría. Asimismo, puede elegir una prueba unilateral o bilateral. Para comprobar que la media de cualquier nivel del factor (excepto la categoría de control) no es igual a la de la categoría de control, utilice una prueba bilateral. Para contrastar si la media en cualquier nivel del factor es menor que la de la categoría de control, seleccione $<$ Control. Asimismo, para contrastar si la media en cualquier nivel del factor es mayor que la de la categoría de control, seleccione $>$ Control.

Ryan, Einot, Gabriel y Welsch (R-E-G-W) desarrollaron dos pruebas de rangos múltiples por pasos. Los procedimientos múltiples por pasos (por tamaño de las distancias) contrastan en primer lugar si todas las medias son iguales. Si no son

iguales, se contrasta la igualdad en los subconjuntos de medias. **R-E-G-W F** se basa en una prueba F y **R-E-G-W Q** se basa en un rango estudentizado. Estas pruebas son más potentes que la prueba de rangos múltiples de Duncan y Student-Newman-Keuls (que también son procedimientos múltiples por pasos), pero no se recomiendan para tamaños de casillas desiguales.

Cuando las varianzas son desiguales, utilice **T2 de Tamhane** (prueba conservadora de comparación por parejas basada en una prueba t), **T3 de Dunnett** (prueba de comparación por parejas basada en el módulo máximo estudentizado), **prueba de comparación por parejas Games-Howell** (a veces liberal), o **C de Dunnett** (prueba de comparación por parejas basada en el rango estudentizado).

La **prueba de rango múltiple de Duncan**, Student-Newman-Keuls (**S-N-K**) y **b de Tukey** son pruebas de rango que asignan rangos a medias de grupo y calculan un valor de rango. Estas pruebas no se utilizan con la misma frecuencia que las pruebas anteriormente mencionadas.

La **prueba t de Waller-Duncan** utiliza la aproximación bayesiana. Esta prueba de rango emplea la media armónica del tamaño muestral cuando los tamaños muestrales no son iguales.

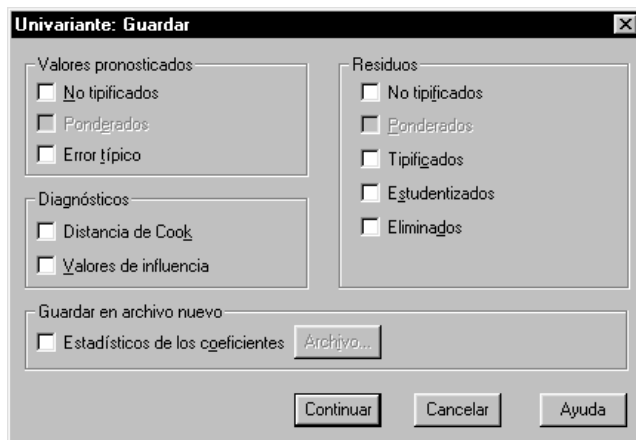
El nivel de significación de la prueba de **Scheffé** está diseñado para permitir todas las combinaciones lineales posibles de las medias de grupo que se van a contrastar, no sólo las comparaciones por parejas disponibles en esta función. El resultado es que la prueba de Scheffé es normalmente más conservadora que otras pruebas, lo que significa que se precisa una mayor diferencia entre las medias para la significación.

La prueba de comparación múltiple por parejas de la diferencia menos significativa (**DMS**) es equivalente a varias pruebas t individuales entre todos los pares de grupos. La desventaja de esta prueba es que no se realiza ningún intento de corregir el nivel crítico para realizar las comparaciones múltiples.

Pruebas mostradas. Se proporcionan comparaciones por parejas para DMS, Sidak, Bonferroni, Games y Howell, T2 y T3 de Tamhane, C de Dunnett y T3 de Dunnett. También se facilitan subconjuntos homogéneos para S-N-K, b de Tukey, Duncan, R-E-G-W F, R-E-G-W Q y Waller. La prueba de la diferencia honestamente significativa de Tukey, la GT2 de Hochberg, la prueba de Gabriel y la prueba de Scheffé son pruebas de comparaciones múltiples y pruebas de rango.

MLG: Guardar

Figura 22-8
Cuadro de diálogo Univariante: Guardar



Es posible guardar los valores pronosticados por el modelo, los residuos y las medidas relacionadas como variables nuevas en el Editor de datos. Muchas de estas variables se pueden utilizar para examinar supuestos sobre los datos. Si desea almacenar los valores para utilizarlos en otra sesión de SPSS, guárdelos en el archivo de datos actual.

Valores pronosticados. Son los valores que predice el modelo para cada caso.

- **No tipificados.** Valor pronosticado por el modelo para la variable dependiente.
- **Ponderados.** Los valores pronosticados no tipificados ponderados. Sólo están disponibles si se seleccionó previamente una variable de ponderación WLS.
- **Error típico.** Estimación de la desviación típica del valor promedio de la variable dependiente para los casos que tengan los mismos valores en las variables independientes.

Diagnósticos. Son medidas para identificar casos con combinaciones poco usuales de valores para los casos y las variables independientes que puedan tener un gran impacto en el modelo.

- **Distancia de Cook.** Medida de cuánto cambiarían los residuos de todos los casos si se excluyera un caso determinado del cálculo de los coeficientes de regresión. Una Distancia de Cook grande indica que la exclusión de ese caso del cálculo de los estadísticos de regresión hará variar substancialmente los coeficientes.
- **Valores de influencia.** Los valores de influencia no centrados. La influencia relativa de cada observación sobre el ajuste del modelo.

Residuos. Un residuo no tipificado es el valor real de la variable dependiente menos el valor pronosticado por el modelo. También se encuentran disponibles residuos eliminados, estudentizados y tipificados. Si ha seleccionado una variable MCP, contará además con residuos no tipificados ponderados.

- **No tipificados.** Diferencia entre un valor observado y el valor pronosticado por el modelo.
- **Ponderados.** Los residuos no tipificados ponderados. Sólo están disponibles si se seleccionó previamente una variable de ponderación WLS.
- **Tipificados.** El residuo dividido por una estimación de su error típico. Los residuos tipificados, que son conocidos también como los residuos de Pearson o residuos estandarizados, tienen una media de 0 y una desviación típica de 1.
- **Método de Student.** Residuo dividido por una estimación de su desviación típica que varía de caso en caso, dependiendo de la distancia de los valores de cada caso en las variables independientes respecto a las medias en las variables independientes.
- **Eliminados.** Residuo para un caso cuando ese caso se excluye del cálculo de los coeficientes de regresión. Es la diferencia entre el valor de la variable dependiente y el valor pronosticado corregido.

Guardar en archivo nuevo. Graba un archivo de datos de SPSS que contiene una matriz de varianzas-covarianzas de las estimaciones de los parámetros del modelo. Asimismo, para cada variable dependiente habrá una fila de estimaciones de los parámetros, una fila de valores de significación para los estadísticos t correspondientes a las estimaciones de los parámetros y una fila de grados de libertad de los residuos. En un modelo multivariante, existen filas similares para cada variable dependiente. Si lo desea, puede usar este archivo matricial en otros procedimientos que lean un archivo matricial de SPSS.

Opciones MLG

Figura 22-9
Cuadro de diálogo Univariante: Opciones



Este cuadro de diálogo contiene estadísticos opcionales. Los estadísticos se calculan utilizando un modelo de efectos fijos.

Medias marginales estimadas. Seleccione los factores e interacciones para los que desee obtener estimaciones de las medias marginales de la población en las casillas. Estas medias se corrigen respecto a las covariables, si las hay.

- **Comparar los efectos principales.** Proporciona comparaciones por parejas no corregidas entre las medias marginales estimadas para cualquier efecto principal del modelo, tanto para los factores inter-sujetos como para los intra-sujetos. Este elemento sólo se encuentra disponible si los efectos principales están seleccionados en la lista Mostrar las medias para.
- **Ajuste del intervalo de confianza.** Seleccione un ajuste de diferencia menor significativa (DMS), Bonferroni o Sidak para los intervalos de confianza y la significación. Este elemento sólo estará disponible si se selecciona Comparar los efectos principales.

Mostrar. Seleccione Estadísticos descriptivos para generar medias observadas, desviaciones típicas y frecuencias para cada variable dependiente en todas las casillas. La opción Estimaciones del tamaño del efecto ofrece un valor parcial de eta-cuadrado para cada efecto y cada estimación de parámetros. El estadístico eta cuadrado describe la proporción de variabilidad total atribuible a un factor. Seleccione Potencia observada para obtener la potencia de la prueba cuando la hipótesis alternativa se ha establecido basándose en el valor observado. Seleccione Estimaciones de los parámetros para generar las estimaciones de los parámetros, los errores típicos, las pruebas t , los intervalos de confianza y la potencia observada para cada prueba. Seleccione Matriz de coeficientes de contraste para obtener la matriz L .

Las pruebas de homogeneidad producen la prueba de homogeneidad de varianzas de Levene para cada variable dependiente en todas las combinaciones de nivel de los factores inter-sujetos sólo para factores inter-sujetos. Las opciones de diagramas de dispersión por nivel y gráfico de los residuos son útiles para comprobar los supuestos sobre los datos. Estos elementos no estarán activado si no hay factores. Seleccione Gráficos de los residuos para generar un gráfico de los residuos observados respecto a los pronosticados respecto a los tipificados para cada variable dependiente. Estos gráficos son útiles para investigar el supuesto de varianzas iguales. Seleccione Falta de ajuste para comprobar si el modelo puede describir de forma adecuada la relación entre la variable dependiente y las variables independientes. La función estimable general permite construir pruebas de hipótesis personales basadas en la función estimable general. Las filas en las matrices de coeficientes de contraste son combinaciones lineales de la función estimable general.

Nivel de significación. Puede que le interese corregir el nivel de significación usado en las pruebas post hoc y el nivel de confianza empleado para construir intervalos de confianza. El valor especificado también se utiliza para calcular la potencia observada para la prueba. Si especifica un nivel de significación, el cuadro de diálogo mostrará el nivel asociado de los intervalos de confianza.

Funciones adicionales de los comandos UNIANOVA

El lenguaje de comandos de SPSS también permite:

- Especificar efectos anidados en el diseño (utilizando el subcomando DESIGN).
- Especificar contrastes de los efectos respecto a una combinación lineal de efectos o un valor (utilizando el subcomando TEST).

- Especificar contrastes múltiples (utilizando el subcomando `CONTRAST`).
- Incluir los valores perdidos definidos por el usuario (utilizando el subcomando `MISSING`).
- Especificar criterios EPS (mediante el subcomando `CRITERIA`).
- Construir una matriz **L**, una matriz **M** o una matriz **K** (utilizando los subcomandos `LMATRIX`, `MMATRIX` y `KMATRIX`).
- Especificar una categoría de referencia intermedia (utilizando el subcomando `CONTRAST` para los contrastes de desviación o simples).
- Especificar la métrica para los contrastes polinómicos (utilizando el subcomando `CONTRAST`).
- Especificar términos de error para las comparaciones post hoc (utilizando el subcomando `POSTHOC`).
- Calcular medias marginales estimadas para cualquier factor o interacción entre los factores en la lista de factores (utilizando el subcomando `EMMEANS`).
- Especificar nombres para las variables temporales (utilizando el subcomando `SAVE`).
- Construir un archivo de datos matricial de correlaciones (utilizando el subcomando `OUTFILE`).
- Construir un archivo de datos matricial que contenga estadísticos de la tabla de ANOVA inter-sujetos (utilizando el subcomando `OUTFILE`).
- Guardar la matriz del diseño en un nuevo archivo de datos (utilizando el subcomando `OUTFILE`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Correlaciones bivariadas

El procedimiento Correlaciones bivariadas calcula el coeficiente de correlación de Pearson, la rho de Spearman y la tau-*b* de Kendall con sus niveles de significación. Las correlaciones miden cómo están relacionadas las variables o los órdenes de los rangos. Antes de calcular un coeficiente de correlación, inspeccione los datos para detectar valores atípicos (que pueden generar resultados equívocos) y evidencias de una relación lineal. El coeficiente de correlación de Pearson es una medida de asociación lineal. Dos variables pueden estar perfectamente relacionadas, pero si la relación no es lineal, el coeficiente de correlación de Pearson no será un estadístico adecuado para medir su asociación.

Ejemplo. ¿Está el número de partidos ganados por un equipo de baloncesto correlacionado con el número medio de puntos anotados por partido? Un diagrama de dispersión indica que existe una relación lineal. Al analizar los datos de la temporada 1994–1995 de la NBA, se descubre que el coeficiente de correlación de Pearson (0,581) es significativo al nivel 0,01. Se puede sospechar que cuantos más partidos se ganen por temporada, menos puntos habrán anotado los adversarios. Estas variables están correlacionadas negativamente (–0,401) y la correlación es significativa al nivel 0,05.

Estadísticos. Para cada variable: número de casos sin valores perdidos, desviación típica y media. Para cada pareja de variables: coeficiente de correlación de Pearson, rho de Spearman, tau-*b* de Kendall, productos cruzados de las desviaciones y covarianzas.

Datos. Utilice variables cuantitativas simétricas para el coeficiente de correlación de Pearson y variables cuantitativas o variables con categorías ordenadas para la rho de Spearman y la tau-*b* de Kendall.

Supuestos. El coeficiente de correlación de Pearson asume que cada pareja de variables es normal bivariada.

Figura 23-1
Resultados del procedimiento *Correlaciones bivariadas*

		Partidos ganados	Puntos obtenidos por partido	Puntos en contra por partido
Correlación de Pearson	Partidos ganados	1.000	.581**	-.401*
	Puntos obtenidos por partido	.581**	1.000	.457*
	Puntos en contra por partido	-.401*	.457*	1.000
Sig. (bilateral)	Partidos ganados	.	.001	.038
	Puntos obtenidos por partido	.001	.	.017
	Puntos en contra por partido	.038	.017	.
N	Partidos ganados	27	27	27
	Puntos obtenidos por partido	27	27	27
	Puntos en contra por partido	27	27	27

** La correlación es significativa al nivel 0,01 (bilateral).

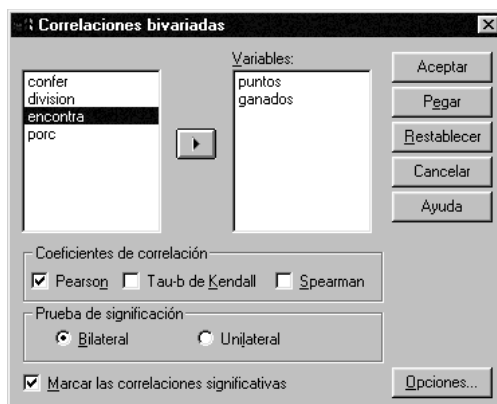
* La correlación es significante al nivel 0,05 (bilateral).

Para obtener correlaciones bivariadas

Elija en los menús:

Analizar
Correlaciones
Bivariadas...

Figura 23-2
Cuadro de diálogo *Correlaciones bivariadas*



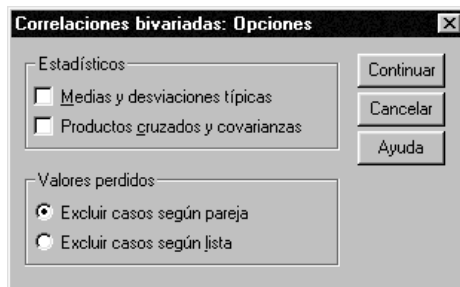
- Seleccione dos o más variables numéricas.

También se encuentran disponibles las siguientes opciones:

- **Coefficientes de correlación.** Para las variables cuantitativas, normalmente distribuidas, seleccione el coeficiente de correlación de Pearson. Si los datos no están normalmente distribuidos o tienen categorías ordenadas, seleccione los correspondientes a la Tau-b de Kendall o Spearman, que miden la asociación entre órdenes de rangos. Los coeficientes de correlación pueden estar entre -1 (una relación negativa perfecta) y $+1$ (una relación positiva perfecta). Un valor 0 indica que no existe una relación lineal. Al interpretar los resultados, se debe evitar extraer conclusiones de causa-efecto a partir de una correlación significativa.
- **Prueba de significación.** Se pueden seleccionar las probabilidades bilaterales o las unilaterales. Si conoce de antemano la dirección de la asociación, seleccione Unilateral. Si no es así, seleccione Bilateral.
- **Marcar las correlaciones significativas.** Los coeficientes de correlación significativos al nivel $0,05$ se identifican por medio de un solo asterisco y los significativos al nivel $0,01$ se identifican con dos asteriscos.

Correlaciones bivariadas: Opciones

Figura 23-3
Cuadro de diálogo Correlaciones bivariadas: Opciones



Estadísticos. Para las correlaciones de Pearson, se puede elegir una o ambas de estas opciones:

- **Medias y desviaciones típicas.** Se muestran para cada variable. También se muestra el número de casos que no tienen valores perdidos. Los valores perdidos se consideran según cada variable individual, sin tener en cuenta la opción elegida para la manipulación de los valores perdidos.
- **Productos cruzados diferenciales y covarianzas.** Se muestran para cada pareja de variables. Cada producto cruzado de las desviaciones es igual a la suma de los productos de las variables corregidas respecto a la media. Éste es el numerador del coeficiente de correlación de Pearson. La covarianza es una medida no tipificada de la relación entre dos variables, igual al producto cruzado diferencial dividido por $N-1$.

Valores perdidos. Puede elegir uno de los siguientes:

- **Excluir casos según pareja.** Se excluyen del análisis los casos con valores perdidos para una o ambas variables de la pareja que forma un coeficiente de correlación. Debido a que cada coeficiente está basado en todos los casos que tienen códigos válidos para esa pareja concreta de variables, en cada cálculo se utiliza la mayor cantidad de información disponible. Esto puede dar como resultado un grupo de coeficientes basados en un número de casos variable.
- **Excluir casos según lista.** Se excluyen de todas las correlaciones los casos con valores perdidos para cualquier variable.

Funciones adicionales de los comandos CORRELATIONS y NONPAR CORR

El lenguaje de comandos de SPSS también permite:

- Escribir una matriz de correlaciones para correlaciones de Pearson que pueda ser utilizada en lugar de los datos brutos, con el fin de obtener otros análisis como el análisis factorial (con el subcomando `MATRIX`).
- Obtener correlaciones de todas las variables de una lista con todas las variables de una segunda lista (utilizando la palabra clave `WITH` en el subcomando `VARIABLES`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Correlaciones parciales

El procedimiento Correlaciones parciales calcula los coeficientes de correlación parcial, los cuales describen la relación lineal existente entre dos variables mientras se controlan los efectos de una o más variables adicionales. Las correlaciones son medidas de asociación lineal. Dos variables pueden estar perfectamente relacionadas, pero si la relación no es lineal, el coeficiente de correlación no es un estadístico adecuado para medir su asociación.

Ejemplo. ¿Existe alguna relación entre la financiación sanitaria y las tasas de enfermedad? Aunque cabe esperar que dicha relación sea negativa, un estudio describe una correlación *positiva* significativa: si la financiación sanitaria aumenta, las tasas de enfermedad parecen disminuir. No obstante, el control de la tasa de visitas médicas elimina virtualmente la correlación positiva observada. La financiación sanitaria y las tasas de enfermedad sólo parecen estar relacionadas positivamente debido a que más personas tienen acceso a la sanidad si la financiación aumenta, lo que tiene como resultado que los médicos y hospitales informen de más enfermedades.

Estadísticos. Para cada variable: número de casos sin valores perdidos, desviación típica y media. Matrices de correlación de orden cero y parcial, con grados de libertad y niveles de significación.

Datos. Utilice variables cuantitativas y simétricas.

Supuestos. El procedimiento Correlaciones parciales supone que cada par de variables es normal bivalente.

Figura 24-1
Resultados de las correlaciones parciales

```

- - - P A R T I A L   C O R R E L A T I O N   C O E F F I C I E N T S   - - -
Zero Order Partial

      GEB_RT   STERB_RT   LOG_BIP
GEB_RT   1.0000   .3670   -.7674
          (  0)   ( 106)   ( 106)
          P= .   P= .000   P= .000

STERB_RT   .3670   1.0000   -.4015
          ( 106)   (  0)   ( 106)
          P= .000   P= .   P= .000

LOG_BIP   -.7674   -.4015   1.0000
          ( 106)   ( 106)   (  0)
          P= .000   P= .000   P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

- - - P A R T I A L   C O R R E L A T I O N   C O E F F I C I E N T S   - - -
Controlling for..   LOG_BIP

      GEB_RT   STERB_RT
GEB_RT   1.0000   .1003
          (  0)   ( 105)
          P= .   P= .304

STERB_RT   .1003   1.0000
          ( 105)   (  0)
          P= .304   P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed
\par \par Controlling for..   LOG_BIP
\par \par      GEB_RT   STERB_RT
\par \par GEB_RT   1.0000   .1003
\par \par      (  0)   ( 105)
\par \par      P= .   P= .304
\par \par STERB_RT   .1003   1.0000
\par \par      ( 105)   (  0)
\par \par      P= .304   P= .
\par \par (Coefficient / (D.F.) / 2-tailed Significance)
\par \par " . " is printed if a coefficient cannot be computed
\par )

```

Para obtener correlaciones parciales

- ▶ Elija en los menús:
 Analizar
 Correlaciones
 Parciales...

Figura 24-2
 Cuadro de diálogo Correlaciones parciales



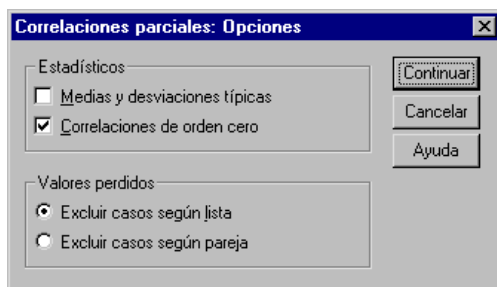
- ▶ Seleccione dos o más variables numéricas para las que se van a calcular las correlaciones parciales.
- ▶ Elija una o más variables numéricas de control.

También se encuentran disponibles las siguientes opciones:

- **Prueba de significación.** Se pueden seleccionar las probabilidades bilaterales o las unilaterales. Si conoce de antemano la dirección de la asociación, seleccione Unilateral. Si no es así, seleccione Bilateral.
- **Mostrar el nivel de significación real.** Por defecto, se muestran la probabilidad y los grados de libertad para cada coeficiente de correlación. Si anula la selección de este elemento, los coeficientes significativos al nivel 0,05 se identifican con un asterisco, los coeficientes significativos al nivel 0,01 se identifican con un asterisco doble y se eliminan los grados de libertad. Este ajuste afecta a las matrices de correlación parcial y de orden cero.

Correlaciones parciales: Opciones

Figura 24-3
Cuadro de diálogo Correlaciones parciales: Opciones



Estadísticos. Puede elegir una o ambas de las siguientes opciones:

- **Medias y desviaciones típicas.** Se muestran para cada variable. También se muestra el número de casos que no tienen valores perdidos.
- **Correlaciones de orden cero.** Se muestra una matriz de las correlaciones simples entre todas las variables, incluyendo las variables de control.

Valores perdidos. Puede elegir una de las siguientes alternativas:

- **Excluir casos según lista.** Se excluyen de todos los cálculos los casos que presenten valores perdidos para cualquier variable, incluso si es para las variables de control.
- **Excluir casos según pareja.** Para el cálculo de las correlaciones de orden cero, en las que se basan las correlaciones parciales, no se utilizará un caso si tiene valores perdidos en una o ambas variables de un par. La eliminación según pareja aprovecha el máximo de los datos que sean posibles. Sin embargo, el número de casos puede variar de unos coeficientes a otros. Cuando se activa esta opción, los grados de libertad para un coeficiente parcial determinado se basan en el número menor de casos utilizado en el cálculo de cualquiera de las correlaciones de orden cero necesarias para el cálculo de dicho coeficiente parcial.

Distancias

Este procedimiento calcula una variedad de estadísticos que miden las similitudes o diferencias (distancias), entre pares de variables o entre pares de casos. Estas medidas de similitud o distancia se pueden utilizar después con otros procedimientos, como análisis factorial, análisis de conglomerados o escalamiento multidimensional, para ayudar en el análisis de conjuntos de datos complejos.

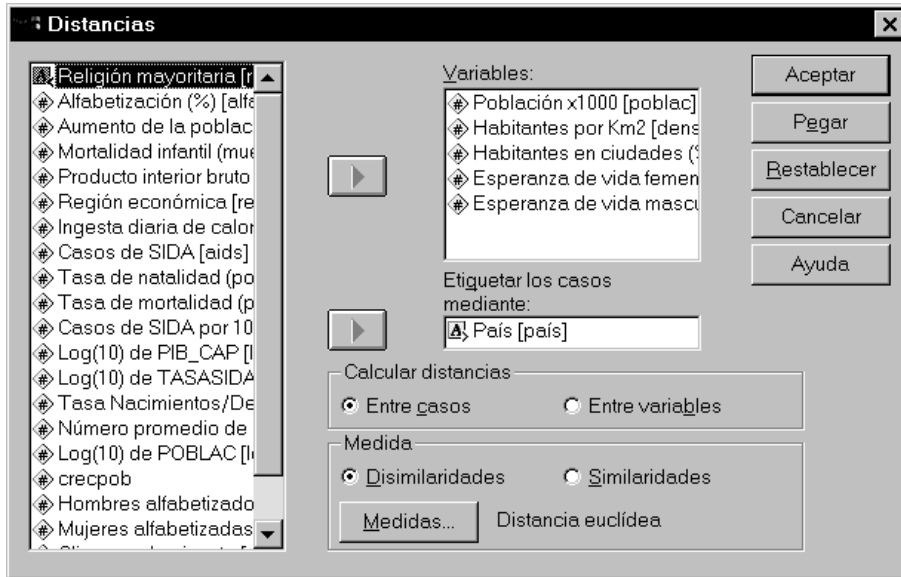
Ejemplo. ¿Es posible medir similitudes entre pares de automóviles en función de ciertas características, como tipo de motor, consumo y potencia? Al calcular las similitudes entre los coches, se puede obtener una noción de qué coches son similares entre sí y cuáles son diferentes. Para un análisis más formal, puede considerar la aplicación de un análisis de conglomeración jerárquico o escalamiento multidimensional a las similitudes para explorar la estructura subyacente.

Estadísticos. Las medidas de diferencia (distancia) para datos de un intervalo son Distancia euclídea, Distancia euclídea al cuadrado, Chebychev, bloque, Minkowski o personalizada; para datos de recuento, medida de chi-cuadrado o phi-cuadrado; para datos binarios, Distancia euclídea, Distancia euclídea al cuadrado, diferencia de tamaño, diferencia de configuración, varianza, forma o Lance y Williams. Las medidas de similitud para datos de intervalos son correlación de Pearson o coseno; para datos binarios, Russel y Rao, concordancia simple, Jaccard, Dice, Rogers y Tanimoto, Sokal y Sneath 1, Sokal y Sneath 2, Sokal y Sneath 3, Kulczynski 1, Kulczynski 2, Sokal y Sneath 4, Hamann, Lambda, D de Anderberg, Y de Yule, Q de Yule, Ochiai, Sokal y Sneath 5, correlación Phi de 4 puntos o dispersión.

Para obtener matrices de distancias

- ▶ Elija en los menús:
 - Analizar
 - Correlaciones
 - Distancias...

Figura 25-1
Cuadro de diálogo Distancias



- ▶ Seleccione al menos una variable numérica para calcular distancias entre casos o seleccione al menos dos variables numéricas para calcular distancias entre variables.
- ▶ Seleccione una alternativa en el grupo Calcular distancias para calcular proximidades entre casos o entre variables.

Distancias: Medidas de disimilaridad

Figura 25-2
Cuadro de diálogo Distancias: Medidas de disimilaridad

En el grupo Medida, seleccione la alternativa que corresponda al tipo de datos (intervalo, recuento o binario); a continuación, de la lista desplegable, seleccione una las medidas que corresponda a dicho tipo de datos. Las medidas disponibles, por tipo de dato, son:

- **Datos de intervalo.** Distancia euclídea, Distancia euclídea al cuadrado, Chebychev, Bloque, Minkowski o Personalizada.
- **Datos de frecuencias.** Medida de chi-cuadrado o Medida de phi-cuadrado.
- **Datos binarios.** Distancia euclídea, Distancia euclídea al cuadrado, Diferencia de tamaño, Diferencia de configuración, Varianza, Forma o Lance y Williams. (Introduzca valores para Presente y Ausente para especificar cuáles son los dos valores representativos; las Distancias ignorarán todos los demás valores.)

El grupo Transformar valores permite estandarizar los valores de los datos para casos o variables *antes* de calcular proximidades. Estas transformaciones no se pueden aplicar a los datos binarios. Los métodos disponibles de estandarización son: Puntuaciones z , Rango -1 a 1 , Rango 0 a 1 , Magnitud máxima de 1 , Media de 1 y Desviación típica 1 .

El grupo Transformar medidas permite transformar los valores generados por la medida de distancia. Se aplican después de calcular la medida de distancia. Las opciones disponibles son: Valores absolutos, Cambiar el signo y Cambiar la escala al rango 0–1.

Distancias: Medidas de similitud

Figura 25-3
Cuadro de diálogo *Distancias: Medidas de similitud*



En el grupo Medida, seleccione la alternativa que corresponda al tipo de datos (intervalo o binario); a continuación, de la lista desplegable, seleccione una las medidas que corresponda a dicho tipo de datos. Las medidas disponibles, por tipo de dato, son:

- **Datos de intervalo.** Correlación de Pearson o Coseno.
- **Datos binarios.** Russel y Rao, Concordancia simple, Jaccard, Dice, Rogers y Tanimoto, Sokal y Sneath 1, Sokal y Sneath 2, Sokal y Sneath 3, Kulczynski 1, Kulczynski 2, Sokal y Sneath 4, Hamann, Lambda, *D* de Anderberg, *Y* de Yule, *Q* de Yule, Ochiai, Sokal y Sneath 5, Correlación Phi de 4 puntos o Dispersión. (Introduzca valores para Presente y Ausente para especificar cuáles son los dos valores representativos; las Distancias ignorarán todos los demás valores.)

El grupo Transformar valores permite estandarizar los valores de los datos para casos o variables antes de calcular proximidades. Estas transformaciones no se pueden aplicar a los datos binarios. Los métodos disponibles de estandarización son: Puntuaciones *z*, Rango -1 a 1, Rango 0 a 1, Magnitud máxima de 1, Media de 1 y Desviación típica 1.

El grupo Transformar medidas permite transformar los valores generados por la medida de distancia. Se aplican después de calcular la medida de distancia. Las opciones disponibles son: Valores absolutos, Cambiar el signo y Cambiar la escala al rango 0-1.

Regresión lineal

La regresión lineal estima los coeficientes de la ecuación lineal, con una o más variables independientes, que mejor prediga el valor de la variable dependiente. Por ejemplo, puede intentar predecir el total de ventas anuales de un vendedor (la variable dependiente) a partir de variables independientes tales como la edad, la formación y los años de experiencia.

Ejemplo. ¿Están relacionados el número de partidos ganados por un equipo de baloncesto en una temporada con la media de puntos que el equipo marca por partido? Un diagrama de dispersión indica que estas variables están relacionadas linealmente. El número de partidos ganados y la media de puntos marcados por el equipo adversario también están relacionados linealmente. Estas variables tienen una relación negativa. A medida que el número de partidos ganados aumenta, la media de puntos marcados por el equipo adversario disminuye. Con la regresión lineal es posible modelar la relación entre estas variables. Puede utilizarse un buen modelo para predecir cuántos partidos ganarán los equipos.

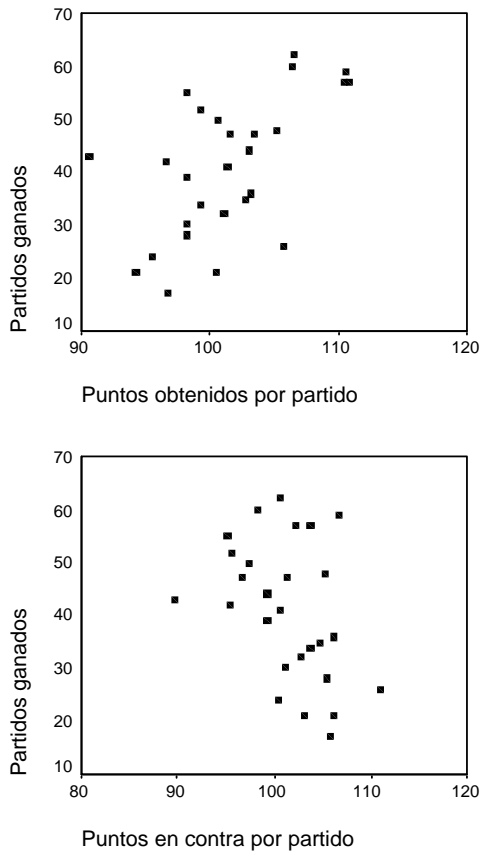
Estadísticos. Para cada variable: número de casos válidos, media y desviación típica. Para cada modelo: coeficientes de regresión, matriz de correlaciones, correlaciones parciales y semiparciales, R múltiple, R^2 , R^2 corregida, cambio en R^2 , error típico de la estimación, tabla de análisis de varianza, valores pronosticados y residuos. Además, intervalos de confianza al 95% para cada coeficiente de regresión, matriz de varianzas-covarianzas, factor de inflación de la varianza, tolerancia, prueba de Durbin-Watson, medidas de distancia (Mahalanobis, Cook y valores de influencia), Df_{β} , Df_{ajuste} , intervalos de pronóstico y diagnósticos por caso. Gráficos: diagramas de dispersión, gráficos parciales, histogramas y gráficos de probabilidad normal.

Datos. Las variables dependiente e independientes deben ser cuantitativas. Las variables categóricas, como la religión, estudios principales o el lugar de residencia, han de recodificarse como variables binarias (dummy) o como otros tipos de variables de contraste.

Supuestos. Para cada valor de la variable independiente, la distribución de la variable dependiente debe ser normal. La varianza de distribución de la variable dependiente debe ser constante para todos los valores de la variable independiente. La relación entre la variable dependiente y cada variable independiente debe ser lineal y todas las observaciones deben ser independientes.

Figura 26-1

Resultados de la regresión lineal



Resumen del modelo ²

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.947 ¹	.898	.889	4.40

1. Variables predictoras: (Constante), Puntos en contra por partido, Puntos obtenidos por partido

2. Variable dependiente: Partidos ganados

ANOVA²

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	4080.533	2	2040.266	105.198	.000 ¹
	Residual	465.467	24	19.394		
	Total	4546.000	26			

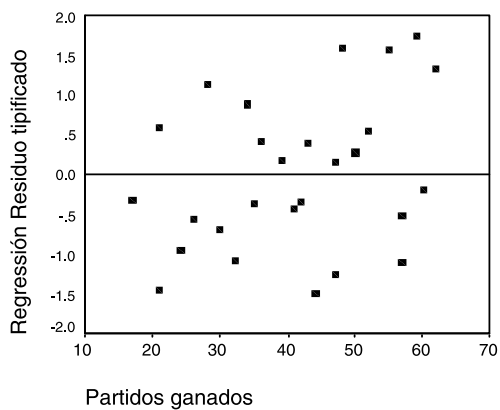
1. Variables predictoras: (Constante), Puntos en contra por partido, Puntos obtenidos por partido

2. Variable dependiente: Partidos ganados

Coeficientes¹

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	28.121	21.404		1.314	.201
	Puntos obtenidos por partido	2.539	.193	.965	13.145	.000
	Puntos en contra por partido	-2.412	.211	-.841	-11.458	.000

1. Variable dependiente: Partidos ganados



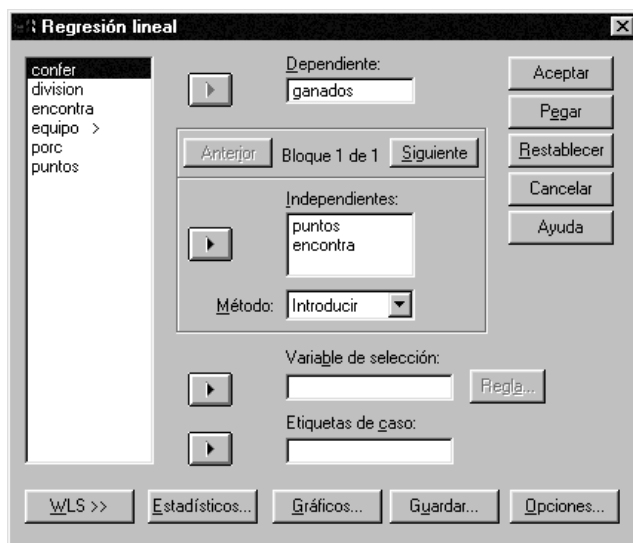
Para obtener un análisis de regresión lineal

- Elija en los menús:

Analizar
Regresión
Lineal...

Figura 26-2

Cuadro de diálogo Regresión lineal



- ▶ En el cuadro de diálogo Regresión lineal, seleccione una variable numérica dependiente.
- ▶ Seleccione una más variables numéricas independientes.

Si lo desea, tiene la posibilidad de:

- Agrupar variables independientes en bloques y especificar distintos métodos de entrada para diferentes subconjuntos de variables.
- Elegir una variable de selección para limitar el análisis a un subconjunto de casos que tengan valores particulares para esta variable.
- Seleccionar una variable de identificación de casos para identificar los puntos en los diagramas.
- Seleccione una variable numérica de Ponderación MCP para el análisis de mínimos cuadrados ponderados.

MCP (WLS). Permite obtener un modelo de mínimos cuadrados ponderados. Los puntos de los datos se ponderan por los inversos de sus varianzas. Esto significa que las observaciones con varianzas grandes tienen menor impacto en el análisis que las observaciones asociadas a varianzas pequeñas.

Métodos de selección de variables en el análisis de regresión lineal

La selección del método permite especificar cómo se introducen las variables independientes en el análisis. Utilizando distintos métodos se pueden construir diversos modelos de regresión a partir del mismo conjunto de variables.

- **Introducir (Regresión).** Procedimiento para la selección de variables en el que todas las variables de un bloque se introducen en un solo paso.
- **Por pasos.** En cada paso se introduce la variable independiente que no se encuentre ya en la ecuación y que tenga la probabilidad para F más pequeña, si esa probabilidad es suficientemente pequeña. Las variables ya introducidas en la ecuación de regresión se eliminan de ella si su probabilidad para F llega a ser suficientemente grande. El método termina cuando ya no hay más variables candidatas a ser incluidas o eliminadas.

- **Eliminar (Selección de variables).** Procedimiento para la selección de variables en el que las variables de un bloque se eliminan en un solo paso.
- **Eliminación hacia atrás.** Procedimiento de selección de variables en el que se introducen todas las variables en la ecuación y después se van excluyendo una tras otra. Aquella variable que tenga la menor correlación parcial con la variable dependiente será la primera en ser considerada para su exclusión. Si satisface el criterio de eliminación, será eliminada. Tras haber excluido la primera variable, se pondrá a prueba aquella variable, de las que queden en la ecuación, que presente una correlación parcial más pequeña. El procedimiento termina cuando ya no quedan en la ecuación variables que satisfagan el criterio de exclusión.
- **Selección hacia adelante.** Procedimiento de selección de variables en el que éstas son introducidas secuencialmente en el modelo. La primera variable que se considerará para ser introducida en la ecuación será aquella que tenga mayor correlación, positiva o negativa, con la variable dependiente. Dicha variable será introducida en la ecuación sólo si satisface el criterio de entrada. Si ha entrado la primera variable, se considerará como próxima candidata la variable independiente que no esté en la ecuación y cuya correlación parcial sea la mayor. El procedimiento termina cuando ya no quedan variables que satisfagan el criterio de entrada.

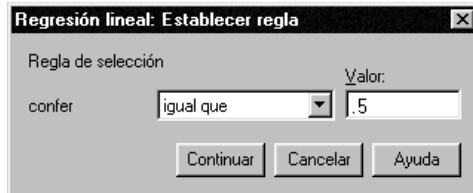
Los valores de significación de los resultados se basan en el ajuste de un único modelo. Por ello, estos valores no suele ser válidos cuando se emplea un método por pasos (Pasos sucesivos, Hacia adelante o Hacia atrás).

Todas las variables deben superar el criterio de tolerancia para que puedan ser introducidas en la ecuación, independientemente del método de entrada especificado. El nivel de tolerancia por defecto es 0,0001. Tampoco se introduce una variable si esto provoca que la tolerancia de otra ya presente en el modelo se sitúe por debajo del criterio de tolerancia.

Todas las variables independientes seleccionadas se añaden a un mismo modelo de regresión. Sin embargo, puede especificar distintos métodos de introducción para diferentes subconjuntos de variables. Por ejemplo, puede introducir en el modelo de regresión un bloque de variables que utilice la selección por pasos sucesivos, y un segundo bloque que emplee la selección hacia adelante. Para añadir un segundo bloque de variables al modelo de regresión, pulse en **Siguiente**.

Regresión lineal: Establecer regla

Figura 26-3
Cuadro de diálogo Regresión lineal: Establecer regla



Los casos definidos por la regla de selección se incluyen en el análisis. Por ejemplo, si selecciona una variable, elija igual que y escriba 5 para el valor; de este modo, solamente se incluirán en el análisis los casos para los cuales la variable seleccionada tenga un valor igual a 5. También se permite un valor de cadena.

Regresión lineal: Gráficos

Figura 26-4
Cuadro de diálogo Regresión lineal: Gráficos



Los gráficos pueden ayudar a validar los supuestos de normalidad, linealidad e igualdad de las varianzas. También son útiles para detectar valores atípicos, observaciones poco usuales y casos de influencia. Tras guardarlos como nuevas variables, dispondrá en el Editor de datos de los valores pronosticados, los residuos y

otros valores diagnósticos, con los cuales podrá poder crear gráficos respecto a las variables independientes. Se encuentran disponibles los siguientes gráficos:

Diagramas de dispersión. Puede representar cualquier combinación por parejas de la lista siguiente: la variable dependiente, los valores pronosticados tipificados, los residuos tipificados, los residuos eliminados, los valores pronosticados corregidos, los residuos estudentizados o los residuos eliminados estudentizados. Represente los residuos tipificados frente a los valores pronosticados tipificados para contrastar la linealidad y la igualdad de las varianzas.

Lista de variables de origen. Muestra una lista con la variable dependiente (DEPENDNT) y las siguientes variables pronosticadas y residuales: Valores pronosticados tipificados (*ZPRED), Residuos tipificados (*ZRESID), Residuos eliminados (*DRESID), Valores pronosticados corregidos (*ADJPRED), Residuos estudentizados (*SRESID) y Residuos estudentizados eliminados (*SDRESID).

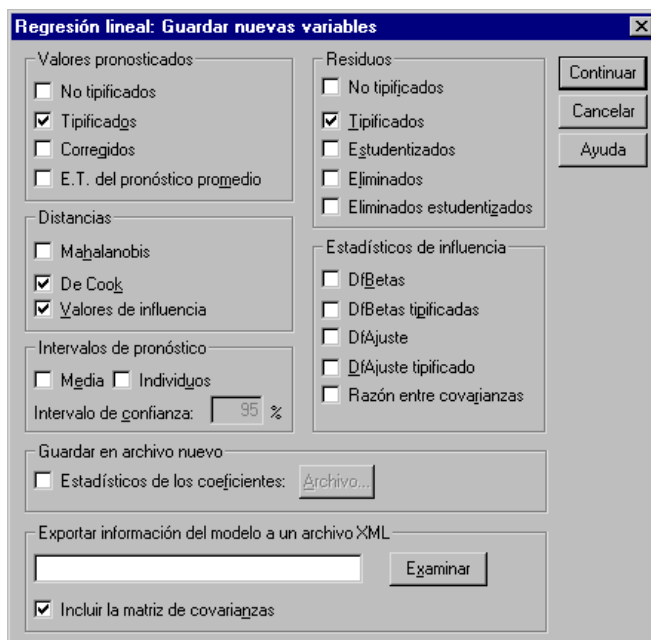
Generar todos los gráficos parciales. Muestra los diagramas de dispersión de los residuos de cada variable independiente y los residuos de la variable dependiente cuando se regresan ambas variables por separado sobre las restantes variables independientes. En la ecuación debe haber al menos dos variables independientes para que se generen los gráficos parciales.

Gráficos de residuos tipificados. Puede obtener histogramas de los residuos tipificados y gráficos de probabilidad normal que comparen la distribución de los residuos tipificados con una distribución normal.

Si se solicita cualquier gráfico, se muestran los estadísticos de resumen para los valores pronosticados tipificados y los residuos tipificados (*ZPRED y *ZRESID).

Regresión lineal: Para guardar variables nuevas

Figura 26-5
Cuadro de diálogo Regresión lineal: Guardar



Puede guardar los valores pronosticados, los residuos y otros estadísticos útiles para los diagnósticos. Cada selección añade una o más variables nuevas a su archivo de datos activo.

Valores pronosticados. Son los valores que el modelo de regresión pronostica para cada caso.

- **No tipificados.** Valor pronosticado por el modelo para la variable dependiente.
- **Tipificados.** Transformación de cada valor pronosticado a su forma tipificada. Es decir, se sustrae el valor pronosticado medio al valor pronosticado y el resultado se divide por la desviación típica de los valores pronosticados. Los valores pronosticados tipificados tienen una media de 0 y una desviación típica de 1.

- **Corregidos.** Valor pronosticado para un caso cuando dicho caso no se incluye en los cálculos de los coeficientes de regresión.
- **E.T. del pronóstico promedio.** Error típico de los valores pronosticados. Estimación de la desviación típica del valor promedio de la variable dependiente para los casos que tienen valores iguales en las variables independientes.

Distancias. Son medidas para identificar casos con combinaciones poco usuales de valores para las variables independientes y casos que puedan tener un gran impacto en el modelo.

- **Mahalanobis.** Medida de cuánto difieren los valores de un caso en las variables independientes respecto al promedio para todos los casos. Una distancia de Mahalanobis grande identificará a un caso que tenga valores extremos en una o más de las variables independientes.
- **De Cook.** Medida de cuánto cambiarían los residuos de todos los casos si se excluyera un caso determinado del cálculo de los coeficientes de regresión. Una Distancia de Cook grande indica que la exclusión de ese caso del cálculo de los estadísticos de regresión hará variar substancialmente los coeficientes.
- **Valores de influencia.** Medida de la influencia de un punto en el ajuste de la regresión. La influencia centrada varía entre 0 (no influye en el ajuste) y $(N-1)/N$.

Intervalos de pronóstico. Los límites superior e inferior para los intervalos de pronóstico individual y promedio.

- **Media (Intervalo de Confianza).** Límites inferior y superior (dos variables) para el intervalo de predicción de la respuesta pronosticada promedio.
- **Individual.** Límites superior e inferior (dos variables) del intervalo de predicción para la variable dependiente para un caso individual.
- **Intervalo de confianza.** Introduzca un valor entre 1 y 99,99 para especificar el nivel de confianza para los dos intervalos de predicción. Debe seleccionar Media o Individuos antes de introducir este valor. Los valores habituales para los intervalos de confianza son 90, 95 y 99.

Residuos. El valor actual de la variable dependiente menos el valor pronosticado por la ecuación de regresión.

- **No tipificados.** Diferencia entre un valor observado y el valor pronosticado por el modelo.

- **Tipificados.** El residuo dividido por una estimación de su error típico. Los residuos tipificados, que son conocidos también como los residuos de Pearson o residuos estandarizados, tienen una media de 0 y una desviación típica de 1.
- **Método de Student.** Residuo dividido por una estimación de su desviación típica que varía de caso en caso, dependiendo de la distancia de los valores de cada caso en las variables independientes respecto a las medias en las variables independientes.
- **Eliminados.** Residuo para un caso cuando ese caso se excluye del cálculo de los coeficientes de regresión. Es la diferencia entre el valor de la variable dependiente y el valor pronosticado corregido.
- **Eliminados estudentizados.** Residuo eliminado para un caso dividido por su error típico. La diferencia entre un residuo eliminado estudentizado y su residuo estudentizado asociado indica la diferencia que implica el eliminar un caso sobre su propia predicción.

Estadísticos de influencia. El cambio en los coeficientes de regresión ($DfBeta$) y en los valores pronosticados ($DfAjuste$) que resulta de la exclusión de un caso particular. También están disponibles los valores tipificados para las $DfBeta$ y para las $DfAjuste$, junto con la razón entre covarianzas.

- **Diferencias en las betas ($DfBet$ s).** La diferencia en el valor de beta es el cambio en el valor de un coeficiente de regresión que resulta de la exclusión de un caso particular. Se calcula un valor para cada término del modelo, incluyendo la constante.
- **$DfBeta$ tipificada.** Valor de la diferencia en beta tipificada. El cambio tipificado en un coeficiente de regresión cuando se elimina del análisis un caso particular. Puede interesarle examinar aquellos casos cuyos valores absolutos sean mayores que 2 dividido por la raíz cuadrada de N , donde N es el número de casos. Se calcula un valor para cada término del modelo, incluyendo la constante.
- **Diferencia en el valor ajustado ($DfFit$).** La diferencia en el valor ajustado es el cambio en el valor pronosticado que resulta de la exclusión de un caso particular.
- **$DfAjuste$ tipificada.** Diferencia tipificada en el valor ajustado. El cambio, tipificado, en el valor pronosticado que resulta de la exclusión de un caso particular. Puede interesarle examinar aquellos valores tipificados cuyo valor

absoluto sea mayor que 2 dividido por la raíz cuadrada de p/N , donde p es el número de variables independientes en la ecuación y N es el número de casos.

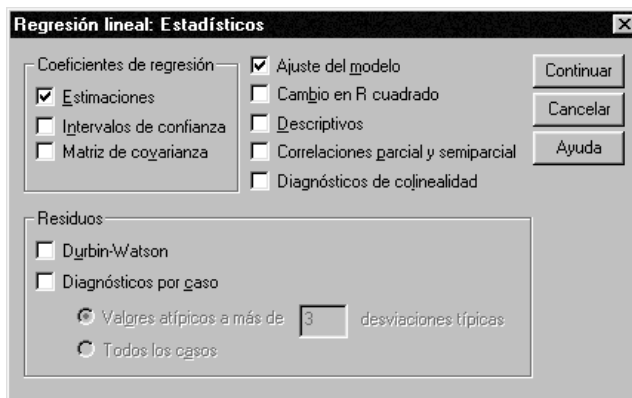
- **Razón entre covarianzas.** Razón del determinante de la matriz de covarianza con un caso particular excluido del cálculo de los coeficientes de regresión, respecto al determinante de la matriz de covarianza con todos los casos incluidos. Si la razón se aproxima a 1, el caso no altera significativamente la matriz de covarianza.

Guardar en archivo nuevo. Almacena los coeficientes de regresión en el archivo que especifique.

Exportar información del modelo a un archivo XML. Las estimaciones de los parámetros y (si lo desea) sus covarianzas se exportan al archivo especificado en formato XML (PMML). *SmartScore* y la versión de servidor de *SPSS* (un producto independiente) pueden utilizar este archivo de modelo para aplicar la información de modelo a otros archivos de datos con fines de puntuación.

Regresión lineal: Estadísticos

Figura 26-6
Cuadro de diálogo Regresión lineal: Estadísticos



Se encuentran disponibles los siguientes estadísticos:

Coeficientes de regresión. La opción Estimaciones muestra el coeficiente de regresión B , el error típico de B , el coeficiente beta tipificado, el valor de t para B y el nivel de significación bilateral t . Intervalos de confianza muestra intervalos de

confianza al 95% para cada coeficiente de regresión o una matriz de covarianzas. Matriz de covarianzas muestra una matriz de varianzas-covarianzas de los coeficientes de regresión, con las covarianzas fuera de la diagonal y las varianzas en la diagonal. También se muestra una matriz de correlaciones.

Ajuste del modelo. Presenta una lista de las variables introducidas y eliminadas del modelo y muestra los siguientes estadísticos de bondad de ajuste: R múltiple, R^2 y R^2 corregida, error típico de la estimación y tabla de análisis de la varianza.

Cambio en R^2 cuadrado. Cambio en el estadístico R^2 cuadrado que se produce al añadir o eliminar una variable independiente. Si es grande el cambio en R^2 cuadrado asociado a una variable, esto significa que esa variable es un buen predictor de la variable dependiente.

Descriptivos. Proporciona el número de casos válidos, la media y la desviación típica para cada variable en el análisis. También muestra una matriz de correlaciones con el nivel de significación unilateral y el número de casos para cada correlación.

Correlación parcial. La correlación que permanece entre dos variables después de eliminar la correlación que es debida a su relación mutua con las otras variables. La correlación entre la variable dependiente y una variable independiente cuando se han eliminado de ambas los efectos lineales de las otras variables independientes presentes en el modelo.

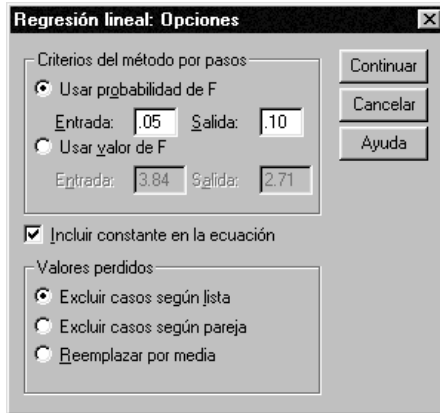
Correlación semi-parcial (Regresión, Tabla pivote). La correlación entre la variable dependiente y una variable independiente cuando se han eliminado de la variable independiente los efectos lineales de las otras variables independientes del modelo. Está relacionada con el cambio en R^2 cuadrado cuando una variable se añade a una ecuación.

Diagnósticos de colinealidad. La colinealidad (o multicolinealidad) es una situación no deseable en la que una de las variables independientes es una función lineal de otras variables independientes. Muestra los autovalores de la matriz de productos cruzados no centrada y escalada, los índices de condición y las proporciones de la descomposición de la varianza junto con los factores de inflación de la varianza (FIV) y las tolerancias para las variables individuales.

Residuos. Presenta la prueba de Durbin-Watson sobre la correlación serial de los residuos y los diagnósticos por casos para los casos que cumplan el criterio de selección (los valores atípicos por encima de n desviaciones típicas).

Regresión lineal: Opciones

Figura 26-7
Cuadro de diálogo Regresión lineal: Opciones



Se encuentran disponibles las siguientes opciones:

Criterios del método por pasos. Estas opciones son aplicables si se ha especificado el método de selección de variables hacia adelante, hacia atrás o por pasos. Las variables se pueden introducir o eliminar del modelo dependiendo de la significación (probabilidad) del valor de F o del propio valor de F .

- **Usar la probabilidad de F .** Una variable se introduce en el modelo si el nivel de significación de su valor de F es menor que el valor de entrada, y se elimina si el nivel de significación de su valor de F es mayor que el valor de salida. La entrada debe ser menor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, eleve el valor de entrada. Para eliminar más variables del modelo, disminuya el valor de salida.
- **Usar valor de F .** Una variable se introduce en el modelo si su valor de F es mayor que el valor de entrada, y se elimina si su valor de F es menor que el valor de salida. La entrada debe ser mayor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, disminuya el valor de entrada. Para eliminar más variables del modelo, eleve el valor de salida.

Incluir la constante en la ecuación. Por defecto, el modelo de regresión incluye un término constante. Si se anula la selección de esta opción se obtiene la regresión que pasan por el origen, lo cual se hace raramente. Algunos resultados de la regresión

que pasan por el origen no son comparables con los resultados de la regresión que sí incluyen una constante. Por ejemplo, R^{cuadrado} no puede interpretarse de la manera usual.

Valores perdidos. Puede elegir uno de los siguientes:

- **Excluir casos según lista.** Sólo se incluirán en el análisis los casos con valores válidos para todas las variables.
- **Excluir casos según pareja.** Los casos con datos completos para la pareja de variables correlacionadas se utilizan para calcular el coeficiente de correlación en el cual se basa el análisis de regresión. Los grados de libertad se basan en el N mínimo de las parejas.
- **Reemplazar por la media.** Se emplean todos los casos en los cálculos, sustituyendo las observaciones perdidas por la media de la variable.

Estimación curvilínea

El procedimiento Estimación Curvilínea genera estadísticos de estimación curvilínea por regresión y gráficos relacionados para 11 modelos diferentes de estimación curvilínea por regresión. Se produce un modelo diferente para cada variable dependiente. También se pueden guardar valores pronosticados, residuos e intervalos pronosticados como nuevas variables.

Ejemplo. Un proveedor de servicios de Internet realiza un seguimiento del porcentaje de tráfico de correo electrónico infectado de virus en la red a lo largo del tiempo. Un diagrama de dispersión revela que la relación es no lineal. Se puede ajustar un modelo lineal a los datos y comprobar la validez de los supuestos y la bondad de ajuste del modelo.

Estadísticos. Para cada modelo: coeficientes de regresión, R múltiple, R^2 , R^2 corregida, error típico de la estimación, tabla de análisis de varianza, valores pronosticados, residuos e intervalos de pronóstico. Modelos: lineal, logarítmico, inverso, cuadrático, cúbico, de potencia, compuesto, curva-S, logístico, de crecimiento y exponencial.

Datos. Las variables dependiente e independientes deben ser cuantitativas. Si selecciona como variable independiente Tiempo, en lugar de una variable del archivo de datos de trabajo, el procedimiento Estimación curvilínea generará una variable de tiempo en la que la distancia temporal entre los casos es uniforme. Si se selecciona Tiempo, la variable dependiente debe ser una medida de serie temporal. El análisis de series temporales requiere una estructura particular para los archivos de datos, de manera que cada caso (cada fila) represente un conjunto de observaciones en un momento determinado del tiempo y que la distancia temporal entre los casos sea uniforme.

Supuestos. Represente los datos gráficamente para determinar cómo se relacionan las variables dependientes e independiente (linealmente, exponencialmente, etc.). Los residuos de un buen modelo deben distribuirse de forma aleatoria y normal. Si

se utiliza un modelo lineal, se deben cumplir los siguientes supuestos. Para cada valor de la variable independiente, la distribución de la variable dependiente debe ser normal. La varianza de distribución de la variable dependiente debe ser constante para todos los valores de la variable independiente. La relación entre la variable dependiente y la variable independiente debe ser lineal y todas las observaciones deben ser independientes.

Figura 27-1

Tabla de resumen de la estimación curvilínea

R	R cuadrado	R cuadrado corregida	Error típico de la estimación
,885	,783	,766	,083

La variable independiente esHours since detection.

Figura 27-2

Estimación curvilínea ANOVA

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Regresión	,949	3	,316	45,736	,000
Residual	,263	38	,007		
Total	1,212	41			

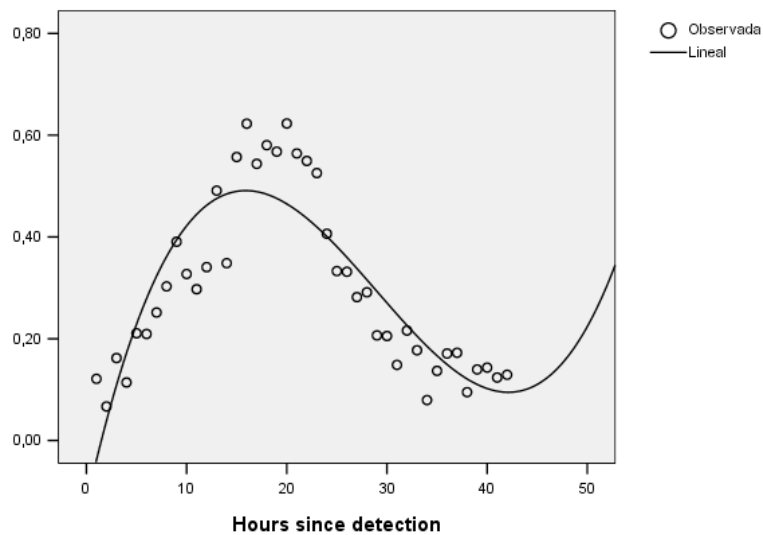
La variable independiente esHours since detection.

Figura 27-3

Coefficientes de estimación curvilínea

	Coefficientes no estandarizados		Coefficientes estandarizados	t	Sig.
	B	Desviación Error	Beta		
Hours since detection	,088	,011	6,305	7,888	,000
Hours since detection ** 2	-,004	,001	-12,111	-6,363	,000
Hours since detection ** 3	4,40E-005	,000	5,638	4,780	,000
(Constante)	-,123	,056		-2,187	,035

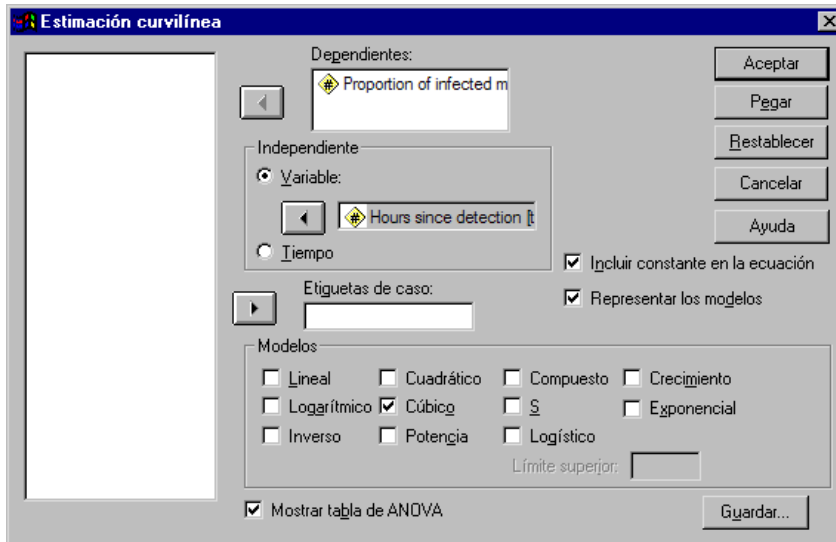
Figura 27-4
Gráfico de estimación curvilínea



Para obtener una estimación curvilínea

- ▶ Elija en los menús:
 - Analizar
 - Regresión
 - Estimación curvilínea...

Figura 27-5
Cuadro de diálogo Estimación curvilínea



- ▶ Seleccione una o más variables dependientes. Se produce un modelo diferente para cada variable dependiente.
- ▶ Seleccione una variable independiente (una variable del archivo de datos de trabajo o Tiempo).

Si lo desea, tiene la posibilidad de:

- Seleccionar una variable para etiquetar los casos en los diagramas de dispersión. Para cada punto en el diagrama de dispersión, se puede utilizar la herramienta de Identificación de puntos para mostrar el valor de la variable utilizada en Etiquetas de caso.
- Pulsar en Guardar para guardar los valores pronosticados, los residuos y los intervalos de pronóstico como nuevas variables.

También se encuentran disponibles las siguientes opciones:

- **Incluir la constante en la ecuación.** Estima un término constante en la ecuación de regresión. La constante se incluye por defecto.

- **Representar los modelos.** Representa los valores de la variable dependiente y cada modelo seleccionado frente a la variable independiente. Se genera un gráfico distinto para cada variable dependiente.
- **Ver tabla de ANOVA.** Muestra una tabla de análisis de varianza de resumen para cada modelo seleccionado.

Modelos del procedimiento Estimación curvilínea

Se puede seleccionar uno o más modelos de estimación curvilínea por regresión. Para determinar qué modelo utilizar, represente los datos. Si las variables parecen estar relacionadas linealmente, utilice un modelo de regresión lineal simple. Cuando las variables no estén relacionadas linealmente, intente transformar los datos. Cuando la transformación no resulte útil, puede necesitar un modelo más complicado. Inspeccione un diagrama de dispersión de los datos; si el diagrama se parece a una función matemática reconocible, ajuste los datos a ese tipo de modelo. Por ejemplo, si los datos se parecen a una función exponencial, utilice un modelo exponencial.

Lineal. Modelo cuya ecuación es $Y = b_0 + (b_1 * t)$. Los valores de la serie se modelan como una función lineal del tiempo.

Logarítmica. Modelo cuya ecuación es $Y = b_0 + (b_1 * \ln(t))$.

Inversa. Modelo cuya ecuación es $Y = b_0 + (b_1 / t)$.

Cuadrático. Modelo cuya ecuación es $Y = b_0 + (b_1 * t) + (b_2 * t^{**2})$. El modelo cuadrático puede utilizarse para modelar una serie que “despega” o una serie que se amortigua.

Cúbico. Modelo definido por la ecuación $Y = b_0 + (b_1 * t) + (b_2 * t^{**2}) + (b_3 * t^{**3})$.

Potencia. Modelo cuya ecuación es $Y = b_0 * (t^{**b_1})$ ó $\ln(Y) = \ln(b_0) + (b_1 * \ln(t))$.

Compuesto. Modelo cuya ecuación es $Y = b_0 * (b_1^{**t})$ ó $\ln(Y) = \ln(b_0) + (\ln(b_1) * t)$.

curva-S. Modelo cuya ecuación es $Y = e^{**(b_0 + (b_1/t))}$ ó $\ln(Y) = b_0 + (b_1/t)$.

Logística. Modelo cuya ecuación es $Y = 1 / (1/u + (b_0 * (b_1^{**t})))$ o $\ln(1/Y - 1/u) = \ln(b_0 + (\ln(b_1) * t))$ donde u es el valor del límite superior. Tras seleccionar Logístico, especifique un valor para el límite superior que será usado en la ecuación de regresión. El valor debe ser un número positivo mayor que el valor máximo de la variable dependiente.

Crecimiento. Modelo cuya ecuación es $Y = e^{b_0 + (b_1 * t)}$ ó $\ln(Y) = b_0 + (b_1 * t)$.

Exponencial. Modelo cuya ecuación es $Y = b_0 * (e^{b_1 * t})$ ó $\ln(Y) = \ln(b_0) + (b_1 * t)$.

Estimación curvilínea: Guardar

Figura 27-6
Cuadro de diálogo Estimación curvilínea: Guardar

Guardar variables. Para cada modelo seleccionado se pueden guardar los valores pronosticados, los residuos (el valor observado de la variable dependiente menos el valor pronosticado por el modelo) y los intervalos de pronóstico (sus límites superior e inferior). En la ventana de resultados, se muestran en una tabla los nombres de las nuevas variables y las etiquetas descriptivas.

Pronosticar casos. Si se selecciona como variable independiente Tiempo en lugar de una variable del archivo de datos de trabajo, se puede especificar un período de predicción que vaya más allá del final de la serie temporal. Puede elegir una de las siguientes alternativas:

- **Desde el período de estimación hasta el último caso.** Pronostica los valores para todos los casos del archivo, basándose en los casos del período de estimación. El período de estimación, que se muestra en la parte inferior del cuadro de diálogo, se define con el subcuadro de diálogo Rango de la opción Seleccionar casos en

el menú Datos. Si no se ha definido un período de estimación, se utilizan todos los casos para pronosticar los valores.

- **Predecir hasta.** Predice los valores hasta la fecha especificada, hora o número de observación, basándose en los casos del período de estimación. Se puede utilizar para predecir valores más allá del último caso de la serie temporal. Los cuadros de texto disponibles para especificar el final del período de predicción dependen de las variables de fecha definidas actualmente. Si no existen variables de fecha definidas, se puede especificar el número de la observación (caso) final.

Utilice la opción de Definir fechas en el menú Datos para crear las variables de fecha.

Análisis discriminante

El análisis discriminante resulta útil para las situaciones en las que se desea construir un modelo predictivo para pronosticar el grupo de pertenencia de un caso a partir de las características observadas de cada caso. El procedimiento genera una función discriminante (o, para más de dos grupos, un conjunto de funciones discriminantes) basada en combinaciones lineales de las variables predictoras que proporcionan la mejor discriminación posible entre los grupos. Las funciones se generan a partir de una muestra de casos para los que se conoce el grupo de pertenencia; posteriormente, las funciones pueden ser aplicadas a nuevos casos que dispongan de medidas para las variables predictoras pero de los que se desconozca el grupo de pertenencia.

Nota: La variable de agrupación puede tener más de dos valores. Los códigos de la variable de agrupación han de ser números enteros y es necesario especificar sus valores máximo y mínimo. Los casos con valores fuera de estos límites se excluyen del análisis.

Ejemplo. Por término medio, las personas de los países de zonas templadas consume más calorías por día que las de los trópicos, y una proporción mayor de la población de las zonas templadas vive en núcleos urbanos. Un investigador desea combinar esta información en una función para determinar cómo de bien un individuo es capaz de discriminar entre los dos grupos de países. El investigador considera además que el tamaño de la población y la información económica también pueden ser importantes. El análisis discriminante permite estimar los coeficientes de la función discriminante lineal, que tiene el aspecto de la parte derecha de una ecuación de regresión lineal múltiple. Es decir, utilizando los coeficientes a , b , c y d , la función es:

$$D = a * \text{clima} + b * \text{urbanos} + c * \text{población} + d * \text{producto interior bruto per capita}$$

Si estas variables resultan útiles para discriminar entre las dos zonas climáticas, los valores de D serán diferentes para los países templados y para los tropicales. Si se utiliza un método de selección de variables por pasos, quizás no se necesite incluir las cuatro variables en la función.

Estadísticos. Para cada variable: medias, desviaciones típicas, ANOVA univariado. Para cada análisis: M de Box, matriz de correlaciones intra-grupos, matriz de covarianzas intra-grupos, matriz de covarianzas de los grupos separados, matriz de covarianzas total. Para cada función discriminante canónica: autovalores, porcentaje de varianza, correlación canónica, lambda de Wilks, chi-cuadrado. Para cada paso: probabilidades previas, coeficientes de la función de Fisher, coeficientes de función no tipificados, lambda de Wilks para cada función canónica.

Datos. La variable de agrupación debe tener un número limitado de categorías distintas, codificadas como números enteros. Las variables independientes que sean nominales deben ser recodificadas a variables dummy o de contraste.

Supuestos. Los casos deben ser independientes. Las variables predictoras deben tener una distribución normal multivariada y las matrices de varianzas-covarianzas intra-grupos deben ser iguales en todos los grupos. Se asume que la pertenencia al grupo es mutuamente exclusiva (es decir, ningún caso pertenece a más de un grupo) y exhaustiva de modo colectivo (es decir, todos los casos son miembros de un grupo). El procedimiento es más efectivo cuando la pertenencia al grupo es una variable verdaderamente categórica; si la pertenencia al grupo se basa en los valores de una variable continua (por ejemplo, un cociente de inteligencia alto respecto a uno bajo), deberá considerar el uso de la regresión lineal para aprovechar la información más rica ofrecida por la propia variable continua.

Figura 28-1
Resultados del análisis discriminante

Autovalores				
Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	1.002	100.0	100.0	.707

Lambda de Wilks

Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1	.499	31.934	4	.000

Matriz de estructura

	Función
	1
Ingesta diaria de calorías	.986
Log(10) de PIB_CAP	.790
Habitantes en ciudades (%)	.488
Log(10) de POBLAC	.082

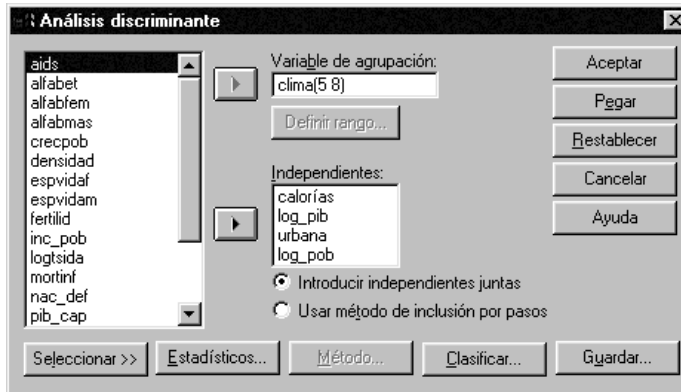
Funciones en los centroides de los grupos

	Función
Clima predominante	1
tropical	-.869
templado	1.107

Para obtener un análisis discriminante

- ▶ Elija en los menús:
 - Analizar
 - Clasificar
 - Discriminante...

Figura 28-2
Cuadro de diálogo Análisis discriminante

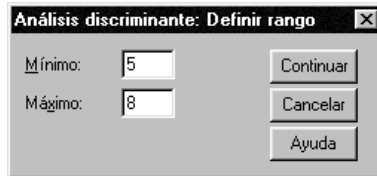


- ▶ Seleccione una variable de agrupación con valores enteros y pulse en Definir rango para especificar las categorías de interés.
- ▶ Seleccione las variables independientes o predictoras. Si la variable de agrupación no tiene valores enteros, la opción Recodificación automática en el menú Transformar creará una que los tenga.
- ▶ Seleccione el método de introducción de las variables independientes.
 - **Introducir independientes juntas.** Método de entrada forzada. Todas las variables independientes que satisfacen el criterio de tolerancia se introducen de forma simultánea.
 - **Usar método de inclusión por pasos.** Utiliza el análisis por pasos para controlar la entrada y la salida de variables.

Si lo desea, tiene la posibilidad de seleccionar casos mediante una variable de selección.

Análisis discriminante: Definir rango

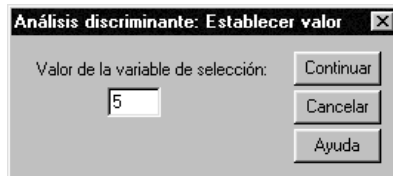
Figura 28-3
Cuadro de diálogo Análisis discriminante: Definir rango



Especifique los valores mínimo y máximo de la variable de agrupación para el análisis. Los casos con valores fuera de este rango no se utilizan en el análisis discriminante, pero sí se clasifican en uno de los grupos existentes a partir de los resultados que obtengan en el análisis. Los valores mínimo y máximo deben ser números enteros.

Análisis discriminante: Seleccionar casos

Figura 28-4
Cuadro de diálogo Análisis discriminante: Establecer valor

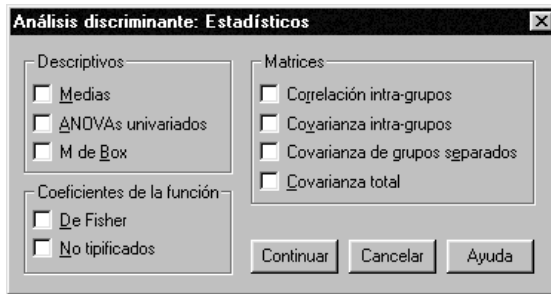


Para seleccionar los casos para el análisis, pulse en Seleccionar, en el cuadro de diálogo principal, elija una variable de selección y, a continuación, pulse en Valor para introducir un entero como valor de selección. Sólo se utilizan los casos con ese valor en la variable de selección para derivar las funciones discriminantes.

Tanto para los casos seleccionados como para los no seleccionados se generan resultados de clasificaciones y estadísticos. De esta forma, se ofrece un mecanismo para clasificar casos nuevos basados en datos previos o para dividir los datos en subconjuntos de contraste y comprobación para realizar procedimientos de validación en el modelo generado.

Análisis discriminante: Estadísticos

Figura 28-5
Cuadro de diálogo Análisis discriminante: Estadísticos



Descriptivos. Las opciones disponibles son: Medias (que incluye las desviaciones típicas), ANOVAs univariados y prueba M de Box.

- **Medias (Análisis discriminante).** Muestra la media y desviación típica totales y las medias y desviaciones típicas de grupo, para las variables independientes.
- **ANOVAs univariados (Análisis discriminante).** Realiza un análisis de varianza de un factor sobre la igualdad de las medias de grupo para cada variable independiente.
- **M de Box.** Contraste sobre la igualdad de las matrices de covarianza de los grupos. Para tamaños de muestra suficientemente grandes, un valor de p no significativo quiere decir que no hay evidencia suficiente de que las matrices difieran. Esta prueba es sensible a las desviaciones de la normalidad multivariada.

Coeficientes de la función. Las opciones disponibles son: Coeficientes de clasificación de Fisher y Coeficientes no tipificados.

- **de Fisher (Análisis discriminante).** Muestra los coeficientes de la función de clasificación de Fisher que pueden utilizarse directamente para la clasificación. Se obtiene un conjunto de coeficientes para cada grupo, y se asigna un caso al grupo para el que tiene una mayor puntuación discriminante.
- **Sin estandarizar (Análisis discriminante).** Muestra los coeficientes de la función discriminante sin estandarizar (los coeficientes brutos).

Matrices. Las matrices de coeficientes disponibles para las variables independientes son las de: Correlación intra-grupos, Covarianza intra-grupos, Covarianza de grupos separados y Covarianza total.

- **Correlación intra-grupos.** Muestra la matriz de correlaciones intra-grupos combinada, que se obtiene de promediar las matrices de covarianza individuales para todos los grupos antes de calcular las correlaciones.
- **Covarianza intra-grupos.** Muestra la matriz de covarianza intra-grupos combinada, la cual puede diferir de la matriz de covarianza total. La matriz se obtiene de promediar, para todos los grupos, las matrices de covarianza individuales.
- **Covarianza de grupos separados.** Muestra las matrices de covarianza de cada grupo por separado.
- **Covarianza total.** Muestra la matriz de covarianza para todos los casos, como si fueran una única muestra.

Análisis discriminante: Método de inclusión por pasos

Figura 28-6

Cuadro de diálogo Análisis discriminante: Método de inclusión por pasos

Método. Seleccione el estadístico que se va a utilizar para introducir o eliminar nuevas variables. Las alternativas disponibles son la lambda de Wilks, la varianza no explicada, la distancia de Mahalanobis, la menor razón F y la V de Rao. Con la V de Rao se puede especificar el incremento mínimo de V para introducir una variable.

- **lambda de Wilks.** Método para la selección de variables por pasos del análisis discriminante que escoge las variables para su introducción en la ecuación basándose en cuánto contribuyen a disminuir la lambda de Wilks. En cada paso se introduce la variable que minimiza la lambda de Wilks global.

- **Varianza no explicada.** En cada paso se introduce la variable que minimiza la suma de la variación no explicada entre los grupos.
- **Distancia de Mahalanobis.** Medida de cuánto difieren los valores de un caso en las variables independientes respecto al promedio para todos los casos. Una distancia de Mahalanobis grande identificará a un caso que tenga valores extremos en una o más de las variables independientes.
- **Menor razón F.** Método para la selección de variables en los análisis por pasos que se basa en maximizar la razón F, calculada a partir de la distancia de Mahalanobis entre los grupos.
- **V de Rao.** Medida de las diferencias entre las medias de los grupos. También llamada traza de Lawley-Hotelling. En cada paso, se introduce la variable que maximiza el incremento en la V de Rao. Después de seleccionar esta opción, introduzca el valor mínimo que debe tener una variable para entrar en el modelo.

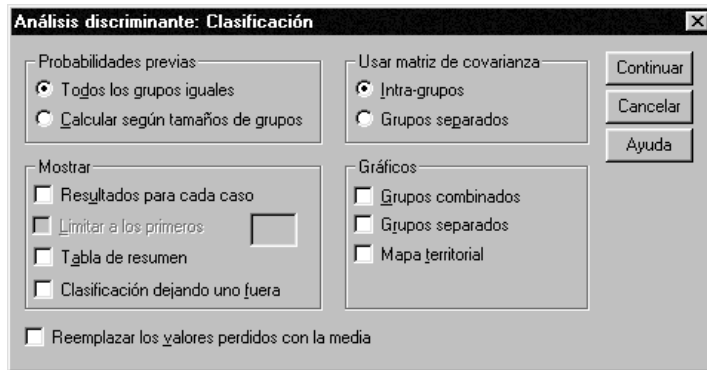
Criterios. Las alternativas disponibles son Usar valor de F y Usar probabilidad de F. Introduzca valores para introducir y eliminar variables.

- **Usar valor de F.** Una variable se introduce en el modelo si su valor de F es mayor que el valor de entrada, y se elimina si su valor de F es menor que el valor de salida. La entrada debe ser mayor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, disminuya el valor de entrada. Para eliminar más variables del modelo, eleve el valor de salida.
- **Usar probabilidad de F.** Una variable se introduce en el modelo si el nivel de significación de su valor de F es menor que el valor de entrada, y se elimina si el nivel de significación de su valor de F es mayor que el valor de salida. La entrada debe ser menor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, eleve el valor de entrada. Para eliminar más variables del modelo, disminuya el valor de salida.

Mostrar. Resumen de los pasos muestra los estadísticos para todas las variables después de cada paso; F para distancias por parejas muestra una matriz de razones F por parejas para cada pareja de grupos.

Análisis discriminante: Clasificar

Figura 28-7
Cuadro de diálogo Análisis discriminante: Clasificar



Probabilidades previas. Estos valores se utilizan para la clasificación. Puede especificar que las probabilidades previas sean iguales para todos los grupos o dejar que los tamaños de grupo observados en la muestra determinen las probabilidades de la pertenencia al grupo.

Mostrar. Las opciones de presentación disponibles son: Resultados por casos, Tabla de resumen y Clasificación dejando uno fuera.

- **Resultados para cada caso.** Se muestran, para cada caso, los códigos del grupo real de pertenencia, el grupo pronosticado, las probabilidades posteriores y las puntuaciones discriminantes.
- **Tabla de resumen.** Número de casos correcta e incorrectamente asignados a cada uno de los grupos, basándose en el análisis discriminante. En ocasiones recibe el nombre de la “Matriz de Confusión”.
- **Clasificación dejando uno fuera.** Se clasifica cada caso del análisis mediante la función derivada a partir de todos los casos, excepto el propio caso. También conocido como método-U.

Reemplazar los valores perdidos con la media. Seleccione esta opción para sustituir la media de una variable independiente para un valor perdido sólo durante la fase de clasificación.

Usar matriz de covarianzas. Existe la opción de clasificar los casos utilizando una matriz de covarianzas intra-grupos o una matriz de covarianzas de los grupos separados.

- **Intra-grupos.** Se utiliza la matriz de covarianza intra-grupos combinada para clasificar los casos.
- **Grupos separados.** Para la clasificación se utilizan las matrices de covarianza de los grupos separados. Dado que la clasificación se basa en las funciones discriminantes y no en las variables originales, esta opción no siempre es equivalente a la discriminación cuadrática.

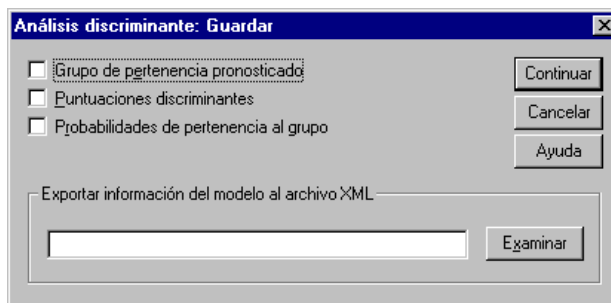
Diagramas. Las opciones de gráficos disponibles son: Grupos combinados, Grupos separados y Mapa territorial.

- **Grupos combinados.** Crea un diagrama de dispersión, con todos los grupos, de los valores en las dos primeras funciones discriminantes. Si sólo hay una función, en su lugar se muestra un histograma.
- **Grupos separados.** Crea diagramas de dispersión, de los grupos por separado, para los valores en las dos primeras funciones discriminantes. Si sólo hay una función se muestra un histograma en su lugar.
- **Mapa territorial.** Gráfico de las fronteras utilizadas para clasificar los casos en grupos a partir de los valores en las funciones. Los números corresponden a los grupos en los que se clasifican los casos. La media de cada grupo se indica mediante un asterisco situado dentro de sus fronteras. No se mostrará el mapa si sólo hay una función discriminante.

Análisis discriminante: Guardar

Figura 28-8

Cuadro de diálogo Análisis discriminante: Guardar nuevas variables



Es posible añadir variables nuevas al archivo de datos activo. Las opciones disponibles son las de grupo de pertenencia pronosticado (una única variable), puntuaciones discriminantes (una variable para cada función discriminante en la solución) y probabilidades de pertenencia al grupo según las puntuaciones discriminantes (una variable para cada grupo).

También se puede exportar información del modelo al archivo especificado en formato XML (PMML). *SmartScore* y la versión de servidor de *SPSS* (un producto independiente) pueden utilizar este archivo de modelo para aplicar la información de modelo a otros archivos de datos con fines de puntuación.

análisis factorial

El análisis factorial intenta identificar variables subyacentes, o **factores**, que expliquen la configuración de las correlaciones dentro de un conjunto de variables observadas. El análisis factorial se suele utilizar en la reducción de los datos para identificar un pequeño número de factores que explique la mayoría de la varianza observada en un número mayor de variables manifiestas. También puede utilizarse para generar hipótesis relacionadas con los mecanismos causales o para inspeccionar las variables para análisis subsiguientes (por ejemplo, para identificar la colinealidad antes de realizar un análisis de regresión lineal).

El procedimiento de análisis factorial ofrece un alto grado de flexibilidad:

- Existen siete métodos de extracción factorial disponibles.
- Existen cinco métodos de rotación disponibles, entre ellos el oblimin directo y el promax para rotaciones no ortogonales.
- Existen tres métodos disponibles para calcular las puntuaciones factoriales; y las puntuaciones pueden guardarse como variables para análisis adicionales.

Ejemplo. ¿Qué actitudes subyacentes hacen que las personas respondan a las preguntas de una encuesta política de la manera en que lo hacen? Examinando las correlaciones entre los elementos de la encuesta se deduce que hay una superposición significativa entre los diversos subgrupos de elementos (las preguntas sobre los impuestos tienden a estar correlacionadas entre sí, las preguntas sobre temas militares también están correlacionadas entre sí, y así sucesivamente). Con el análisis factorial, se puede investigar el número de factores subyacentes y, en muchos casos, se puede identificar lo que los factores representan conceptualmente. Adicionalmente, se pueden calcular las puntuaciones factoriales para cada encuestado, que pueden utilizarse en análisis subsiguientes. Por ejemplo, es posible construir un modelo de regresión logística para predecir el comportamiento de voto basándose en las puntuaciones factoriales.

Estadísticos. Para cada variable: número de casos válidos, media y desviación típica. Para cada análisis factorial: matriz de correlaciones de variables, incluidos niveles de significación, determinante, inversa; matriz de correlaciones reproducida, que incluye anti-imagen; solución inicial (comunalidades, autovalores y porcentaje de varianza explicada); KMO (medida de la adecuación muestral de Kaiser-Meyer-Olkin) y prueba de esfericidad de Bartlett; solución sin rotar, que incluye saturaciones factoriales, comunalidades y autovalores; solución rotada, que incluye la matriz de configuración rotada y la matriz de transformación; para rotaciones oblicuas: Para las rotaciones oblicuas: las matrices de estructura y de configuración rotadas; matriz de coeficientes para el cálculo de las puntuaciones factoriales y matriz de covarianzas entre los factores. Gráficos: gráfico de sedimentación y gráfico de las saturaciones de los dos o tres primeros factores.

Datos. Las variables deben ser cuantitativas a nivel de **intervalo** o de **razón**. Los datos categóricos (como la religión o el país de origen) no son adecuados para el análisis factorial. Los datos para los cuales razonablemente se pueden calcular los coeficientes de correlación de Pearson, deberían ser adecuados para el análisis factorial.

Supuestos. Los datos deben tener una distribución normal bivariada para cada pareja de variables y las observaciones deben ser independientes. El modelo de análisis factorial especifica que las variables vienen determinadas por los factores comunes (los factores estimados por el modelo) y por factores únicos (los cuales no se superponen entre las distintas variables observadas); las estimaciones calculadas se basan en el supuesto de que ningún factor único está correlacionado con los demás, ni con los factores comunes.

Figura 29-1
Resultados del análisis factorial

Estadísticos descriptivos

	Media	Desviación típica	N del análisis
Esperanza de vida femenina	72.63	8.03	24
Mortalidad infantil (muertes por 1000 nacimientos vivos)	36.892	31.104	24
Alfabetización (%)	85.54	19.29	24
Tasa de natalidad (por 1.000 habitantes)	22.625	8.747	24
Número promedio de hijos	2.929	1.277	24
Habitantes en ciudades (%)	61.38	20.61	24
Log(10) de PIB_CAP	3.4876	.6474	24
Aumento de la población (% anual)	1.456	.855	24
Tasa Nacimientos/Defunciones	3.1376	1.7112	24
Tasa de mortalidad (por 1.000 habitantes)	8.04	2.18	24
Log(10) de POBLAC	4.3088	.7324	24

Comunalidades

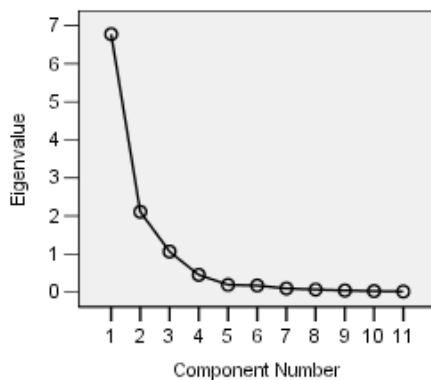
	Inicial	Extracción
Esperanza de vida femenina	1.000	.950
Mortalidad infantil (muertes por 1000 nacimientos vivos)	1.000	.949
Alfabetización (%)	1.000	.842
Tasa de natalidad (por 1.000 habitantes)	1.000	.975
Número promedio de hijos	1.000	.891
Habitantes en ciudades (%)	1.000	.691
Log(10) de PIB_CAP	1.000	.884
Aumento de la población (% anual)	1.000	.974
Tasa Nacimientos/Defunciones	1.000	.956
Tasa de mortalidad (por 1.000 habitantes)	1.000	.776
Log(10) de POBLAC	1.000	.652

Método de extracción: Análisis de Componentes principales.

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	6.960	63.275	63.275	6.960	63.275	63.275	6.910	62.819	62.819
2	2.579	23.446	86.721	2.579	23.446	86.721	2.629	23.902	86.721
3	.612	5.562	92.283						
4	.361	3.281	95.564						
5	.218	1.985	97.549						
6	.130	1.183	98.732						
7	6.131E-02	.557	99.290						
8	3.775E-02	.343	99.633						
9	2.610E-02	.237	99.870						
10	9.840E-03	8.946E-02	99.959						
11	4.461E-03	4.055E-02	100.000						

Método de extracción: Análisis de Componentes principales.

**Matriz de componentes rotados**

	Componente	
	1	2
Esperanza de vida femenina	-.964	
Mortalidad infantil (muertes por 1000 nacimientos vivos)	.962	
Log(10) de PIB_CAP	-.937	
Tasa de natalidad (por 1.000 habitantes)	.910	.384
Número promedio de hijos	.906	.264
Alfabetización (%)	-.898	
Aumento de la población (% anual)	.830	.535
Habitantes en ciudades (%)	-.789	.263
Tasa de mortalidad (por 1.000 habitantes)		-.876
Tasa Nacimientos/Defunciones	.464	.861
Log(10) de POBLAC	.434	-.681

Método de extracción: Análisis de componentes principales.

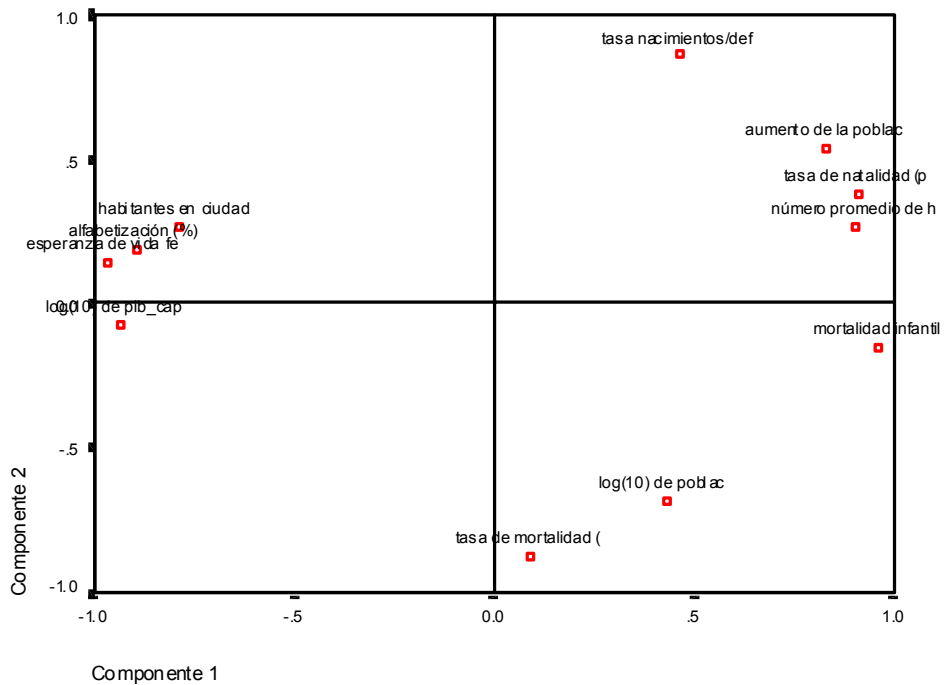
Método de rotación: Normalización Varimax con Kaiser.

Matriz de transformación de las componentes

Componente	1	2
1	.994	.107
2	-.107	.994

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.

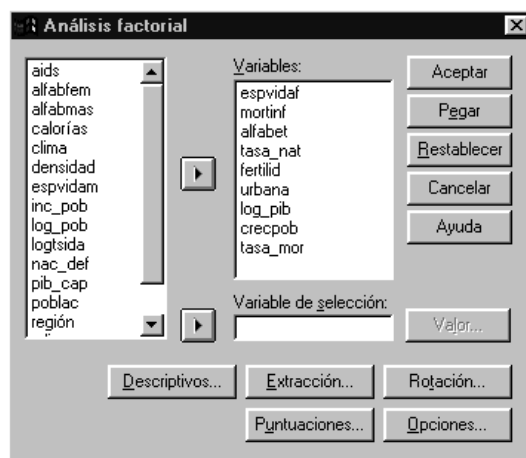
Gráfico de componentes en espacio rotado



Para obtener un análisis factorial

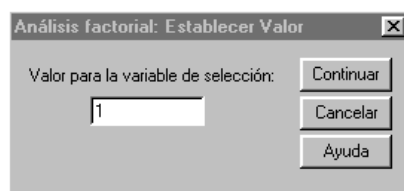
- ▶ Elija en los menús:
 - Analizar
 - Reducción de datos
 - Análisis factorial...
- ▶ Seleccione las variables para el análisis factorial.

Figura 29-2
Cuadro de diálogo Análisis factorial



Selección de casos en el análisis factorial

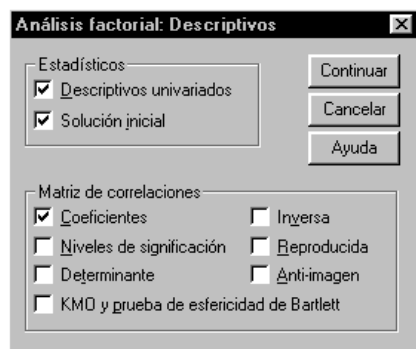
Figura 29-3
Cuadro de diálogo Análisis factorial: Establecer valor



Para seleccionar los casos para el análisis, elija una variable de selección y pulse en Valor para introducir un entero como el valor de selección. En el análisis factorial, sólo se usarán los casos con ese valor para la variable de selección.

Análisis factorial: Descriptivos

Figura 29-4
Cuadro de diálogo Análisis factorial: Descriptivos



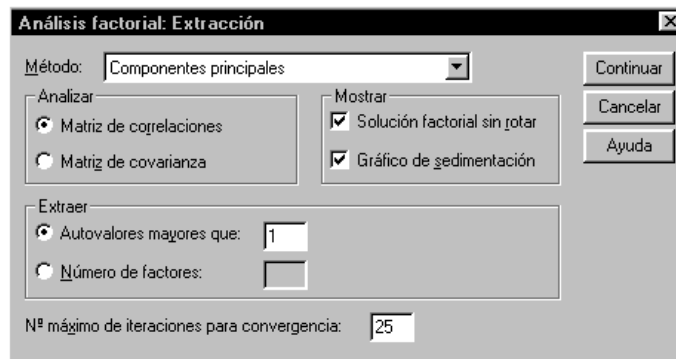
Estadísticos. Los descriptivos univariados incluyen la media, la desviación típica y el número de casos válidos para cada variable. La solución inicial muestra las comunalidades iniciales, los autovalores y el porcentaje de varianza explicada.

Matriz de correlaciones. Las opciones disponibles son: coeficientes, niveles de significación, determinante, inversa, reproducida, anti-imagen y KMO y prueba de esfericidad de Bartlett.

- **KMO y prueba de esfericidad de Bartlett.** La medida de la adecuación muestral de Kaiser-Meyer-Olkin contrasta si las correlaciones parciales entre las variables son pequeñas. La prueba de esfericidad de Bartlett contrasta si la matriz de correlaciones es una matriz identidad, que indicaría que el modelo factorial es inadecuado.
- **Reproducida.** Matriz de correlaciones estimada a partir de la solución factorial. También se muestran las correlaciones residuales (la diferencia entre la correlación observada y la reproducida).
- **Anti-imagen.** La matriz de correlaciones anti-imagen contiene los negativos de los coeficientes de correlación parcial y la matriz de covarianza anti-imagen contiene los negativos de las covarianzas parciales. En un buen modelo factorial la mayoría de los elementos no diagonales deben ser pequeños. En la diagonal de la matriz de correlaciones anti-imagen se muestra la medida de adecuación muestral para esa variable.

Análisis factorial: Extracción

Figura 29-5
Cuadro de diálogo Análisis factorial: Extracción



Método. Permite especificar el método de extracción factorial. Los métodos disponibles son: Componentes principales, Mínimos cuadrados no ponderados, Mínimos cuadrados generalizados, Máxima verosimilitud, factorización de Ejes principales, factorización Alfa y factorización Imagen.

- **Análisis de componentes principales (Análisis factorial).** Método para la extracción de factores utilizada para formar combinaciones lineales independientes de las variables observadas. La primera componente tiene la varianza máxima. Las componentes sucesivas explican progresivamente proporciones menores de la varianza y no están correlacionadas las unas con las otras. El análisis de componentes principales se utiliza para obtener la solución factorial inicial. Puede utilizarse cuando una matriz de correlaciones es singular.
- **Método de mínimos cuadrados no ponderados (Análisis factorial).** Método de extracción factorial que minimiza la suma de los cuadrados de las diferencias entre las matrices de correlaciones observada y reproducida, ignorando las diagonales.
- **Método de Mínimos cuadrados generalizados (Análisis factorial).** Método de extracción de factores que minimiza la suma de los cuadrados de las diferencias entre las matrices de correlación observada y reproducida. Las correlaciones se ponderan por el inverso de su unicidad, de manera que las variables que tengan un valor alto de unicidad reciban un peso menor que aquellas que tengan un valor bajo de unicidad.

- **Método de máxima verosimilitud (Análisis factorial).** Método de extracción factorial que proporciona las estimaciones de los parámetros que con mayor probabilidad han producido la matriz de correlaciones observada, si la muestra procede de una distribución normal multivariada. Las correlaciones se ponderan por el inverso de la unicidad de las variables, y se emplea un algoritmo iterativo.
- **Factorización de ejes principales (Análisis factorial).** Método para la extracción de factores que parte de la matriz de correlaciones original con los cuadrados de los coeficientes de correlación múltiple insertados en la diagonal principal como estimaciones iniciales de las comunalidades. Las saturaciones factoriales resultantes se utilizan para estimar de nuevo las comunalidades y reemplazan a las estimaciones previas en la diagonal de la matriz. Las iteraciones continúan hasta que el cambio en las comunalidades, de una iteración a la siguiente, satisfaga el criterio de convergencia para la extracción.
- **Alfa (Análisis factorial).** Método de extracción factorial que considera a las variables incluidas en el análisis como una muestra del universo de las variables posibles. Este método maximiza el Alfa de Cronbach para los factores.
- **Factorización imagen (Análisis factorial).** Método para la extracción de factores, desarrollado por Guttman y basado en la teoría de las imágenes. La parte común de una variable, llamada la imagen parcial, se define como su regresión lineal sobre las restantes variables, en lugar de ser una función de los factores hipotéticos.

Analizar. Permite especificar o una matriz de correlaciones o una matriz de covarianzas.

- **Matriz de correlaciones.** Es útil si las variables de su análisis se miden sobre escalas distintas.
- **Matriz de covarianzas.** Es útil si se desea aplicar el análisis factorial a varios grupos con distintas varianzas para cada variable.

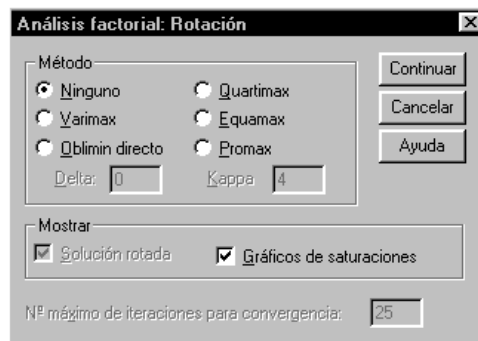
Extraer. Se pueden retener todos los factores cuyos autovalores excedan un valor especificado o retener un número específico de factores.

Mostrar. Permite solicitar la solución factorial sin rotar y el gráfico de sedimentación de los autovalores.

- **Solución factorial sin rotar (Análisis factorial).** Muestra las saturaciones factoriales sin rotar (la matriz de configuración factorial), las comunalidades y los autovalores de la solución factorial.
 - **Gráfico de sedimentación.** Gráfico de la varianza asociada a cada factor. Se utiliza para determinar cuántos factores deben retenerse. Típicamente el gráfico muestra la clara ruptura entre la pronunciada pendiente de los factores más importantes y el descenso gradual de los restantes (los sedimentos).
- Nº máximo de iteraciones para convergencia.** Permite especificar el número máximo de pasos que el algoritmo puede seguir para estimar la solución.

Análisis factorial: Rotación

Figura 29-6
Cuadro de diálogo Análisis factorial: Rotación



Método. Permite seleccionar el método de rotación factorial. Los métodos disponibles son: varimax, equamax, quartimax, oblimin directo y promax.

- **Método varimax (Análisis factorial).** Método de rotación ortogonal que minimiza el número de variables que tienen saturaciones altas en cada factor. Simplifica la interpretación de los factores.
- **Criterio Oblimin directo.** Método para la rotación oblicua (no ortogonal). Cuando delta es igual a cero (el valor por defecto) las soluciones son las más oblicuas. A medida que delta se va haciendo más negativo, los factores son menos oblicuos. Para anular el valor por defecto 0 para delta, introduzca un número menor o igual que 0,8.

- **Método quartimax (Análisis factorial).** Método de rotación que minimiza el número de factores necesarios para explicar cada variable. Simplifica la interpretación de las variables observadas.
- **Método equamax (Análisis factorial).** Método de rotación que es combinación del método varimax, que simplifica los factores, y el método quartimax, que simplifica las variables. Se minimiza tanto el número de variables que saturan alto en un factor como el número de factores necesarios para explicar una variable.
- **Rotación promax (Análisis factorial).** Rotación oblicua que permite que los factores estén correlacionados. Puede calcularse más rápidamente que una rotación oblimin directa, por lo que es útil para conjuntos de datos grandes.

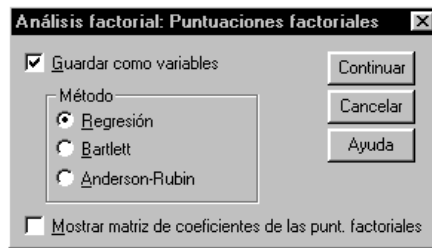
Mostrar. Permite incluir los resultados de la solución rotada, así como los gráficos de las saturaciones para los dos o tres primeros factores.

- **Solución rotada (Análisis factorial).** Debe seleccionarse un método de rotación para obtener la solución rotada. Para las rotaciones ortogonales, se muestran la matriz de configuración rotada y la matriz de transformación. Para las rotaciones oblicuas, se muestran las matrices de estructura, de configuración y de correlaciones entre los factores.
 - **Diagrama de las saturaciones factoriales.** Representación tridimensional de las saturaciones factoriales para los tres primeros factores. Para una solución de dos factores, se representa un diagrama bidimensional. No se muestra el gráfico si sólo se extrae un factor. Si se solicita la rotación, los diagramas representan las soluciones rotadas.
- Nº máximo de iteraciones para convergencia.** Permite especificar el número máximo de pasos que el algoritmo puede seguir para llevar a cabo la rotación.

Análisis factorial: Puntuaciones factoriales

Figura 29-7

Cuadro de diálogo Análisis factorial: Puntuaciones factoriales



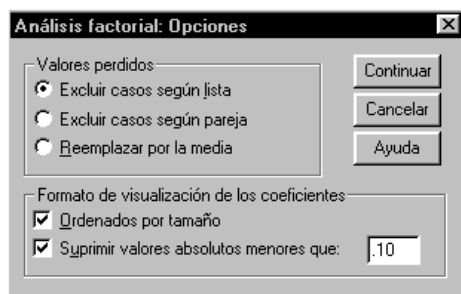
Guardar como variables. Crea una nueva variable para cada factor en la solución final. Seleccione uno de los siguientes métodos alternativos para calcular las puntuaciones factoriales: Regresión, Bartlett o Anderson-Rubin.

- **Método de regresión (Análisis factorial).** Método para estimar los coeficientes de las puntuaciones factoriales. Las puntuaciones resultantes tienen de media 0 y varianza igual al cuadrado de la correlación múltiple entre las puntuaciones factoriales estimadas y los valores factoriales verdaderos. Las puntuaciones pueden estar correlacionadas incluso cuando los factores son ortogonales.
- **Puntuaciones de Bartlett.** Método para estimar los coeficientes de las puntuaciones factoriales. Las puntuaciones resultantes tienen una media de 0. Se minimiza la suma de cuadrados de los factores únicos sobre el rango de las variables.
- **Método de Anderson-Rubin (Análisis factorial).** Método para calcular los coeficientes para las puntuaciones factoriales; es una modificación del método de Bartlett, que asegura la ortogonalidad de los factores estimados. Las puntuaciones resultantes tienen una media 0, una desviación típica de 1 y no correlacionan entre sí.

Mostrar matriz de coeficientes de las puntuaciones factoriales. Muestra los coeficientes por los cuales se multiplican las variables para obtener puntuaciones factoriales. También muestra las correlaciones entre las puntuaciones factoriales.

Análisis factorial: Opciones

Figura 29-8
Cuadro de diálogo Análisis factorial: Opciones



Valores perdidos. Permite especificar el tratamiento que reciben los valores perdidos. Las alternativas disponibles son: Excluir casos **según lista**, Excluir casos **según pareja** y Reemplazar por la media.

Formato de presentación de los coeficientes. Permite controlar aspectos de las matrices de resultados. Los coeficientes se ordenan por tamaño y se suprimen aquéllos cuyos valores absolutos sean menores que el valor especificado.

Selección de procedimientos para la conglomeración

Los análisis de conglomerados se pueden realizar mediante los procedimientos de análisis de conglomerados en dos fases, jerárquico o de K-medias. Cada uno de estos procedimientos emplea un algoritmo distinto en la creación de conglomerados y contiene opciones que no están disponibles en los otros.

Análisis de conglomerados en dos fases. En algunas aplicaciones, se puede seleccionar como método el procedimiento Análisis de conglomerados en dos fases. Ofrece una serie de funciones únicas que se detallan a continuación:

- Selección automática del número más apropiado de conglomerados y medidas para la selección de los distintos modelos de conglomerado.
- Posibilidad de crear modelos de conglomerado basados al mismo tiempo en variables categóricas y continuas.
- Posibilidad de guardar el modelo de conglomerados en un archivo XML externo y, a continuación, leer el archivo y actualizar el modelo de conglomerados con datos más recientes.

Asimismo, el procedimiento Análisis de conglomerados en dos fases puede analizar archivos de datos grandes.

Análisis de conglomerados jerárquico. El uso del procedimiento Análisis de conglomerados jerárquico se limita a archivos de datos más pequeños (cientos de objetos por conglomerar) y ofrece una serie de funciones únicas que se detallan a continuación:

- Posibilidad de conglomerar casos o variables.

- Posibilidad de calcular un rango de soluciones posibles y guardar los conglomerados de pertenencia para cada una de dichas soluciones.
- Distintos métodos de formación de conglomerados, transformación de variables y medida de disimilaridad entre conglomerados.

Siempre que todas las variables sean del mismo tipo, el procedimiento Análisis de conglomerados jerárquico podrá analizar variables de intervalo (continuas), de recuento o binarias.

Análisis de conglomerados de K-medias. El uso del procedimiento Análisis de conglomerados de K-medias se limita a datos continuos y requiere que el usuario especifique previamente el número de conglomerados y ofrece una serie de funciones únicas que se detallan a continuación:

- Posibilidad de guardar las distancias desde los centros de los conglomerados hasta los distintos objetos.
- Posibilidad de leer los centros de los conglomerados iniciales y guardar los centros de los conglomerados finales desde un archivo SPSS externo.

Asimismo, el procedimiento Análisis de conglomerados de K-medias puede analizar archivos de datos grandes.

Análisis de conglomerados en dos fases

El procedimiento Análisis de conglomerados en dos fases es una herramienta de exploración diseñada para descubrir las agrupaciones naturales (o conglomerados) de un conjunto de datos que, de otra manera, no sería posible detectar. El algoritmo que emplea este procedimiento incluye varias atractivas funciones que lo hacen diferente de las técnicas de conglomeración tradicionales:

- **Tratamiento de variables categóricas y continuas.** Al suponer que las variables son independientes, es posible aplicar una distribución normal multinomial conjunta en las variables continuas y categóricas.
- **Selección automática del número de conglomerados.** Mediante la comparación de los valores de un criterio de selección del modelo para diferentes soluciones de conglomeración, el procedimiento puede determinar automáticamente el número óptimo de conglomerados.
- **Escalabilidad.** Mediante la construcción de un árbol de características de conglomerados (CF) que resume los registros, el algoritmo en dos fases puede analizar archivos de datos de gran tamaño.

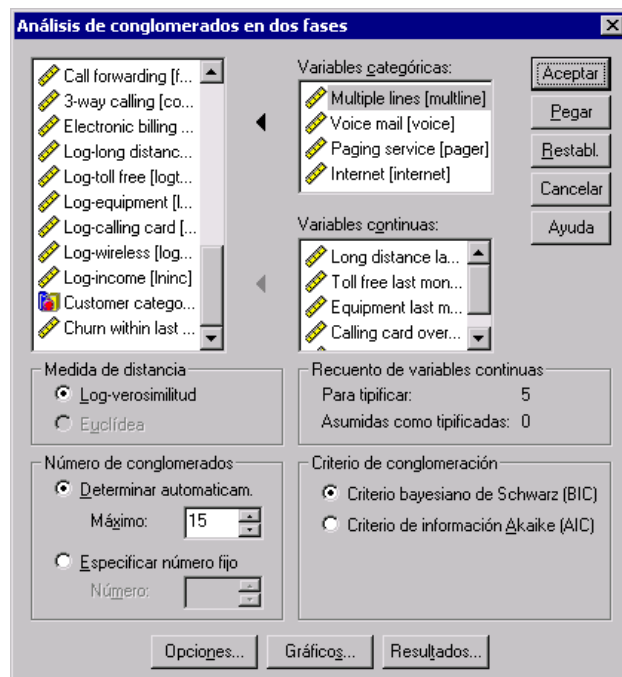
Ejemplo. Las empresas minoristas y de venta de productos para el consumidor suelen aplicar técnicas de conglomeración a los datos que describen los hábitos de consumo, sexo, edad, nivel de ingresos, etc. de los clientes. Estas empresas adaptan sus estrategias de desarrollo de productos y de marketing en función de cada grupo de consumidores para aumentar las ventas y el nivel de fidelidad a la marca.

Estadísticos. El procedimiento genera criterios de información (AIC o BIC) según el número de conglomerados de la solución, las frecuencias de los conglomerados para la conglomeración final y los estadísticos descriptivos por conglomerado para la conglomeración final.

Diagramas. El procedimiento genera gráficos de barras para las frecuencias de los conglomerados, gráficos de sectores para las frecuencias de los conglomerados y gráficos de la importancia de las variables.

Figura 31-1

Cuadro de diálogo *Análisis de conglomerados en dos fases*



Medida de distancia. Esta opción determina cómo se calcula la similaridad entre dos conglomerados.

- **Log-verosimilitud.** La medida de la verosimilitud realiza una distribución de probabilidad entre las variables. Las variables continuas se supone que tienen una distribución normal, mientras que las variables categóricas se supone que son multinomiales. Se supone que todas las variables son independientes.
- **Euclídea.** La medida euclídea es la distancia según una “línea recta” entre dos conglomerados. Sólo se puede utilizar cuando todas las variables son continuas.

Número de conglomerados. Esta opción permite especificar cómo se va a determinar el número de conglomerados.

- **Determinar automáticamente.** El procedimiento determinará automáticamente el número “óptimo” de conglomerados, utilizando el criterio especificado en el grupo Criterio de conglomeración. Si lo desea, introduzca un entero positivo para especificar el número máximo de conglomerados que el procedimiento debe tener en cuenta.
- **Especificar número fijo.** Permite fijar el número de conglomerados de la solución. Introduzca un número entero positivo.

Recuento de variables continuas. Este grupo proporciona un resumen de las especificaciones acerca de la tipificación de variables continuas realizadas en el cuadro de diálogo Opciones. Si desea obtener más información, consulte “Opciones del análisis de conglomerados en dos fases” en p. 491.

Criterio de conglomeración. Esta opción determina cómo el algoritmo de conglomeración determina el número de conglomerados. Se puede especificar tanto el criterio de información bayesiano (BIC) como el criterio de información de Akaike (AIC).

Datos. Este procedimiento trabaja tanto con variables continuas como categóricas. Los casos representan los objetos que se van a conglomerar y las variables representan los atributos en los que se va a basar la conglomeración.

Orden de casos. Observe que el árbol de características de conglomerados y la solución final pueden depender del orden de los casos. Para minimizar los efectos del orden, ordene los casos aleatoriamente. Puede que desee obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada. En situaciones en que esto resulta difícil debido a unos tamaños de archivo demasiado grandes, se pueden sustituir varias ejecuciones por una muestra de casos ordenados con distintos órdenes aleatorios.

Supuestos. La medida de la distancia de la verosimilitud supone que las variables del modelo de conglomerados son independientes. Además, se supone que cada variable continua tiene una distribución normal (de Gauss) y que cada variable categórica tiene una distribución multinomial. Las comprobaciones empíricas internas indican que este procedimiento es bastante robusto frente a las violaciones tanto del supuesto de independencia como de las distribuciones, pero aún así es preciso tener en cuenta hasta qué punto se cumplen estos supuestos.

Utilice el procedimiento Correlaciones bivariadas para comprobar la independencia de dos variables continuas. Utilice el procedimiento Tablas de contingencia para comprobar la independencia de dos variables categóricas. Utilice el procedimiento

Medias para comprobar la independencia existente entre una variable continua y otra categórica. Utilice el procedimiento Explorar para comprobar la normalidad de una variable continua. Utilice el procedimiento Prueba de chi-cuadrado para comprobar si una variable categórica tiene una determinada distribución multinomial.

Para obtener un análisis de conglomerados en dos fases

- ▶ Elija en los menús:
 - Analizar
 - Clasificar
 - Conglomerado en dos fases...

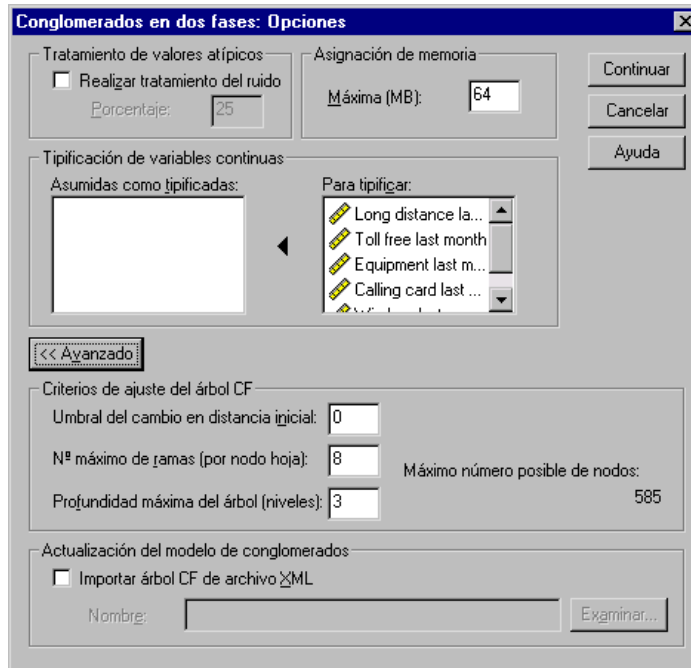
- ▶ Seleccione una o varias variables categóricas o continuas.

Si lo desea, tiene la posibilidad de:

- Ajustar los criterios utilizados para generar los conglomerados.
- Seleccionar los ajustes para el tratamiento del ruido, la asignación de memoria, la tipificación de las variables y la entrada del modelo de conglomerados.
- Solicitar gráficos y tablas opcionales.
- Guardar los resultados del modelo en el archivo de trabajo o en un archivo XML externo.

Opciones del análisis de conglomerados en dos fases

Figura 31-2
Cuadro de diálogo Conglomerados en dos fases: Opciones



Tratamiento de valores atípicos. Este grupo permite tratar los valores atípicos de manera especial durante la conglomeración si se llena el árbol de características de los conglomerados (CF). El árbol CF se considera lleno si no puede aceptar ningún caso más en un nodo hoja y no hay ningún nodo hoja que se pueda dividir.

- Si selecciona el tratamiento del ruido y el árbol CF se llena, se hará volver a crecer después de colocar los casos existentes en hojas poco densas en una hoja de “ruido”. Se considera que una hoja es poco densa si contiene un número de casos inferior a un determinado porcentaje de casos del máximo tamaño de hoja. Tras volver a hacer crecer el árbol, los valores atípicos se colocarán en el árbol CF en caso de que sea posible. Si no es así, se descartarán los valores atípicos.
- Si no selecciona el tratamiento del ruido y el árbol CF se llena, se hará volver a crecer utilizando un umbral del cambio en distancia mayor. Tras la conglomeración final, los valores que no se puedan asignar a un conglomerado

se considerarán como valores atípicos. Al conglomerado de valores atípicos se le asigna un número de identificación de -1 y no se incluirá en el recuento del número de conglomerados.

Asignación de memoria. Este grupo permite especificar la cantidad máxima de memoria en megabytes (MB) que puede utilizar el algoritmo de conglomeración. Si el procedimiento supera este máximo, utilizará el disco para almacenar la información que no se pueda colocar en la memoria. Especifique un número mayor o igual que 4.

- Consulte con el administrador del sistema si desea conocer el valor máximo que puede especificar en su sistema.
- Si este valor es demasiado bajo, es posible que el algoritmo no consiga obtener el número correcto o deseado de conglomerados.

Tipificación de variables. El algoritmo de conglomeración trabaja con variables continuas tipificadas. Todas las variables continuas que no estén tipificadas deben dejarse como variables en la lista Para tipificar. Para ahorrar algún tiempo y trabajo para el ordenador, puede seleccionar todas las variables continuas que ya haya tipificado como variables en la lista Asumidas como tipificadas.

Opciones avanzadas

Criterios de ajuste del árbol CF. Los siguientes ajustes del algoritmo de conglomeración se aplican específicamente al árbol de características de conglomerados (CF) y deberán cambiarse con cuidado:

- **Umbral del cambio en distancia inicial.** Éste es el umbral inicial que se utiliza para hacer crecer el árbol CF. Si se ha insertado una determinada hoja en el árbol CF que produciría una densidad inferior al umbral, la hoja no se dividirá. Si la densidad supera el umbral, se dividirá la hoja.
- **Nº máximo de ramas (por nodo hoja).** Número máximo de nodos filiales que puede tener una hoja.
- **Máxima profundidad de árbol.** Número máximo de niveles que puede tener un árbol CF.
- **Máximo número posible de nodos.** Indica el número máximo de nodos del árbol CF que puede generar potencialmente el procedimiento, de acuerdo con la función $(b^{d+1} - 1) / (b - 1)$, donde b es el número máximo de ramas y d es la profundidad máxima del árbol. Tenga en cuenta que un árbol CF excesivamente grande

puede agotar los recursos del sistema y afectar negativamente al rendimiento del procedimiento. Como mínimo, cada nodo requiere 16 bytes.

Actualización del modelo de conglomerados. Este grupo permite importar y actualizar un modelo de conglomerados generado en un análisis anterior. El archivo de entrada contiene el árbol CF en formato XML. A continuación, se actualizará el modelo con los datos existentes en el archivo activo. Debe seleccionar los nombres de variable en el cuadro de diálogo principal en el mismo orden en que se especificaron en el análisis anterior. El archivo XML permanecerá inalterado, a no ser que escriba específicamente la nueva información del modelo en el mismo nombre de archivo. Si desea obtener más información, consulte “Resultados de análisis de conglomerados en dos fases” en p. 495.

Si se ha especificado una actualización del modelo de conglomerados, se utilizarán las opciones pertenecientes a la generación del árbol CF que se especificaron para el modelo original. Concretamente, se utilizarán los ajustes del modelo guardado acerca de la medida de distancia, el tratamiento del ruido, la asignación de memoria y los criterios de ajuste del árbol CF, por lo que se ignorarán todos los ajustes de estas opciones que se hayan especificado en los cuadros de diálogo.

Nota: Al realizar una actualización del modelo de conglomerados, el procedimiento supone que ninguno de los casos seleccionados en el archivo de datos de trabajo se utilizaron para crear el modelo de conglomerados original. El procedimiento también supone que los casos utilizados en la actualización del modelo proceden de la misma población que los casos utilizados para crear el modelo; es decir, se supone que las medias y las varianzas de las variables continuas y los niveles de las variables categóricas son los mismos en ambos conjuntos de casos. Si los conjuntos de casos “nuevo” y “antiguo” proceden de poblaciones heterogéneas, deberá ejecutar el procedimiento Análisis de conglomerados en dos fases para los conjuntos combinados de casos para obtener los resultados óptimos.

Gráficos de análisis de conglomerados en dos fases

Figura 31-3
Cuadro de diálogo Conglomerados en dos fases: Gráficos

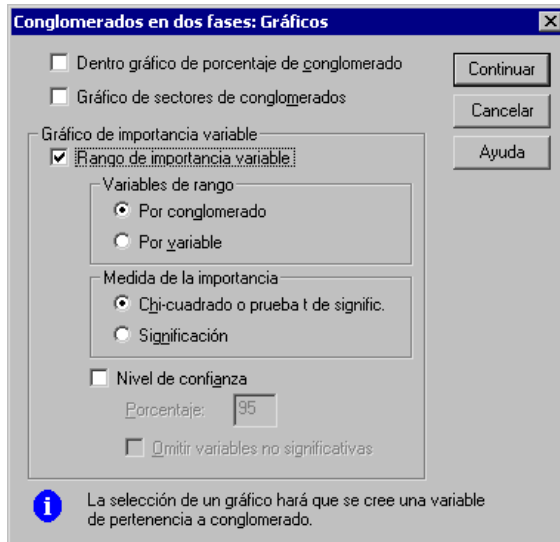


Gráfico del porcentaje intra-conglomerado. Muestra los gráficos que indican la variación dentro del conglomerado de cada variable. Para cada variable categórica, se genera un gráfico de barras agrupado, mostrando la frecuencia de la categoría por identificador de conglomerado. Para cada variable continua, se genera un gráfico de barras de error, indicando las barras de error por identificador de conglomerado.

Gráfico de sectores de los conglomerados. Muestra un gráfico de sectores que muestra el porcentaje y las frecuencias de observaciones correspondientes a cada conglomerado.

Gráfico de la importancia de las variables. Muestra varios gráficos diferentes que indican la importancia de cada variable dentro de cada conglomerado. Los resultados se ordenan según el nivel de importancia de cada variable.

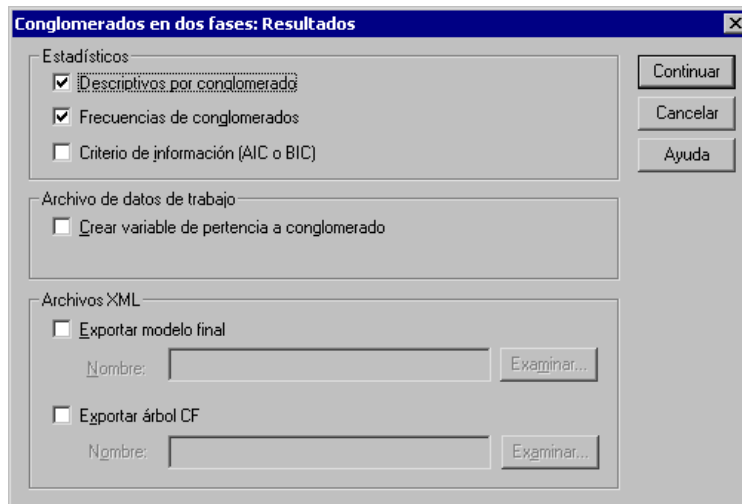
- **Ordenar variables.** Esta opción determina si los gráficos que se crearán para cada conglomerado (Por variable) o para cada variable (Por conglomerado).

- **Medida de la importancia.** Esta opción le permite seleccionar la medida de la importancia de las variables que se va a representar en el gráfico. Chi-cuadrado o prueba t de significación muestra un estadístico chi-cuadrado de Pearson como la importancia de una variable categórica y un estadístico t como importancia de una variable continua. Significación muestra uno menos el valor p para la prueba de igualdad de las medias para una variable continua y la frecuencia esperada del conjunto de datos global para una variable categórica.
- **Nivel de confianza.** Esta opción permite establecer el nivel de confianza para la prueba de igualdad de la distribución de una variable dentro de un conglomerado frente a la distribución global de la variable. Especifique un número inferior a 100 y superior o igual a 50. El valor del nivel de confianza se muestra como una línea vertical en los gráficos de la importancia de las variables, si los gráficos se crean por variable o si se representa gráficamente la medida de la significación.
- **Omitir variables no significativas.** Las variables que no son significativas para el nivel de confianza especificado no aparecen en los gráficos de la importancia de las variables.

Resultados de análisis de conglomerados en dos fases

Figura 31-4

Cuadro de diálogo Resultados de análisis de conglomerados en dos fases



Estadísticos. Este grupo proporciona opciones para la presentación de tablas con los resultados de la conglomeración. Se generan los estadísticos descriptivos y las frecuencias de los conglomerados para el modelo de conglomerados final, mientras que la tabla de criterio de información muestra los resultados correspondientes a varias soluciones de conglomerados.

- **Descriptivos por conglomerado.** Muestra dos tablas que describen las variables de cada conglomerado. En una tabla, se informa de las medias y las desviaciones típicas para las variables continuas por conglomerado. La otra tabla informa de las frecuencias de las variables categóricas por conglomerado.
- **Frecuencias de los conglomerados.** Muestra una tabla que informa del número de observaciones existentes en cada conglomerado.
- **Criterio de información (AIC o BIC).** Muestra una tabla que contiene los valores de AIC o BIC, dependiendo del criterio elegido en el cuadro de diálogo principal, para diferentes números de conglomerados. Sólo se proporciona esta tabla cuando el número de conglomerados se determina automáticamente. Si el número de conglomerados es fijo, se ignora este ajuste y no se proporciona la tabla.

Archivo de datos de trabajo. Este grupo permite guardar las variables en el archivo de datos de trabajo.

- **Crear variable del conglomerado de pertenencia.** Esta variable contiene un número de identificación de conglomerado para cada caso. El nombre de esta variable es *tsc_n*, donde *n* es un número entero positivo que indica el ordinal de la operación de almacenamiento del archivo de datos de trabajo realizada por este procedimiento en una determinada sesión.

Archivos XML. El modelo de conglomerados final y el árbol CF son dos tipos de archivos de resultados que se pueden exportar en formato XML.

- **Exportar modelo final.** También se puede exportar el modelo de conglomerado final al archivo especificado en formato XML (PMML). *SmartScore* y la versión de servidor de *SPSS* (un producto independiente) pueden utilizar este archivo de modelo para aplicar la información de modelo a otros archivos de datos con fines de puntuación.
- **Exportar árbol CF.** Esta opción permite guardar el estado actual del árbol de conglomerados y actualizarlo más tarde utilizando nuevos datos.

Análisis de conglomerados jerárquico

Este procedimiento intenta identificar grupos relativamente homogéneos de casos (o de variables) basándose en las características seleccionadas, mediante un algoritmo que comienza con cada caso (o cada variable) en un conglomerado diferente y combina los conglomerados hasta que sólo queda uno. Es posible analizar las variables brutas o elegir de entre una variedad de transformaciones de estandarización. Las medidas de distancia o similitud se generan mediante el procedimiento Proximidades. Los estadísticos se muestran en cada etapa para ayudar a seleccionar la mejor solución.

Ejemplo. ¿Existen grupos identificables de programas televisivos que atraigan a audiencias similares dentro de cada grupo? Con el análisis de conglomerados jerárquico, podría agrupar los programas de TV (los casos) en grupos homogéneos basados en las características del espectador. Esto se puede utilizar para identificar segmentos de mercado. También puede agrupar ciudades (los casos) en grupos homogéneos, de manera que se puedan seleccionar ciudades comparables para probar diversas estrategias de marketing.

Estadísticos. Historial de conglomeración, matriz de distancias (o similitudes) y pertenencia a los conglomerados para una solución única o una serie de soluciones. Gráficos: dendrogramas y diagramas de témpanos.

Datos. Las variables pueden ser cuantitativas, binarias o datos de recuento (frecuencias). El escalamiento de las variables es un aspecto importante, ya que las diferencias en el escalamiento pueden afectar a las soluciones de conglomeración. Si las variables muestran grandes diferencias en el escalamiento (por ejemplo, una variable se mide en dólares y la otra se mide en años), debería considerar la posibilidad de estandarizarlas (esto puede llevarse a cabo automáticamente mediante el propio procedimiento Análisis de conglomerados jerárquico).

Orden de casos. Si hay distancias empatadas o similitudes en los datos de entrada o bien se producen entre los conglomerados actualizados durante la unión, la solución de conglomerado resultante puede depender del orden de los casos del archivo. Puede que desee obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada.

Supuestos. Las medidas de distancia o similitud empleadas deben ser adecuadas para los datos analizados (véase el procedimiento Proximidades para obtener más información sobre la elección de las medidas de distancia y similitud). Asimismo, debe incluir todas las variables relevantes en el análisis. Si se omiten variables de interés la solución obtenida puede ser equívoca. Debido a que el análisis de conglomerados jerárquico es un método exploratorio, los resultados deben considerarse provisionales hasta que sean confirmados mediante otra muestra independiente.

Figura 32-1

Resultados del procedimiento Análisis de conglomerados jerárquico

Agglomeration Schedule							
		Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
		Cluster 1	Cluster 2		Cluster 1	Cluster 2	
Stage	1	11	12	.112	0	0	2
	2	6	11	.132	0	1	4
	3	7	9	.185	0	0	5
	4	6	8	.227	2	0	7
	5	7	10	.274	3	0	7
	6	1	3	.423	0	0	10
	7	6	7	.438	4	5	14
	8	13	14	.484	0	0	15
	9	2	5	.547	0	0	11
	10	1	4	.691	6	0	11
	11	1	2	1.023	10	9	13
	12	15	16	1.370	0	0	13
	13	1	15	1.716	11	12	14
	14	1	6	2.642	13	7	15
	15	1	13	4.772	14	8	0

- ▶ Si está aglomerando casos, seleccione al menos una variable numérica. Si está aglomerando variables, seleccione al menos tres variables numéricas.

Si lo desea, puede seleccionar una variable de identificación para etiquetar los casos.

Análisis de conglomerados jerárquico: Método

Figura 32-3

Cuadro de diálogo Análisis de conglomerados jerárquico: Método

Método de conglomeración. Las opciones disponibles son: Vinculación inter-grupos, Vinculación intra-grupos, Vecino más próximo, Vecino más lejano, Agrupación de centroides, Agrupación de medianas y Método de Ward.

Medida. Permite especificar la medida de distancia o similitud que será empleada en la aglomeración. Seleccione el tipo de datos y la medida de distancia o similitud adecuada:

- **Datos de intervalo.** Distancia euclídea, Distancia euclídea al cuadrado, Coseno, Correlación de Pearson, Chebychev, Bloque, Minkowski y Personalizada.
- **Datos de frecuencias.** Las opciones disponibles son: Medida de chi-cuadrado y Medida de phi-cuadrado.
- **Datos binarios.** Las opciones disponibles son: Distancia euclídea, Distancia euclídea al cuadrado, Diferencia de tamaño, Diferencia de configuración, Varianza, Dispersión, Forma, Concordancia simple, Correlación phi de 4 puntos,

Lambda, D de Anderberg, Dice, Hamann, Jaccard, Kulczynski 1, Kulczynski 2, Lance y Williams, Ochiai, Rogers y Tanimoto, Russel y Rao, Sokal y Sneath 1, Sokal y Sneath 2, Sokal y Sneath 3, Sokal y Sneath 4, Sokal y Sneath 5, Y de Yule y Q de Yule.

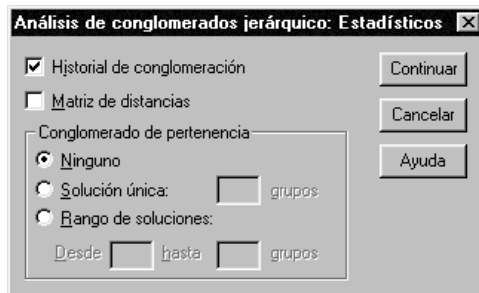
Transformar valores. Permite estandarizar los valores de los datos, para los casos o las variables, antes de calcular las proximidades (no está disponible para datos binarios). Los métodos disponibles de estandarización son: Puntuaciones z , Rango -1 a 1, Rango 0 a 1, Magnitud máxima de 1, Media de 1 y Desviación típica 1.

Transformar medidas. Permite transformar los valores generados por la medida de distancia. Se aplican después de calcular la medida de distancia. Las opciones disponibles son: Valores absolutos, Cambiar el signo y Cambiar la escala al rango 0–1.

Análisis de conglomerados jerárquico: Estadísticos

Figura 32-4

Cuadro de diálogo Análisis de conglomerados jerárquico: Estadísticos



Historial de conglomeración. Muestra los casos o conglomerados combinados en cada etapa, las distancias entre los casos o los conglomerados que se combinan, así como el último nivel del proceso de aglomeración en el que cada caso (o variable) se unió a su conglomerado correspondiente.

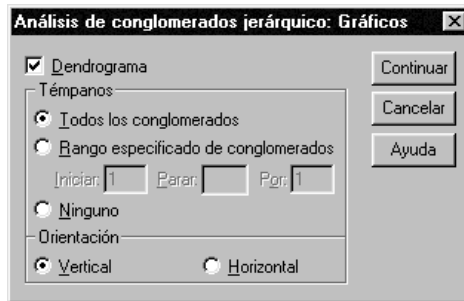
Matriz de distancias. Proporciona las distancias o similitudes entre los elementos.

Conglomerado de pertenencia. Muestra el conglomerado al cual se asigna cada caso en una o varias etapas de la combinación de los conglomerados. Las opciones disponibles son: Solución única y Rango de soluciones.

Análisis de conglomerados jerárquico: Gráficos

Figura 32-5

Cuadro de diálogo Análisis de conglomerados jerárquico: Gráficos



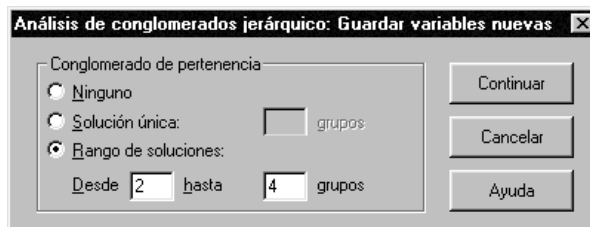
Dendrograma. Muestra un **dendrograma**. Los dendrogramas pueden emplearse para evaluar la cohesión de los conglomerados que se han formado y proporcionar información sobre el número adecuado de conglomerados que deben conservarse.

Témpanos. Muestra un **diagrama de témpanos**, que incluye todos los conglomerados o un rango especificado de conglomerados. Los diagramas de témpanos muestran información sobre cómo se combinan los casos en los conglomerados, en cada iteración del análisis. La orientación permite seleccionar un diagrama vertical u horizontal.

Análisis de conglomerados jerárquico: Guardar variables nuevas

Figura 32-6

Cuadro de diálogo Análisis de conglomerados jerárquico: Guardar variables nuevas



Conglomerado de pertenencia. Permite guardar los conglomerados de pertenencia para una solución única o un rango de soluciones. Las variables guardadas pueden emplearse en análisis posteriores para explorar otras diferencias entre los grupos.

Análisis de conglomerados de K-medias

Este procedimiento intenta identificar grupos de casos relativamente homogéneos basándose en las características seleccionadas y utilizando un algoritmo que puede gestionar un gran número de casos. Sin embargo, el algoritmo requiere que el usuario especifique el número de conglomerados. Puede especificar los centros iniciales de los conglomerados si conoce de antemano dicha información. Puede elegir uno de los dos métodos disponibles para clasificar los casos: la actualización de los centros de los conglomerados de forma iterativa o sólo la clasificación. Asimismo, puede guardar la pertenencia a los conglomerados, información de la distancia y los centros de los conglomerados finales. Si lo desea, puede especificar una variable cuyos valores sean utilizados para etiquetar los resultados por casos. También puede solicitar los estadísticos F de los análisis de varianza. Aunque estos estadísticos son oportunistas (ya que el procedimiento trata de formar grupos que de hecho difieran), el tamaño relativo de los estadísticos proporciona información acerca de la contribución de cada variable a la separación de los grupos.

Ejemplo. ¿Cuáles son los grupos identificables de programas de televisión que atraen audiencias parecidas dentro de cada grupo? Con el análisis de conglomerados de k -medias, podría agrupar los programas de televisión (los casos) en k grupos homogéneos, basados en las características del televidente. Esto se puede utilizar para identificar segmentos de mercado. También puede agrupar ciudades (los casos) en grupos homogéneos, de manera que se puedan seleccionar ciudades comparables para probar diversas estrategias de marketing.

Estadísticos. Para la solución completa: centros iniciales de los conglomerados, tabla de ANOVA. Para cada caso: información del conglomerado, distancia desde el centro del conglomerado.

Datos. Las variables deben ser cuantitativas en el nivel de intervalo o de razón. Si las variables son binarias o recuentos, utilice el procedimiento Análisis de conglomerados jerárquicos.

Orden de casos y centro de conglomerados iniciales. El algoritmo por defecto para elegir centros de conglomerados iniciales no es invariable con respecto a la ordenación de casos. La opción Usar medias actualizadas del cuadro de diálogo Iterar hace que la solución resultante dependa potencialmente del orden de casos con independencia de cómo se eligen los centros de conglomerados iniciales. Si va a utilizar alguno de estos métodos, puede que desee obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada. La especificación de los centros de conglomerados iniciales y la no utilización de la opción Usar medias actualizadas evita los problemas relacionados con el orden de casos. No obstante, la ordenación de los centros de conglomerados iniciales puede afectar a la solución en caso de haber distancias empatadas desde los casos a los centros de conglomerados. La comparación de los resultados de los análisis con las distintas permutaciones de los valores de centros iniciales se puede utilizar para evaluar la estabilidad de una solución determinada.

Supuestos. Las distancias se calculan utilizando la distancia euclídea simple. Si desea utilizar otra medida de distancia o de similaridad, utilice el procedimiento Análisis de conglomerados jerárquicos. El escalamiento de las variables es una consideración importante; si sus variables utilizan diferentes escalas (por ejemplo, una variable se expresa en dólares y otra en años), los resultados podrían ser equívocos. En estos casos, debería considerar la estandarización de las variables antes de realizar el análisis de conglomerados de k -medias (esto se puede hacer en el procedimiento Descriptivos). Este procedimiento supone que ha seleccionado el número apropiado de conglomerados y que ha incluido todas las variables relevantes. Si ha seleccionado un número inapropiado de conglomerados o ha omitido variables relevantes, los resultados podrían ser equívocos.

Figura 33-1
Resultados del análisis de conglomerados de K-medias

Centros iniciales de los conglomerados

	Conglomerado			
	1	2	3	4
ZURBAN	-1,88606	-1,54314	1,45741	,55724
ZLIFEEXP	-3,52581	-1,69358	,62725	,99370
ZLITERAC	-2,89320	-1,65146	-,51770	,88601
ZPOP_INC	,93737	,16291	3,03701	-1,12785
ZBABYMOR	4,16813	1,38422	-,69589	-,88983
ZBIRTH_R	2,68796	,42699	,33278	-1,08033
ZDEATH_R	4,41517	,63185	-1,89037	,63185
ZLOG_GDP	-1,99641	-1,78455	,53091	1,22118
ZB_TO_D	-,52182	- 31333	4 40082	- 99285
ZFERTILT	2,24070	,75481	,46008	-,76793
ZLOG_POP	,24626	2,65246	-1,29624	-,74406

Historial de iteraciones

Iteración	Cambio en los centros de los conglomerados			
	1	2	3	4
1	1.932	1.596	3.343	2.724
2	.000	.314	.466	.471
3	.861	.195	.172	.414
4	.604	.150	.000	.337
5	.000	.167	.237	.253
6	.000	.071	.287	.199
7	.623	.000	.000	.160
8	.000	.074	.000	.084
9	.000	.077	.000	.080
10	.000	.000	.185	.097

Centros de los conglomerados finales

	Conglomerado			
	1	2	3	4
ZURBANA	-1.70745	.62767	.16816	-.30863
ZESPVIDA	-2.52826	.80611	-.28417	-.15939
ZALFABET	-2.30833	.73368	-.81671	.13880
ZINC_POB	.59747	-.95175	1.45301	.13400
ZMORTINF	2.43210	-.80817	.25622	.22286
ZTASA_NA	1.52607	-.99285	1.13716	.12929
ZTASA_M	2.10314	.31319	-.71414	-.44640
ZLOG_PIB	-1.77704	.94249	-.16871	-.58745
ZNAC_DEF	-.29856	-.84758	1.45251	.19154
ZFERTILI	1.51003	-.87669	1.27010	-.12150
ZLOG_POB	.83475	-.22199	-.49499	.34577

Distancias entre los centros de los conglomerados finales

Conglomerado	1	2	3	4
1		7.924	5.640	5.627
2	7.924		5.246	3.249
3	5.640	5.246		2.897
4	5.627	3.249	2.897	

ANOVA

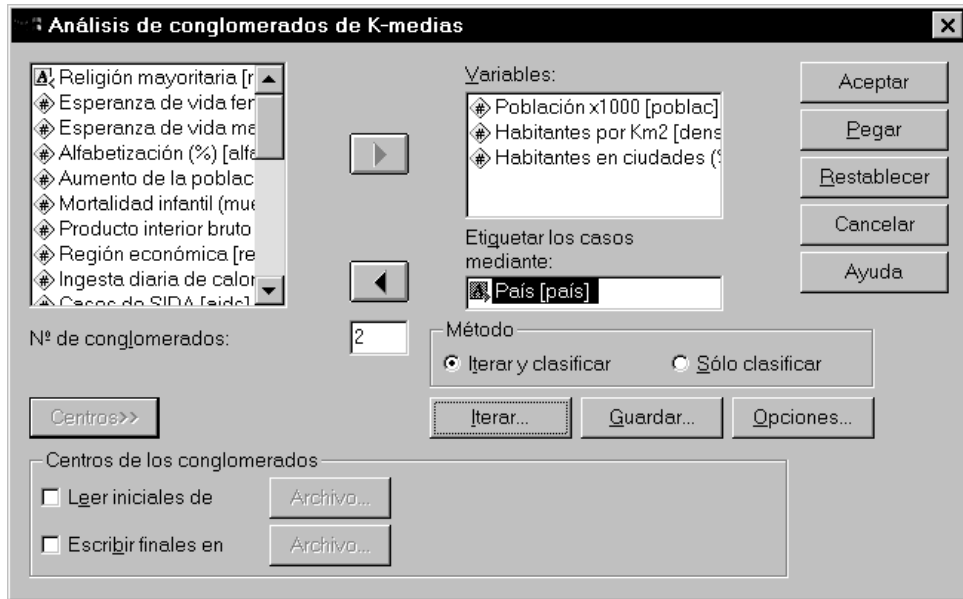
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
ZURBAN	10.409	3	.541	68	19.234	.000
ZLIFEEXP	19.410	3	.210	68	92.614	.000
ZLITERAC	18.731	3	.229	68	81.655	.000
ZPOP_INC	18.464	3	.219	68	84.428	.000
ZBABYMOR	18.621	3	.239	68	77.859	.000
ZBIRTH_R	19.599	3	.167	68	117.339	.000
ZDEATH_R	13.628	3	.444	68	30.676	.000
ZLOG_GDP	17.599	3	.287	68	61.313	.000
ZB_TO_D	16.316	3	.288	68	56.682	.000
ZFERTILT	18.829	3	.168	68	112.273	.000
ZLOG_POP	3.907	3	.877	68	4.457	.006

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Para obtener un análisis de conglomerados de K-medias

- ▶ Elija en los menús:
 - Analizar
 - Clasificar
 - Conglomerado de K medias...

Figura 33-2
Cuadro de diálogo Análisis de conglomerados de K-medias



- ▶ Seleccione las variables que se van a utilizar en el análisis de conglomerados.
- ▶ Especifique el número de conglomerados. Este número no debe ser inferior a dos ni superior al número de casos del archivo de datos.
- ▶ Seleccione Iterar y clasificar o Sólo clasificar.

Si lo desea, puede seleccionar una variable de identificación para etiquetar los casos.

Eficacia del análisis de conglomerados de K-medias

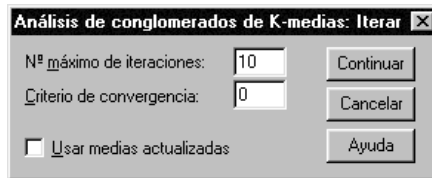
El comando de análisis de conglomerados de *k*-medias es eficaz principalmente porque no calcula las distancias entre todos los pares de casos, como hacen muchos algoritmos de conglomeración, como el utilizado por el comando de conglomeración jerárquica.

Para obtener la máxima eficacia, tome una muestra de casos y utilice el método Iterar y clasificar para determinar los centros de los conglomerados. Seleccione Escribir finales en Archivo. Después restaure el archivo de datos completo y seleccione el método Sólo clasificar. Pulse en Centros y pulse en Leer iniciales de Archivo para clasificar el archivo completo utilizando los centros estimados a partir de la muestra.

Análisis de conglomerados de *K-medias*: Iterar

Figura 33-3

Cuadro de diálogo Análisis de conglomerados de *K-medias*: Iterar



Estas opciones sólo están disponibles si se selecciona el método Iterar y clasificar en el cuadro de diálogo principal.

Nº máximo de iteraciones. Limita el número de iteraciones en el algoritmo *k-medias*. La iteración se detiene después de este número de iteraciones, incluso si no se ha satisfecho el criterio de convergencia. Este número debe estar entre el 1 y el 999.

Para reproducir el algoritmo utilizado por el comando Quick Cluster en las versiones previas a la 5.0, establezca en 1 el número máximo de iteraciones.

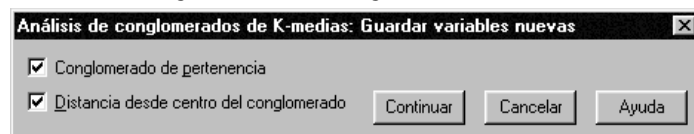
Criterio de convergencia. Determina cuándo cesa la iteración. Representa una proporción de la distancia mínima entre los centros iniciales de los conglomerados, por lo que debe ser mayor que 0 pero no mayor que 1. Por ejemplo, si el criterio es igual a 0,02, la iteración cesará si una iteración completa no mueve ninguno de los centros de los conglomerados en una distancia superior al dos por ciento de la distancia menor entre cualquiera de los centros iniciales.

Usar medias actualizadas. Permite solicitar la actualización de los centros de los conglomerados tras la asignación de cada caso. Si no selecciona esta opción, los nuevos centros de los conglomerados se calcularán después de la asignación de todos los casos.

Análisis de conglomerados de K-medias: Guardar

Figura 33-4

Cuadro de diálogo *Análisis de conglomerados de K-medias: Guardar*



Puede guardar información sobre la solución como nuevas variables para que puedan ser utilizadas en análisis subsiguientes:

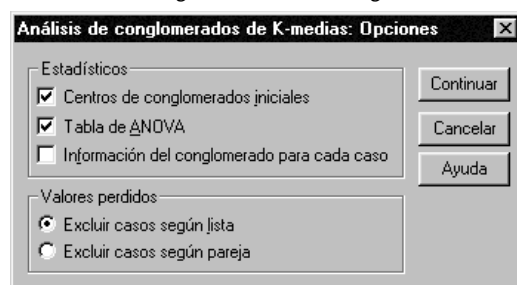
Conglomerado de pertenencia. Crea una nueva variable que indica el conglomerado final al que pertenece cada caso. Los valores de la nueva variable van desde el 1 hasta el número de conglomerados.

Distancia desde centro del conglomerado. Crea una nueva variable que indica la distancia euclídea entre cada caso y su centro de clasificación.

Análisis de conglomerados de K-medias: Opciones

Figura 33-5

Cuadro de diálogo *Análisis de conglomerados de K-medias: Opciones*



Estadísticos. Puede seleccionar los siguientes estadísticos: Centros de conglomerados iniciales, Tabla de ANOVA e Información del conglomerado para cada caso.

- **Centros de conglomerados iniciales.** Primera estimación de las medias de las variables para cada uno de los conglomerados. Por defecto se selecciona entre los datos un número de casos debidamente espaciados igual al número de

conglomerados. Los centros iniciales de los conglomerados se utilizan como criterio para una primera clasificación y, a partir de ahí, se van actualizando.

- **Tabla de ANOVA (Conglomerado de K-medias, Opciones).** Muestra una tabla de análisis de varianza que incluye las pruebas F univariadas para cada variable de aglomeración. Las pruebas F son sólo descriptivas y las probabilidades resultantes no deben ser interpretadas. La tabla de ANOVA no se mostrará si se asignan todos los casos a un único conglomerado.
- **Información del conglomerado para cada caso.** Muestra, para cada caso, el conglomerado final asignado y la distancia euclídea entre el caso y el centro del conglomerado utilizado para clasificar el caso. También muestra la distancia euclídea entre los centros de los conglomerados finales.

Valores perdidos. Las opciones disponibles son: Excluir casos según lista o Excluir casos según pareja.

- **Excluir casos según lista.** Excluye los casos con valores perdidos para cualquier variable de agrupación del análisis.
- **Excluir casos según pareja.** Asigna casos a los conglomerados en función de las distancias calculadas desde todas las variables con valores no perdidos.

Pruebas no paramétricas

El procedimiento Pruebas no paramétricas proporciona varias pruebas que no necesitan supuestos sobre la forma de la distribución subyacente:

Prueba de chi-cuadrado. Tabula una variable en categorías y calcula un estadístico de chi-cuadrado basándose en las diferencias entre las frecuencias observadas y las esperadas.

Prueba binomial. Compara la frecuencia observada en cada categoría de una variable dicotómica con las frecuencias esperadas en la distribución binomial.

Prueba de rachas. Comprueba si el orden de aparición de dos valores de una variable es aleatorio.

Prueba de Kolmogorov-Smirnov para una muestra. Compara la función de distribución acumulativa observada de una variable con una distribución teórica especificada, que puede ser normal, uniforme o de Poisson.

Pruebas para dos muestras independientes. Compara dos grupos de casos en una variable. Se encuentran disponibles la prueba U de Mann-Whitney, la prueba de Kolmogorov-Smirnov para dos muestras, la prueba de Moses de reacciones extremas y la prueba de rachas de Wald-Wolfowitz.

Pruebas para varias muestras independientes. Compara dos o más grupos de casos en una variable. Se encuentran disponibles la prueba de Kruskal-Wallis, la prueba de la mediana y la prueba de Jonckheere-Terpstra.

Pruebas para dos muestras relacionadas. Compara las distribuciones de dos variables. La prueba de Wilcoxon de los rangos con signo, la prueba de signos y la prueba de McNemar.

Pruebas para varias muestras relacionadas. Compara las distribuciones de dos o más variables. Se encuentran disponibles la prueba de Friedman, la W de Kendall y la Q de Cochran.

Además, para todas las pruebas citadas anteriormente están disponibles los cuartiles y la media, la desviación típica, el mínimo, el máximo y, por último, el número de casos no perdidos.

Prueba de chi-cuadrado

El procedimiento Prueba de chi-cuadrado tabula una variable en categorías y calcula un estadístico de chi-cuadrado. Esta prueba de bondad de ajuste compara las frecuencias observadas y esperadas en cada categoría para contrastar si todas las categorías contienen la misma proporción de valores o si cada categoría contiene una proporción de valores especificada por el usuario.

Ejemplos. La prueba de chi-cuadrado podría utilizarse para determinar si una bolsa de caramelos contiene en igualdad de proporción caramelos de color azul, marrón, verde, naranja, rojo y amarillo. También podría utilizarse para ver si una bolsa de caramelos contiene un 5% de color azul, un 30% de color marrón, un 10% de color verde, un 20% de color naranja, un 15% de color rojo y un 15% de color amarillo.

Estadísticos. Media, desviación típica, mínimo, máximo y cuartiles. Número y el porcentaje de casos perdidos y no perdidos, número de casos observados y esperados de cada categoría, residuos y estadístico de chi-cuadrado.

Datos. Use variables categóricas numéricas ordenadas o no ordenadas (niveles de medida ordinal o nominal). Para convertir las variables de cadena en variables numéricas, utilice el procedimiento Recodificación automática, disponible en el menú Transformar.

Supuestos. Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Se asume que los datos son una muestra aleatoria. Las frecuencias esperadas para cada categoría deberán ser 1 como mínimo. No más de un 20% de las categorías deberán tener frecuencias esperadas menores que 5.

Figura 34-1
Resultados del procedimiento Prueba de chi-cuadrado

Color de la jalea

	N observado	N esperado	Residual
Azul	6	5.9	.1
Marrón	33	35.7	-2.7
Verde	9	11.9	-2.9
Amarillo	17	23.8	-6.8
Naranja	22	17.8	4.2
Rojo	26	17.8	8.2
Total	113		

Estadísticos de contraste

	Color
Chi-cuadrado ¹	7.544
gl	5
Sig. asintót.	.183

1. 0 casillas (.0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es 5.9.

Color de la jalea

	N observado	N esperado	Residual
Azul	6	18.8	-12.8
Marrón	33	18.8	14.2
Verde	9	18.8	-9.8
Amarillo	17	18.8	-1.8
Naranja	22	18.8	3.2
Rojo	26	18.8	7.2
Total	113		

Estadísticos de contraste

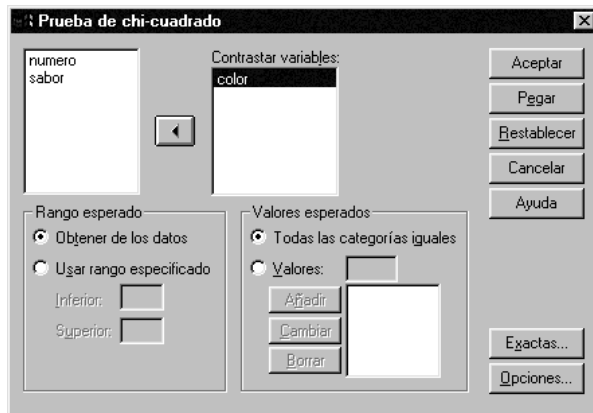
	Color
Chi-cuadrado ¹	27.973
gl	5
Sig. asintót.	.000

1. 0 casillas (.0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es 18.8.

Para obtener una prueba de chi-cuadrado

- ▶ Elija en los menús:
 - Analizar
 - Pruebas no paramétricas
 - Chi-cuadrado...

Figura 34-2
Cuadro de diálogo Prueba de chi-cuadrado



- ▶ Seleccione una o más variables de contraste. Cada variable genera una prueba independiente.

Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y el control sobre el tratamiento de los datos perdidos.

Prueba de chi-cuadrado: Rango y valores esperados

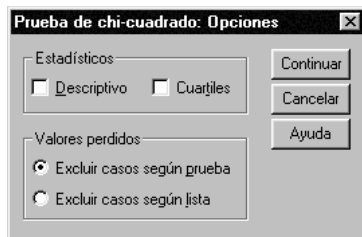
Rango esperado. Por defecto, cada valor distinto de la variable se define como una categoría. Para establecer categorías dentro de un rango específico, seleccione Usar rango especificado e introduzca valores enteros para los límites inferior y superior. Se establecerán categorías para cada valor entero dentro del rango inclusivo y los casos con valores fuera de los límites se excluirán. Por ejemplo, si se especifica 1 como límite inferior y 4 como límite superior, únicamente se utilizarán los valores enteros entre 1 y 4 para la prueba de chi-cuadrado.

Valores esperados. Por defecto, todas las categorías tienen valores esperados iguales. Las categorías pueden tener proporciones esperadas especificadas por el usuario. Seleccione Valores, introduzca un valor mayor que 0 para cada categoría de la variable de contraste y pulse en Añadir. Cada vez que se añade un valor, éste aparece al final de la lista de valores. El orden de los valores es importante; corresponde al orden ascendente de los valores de categoría de la variable de contraste. El primer valor de la lista corresponde al valor de grupo mínimo de la variable de contraste y el último valor corresponde al valor máximo. Los elementos de la lista de valores se suman y, a continuación, cada valor se divide por esta suma para calcular la proporción de casos esperados en la categoría correspondiente. Por ejemplo, una lista de valores de 3, 4, 5, 4 especifica unas proporciones esperadas de 3/16, 4/16, 5/16 y 4/16.

Prueba de chi-cuadrado: Opciones

Figura 34-3

Cuadro de diálogo Prueba de chi-cuadrado: Opciones



Estadísticos. Puede seleccionar uno o ambos de los estadísticos de resumen siguientes:

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

Funciones adicionales del comando NPAR TESTS (prueba de chi-cuadrado)

Con el lenguaje de comandos también podrá:

- Especificar diferentes valores mínimo y máximo o las frecuencias esperadas para diferentes variables (mediante el subcomando `CHISQUARE`).
- Contrastar la misma variable respecto a las diferentes frecuencias esperadas, o bien usar diferentes rangos (mediante el subcomando `EXPECTED`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Prueba binomial

El procedimiento Prueba binomial compara las frecuencias observadas de las dos categorías de una variable dicotómica con las frecuencias esperadas en una distribución binomial con un parámetro de probabilidad especificado. Por defecto, el parámetro de probabilidad para ambos grupos es 0,5. Para cambiar las probabilidades, puede introducirse una proporción de prueba para el primer grupo. La probabilidad del segundo grupo será 1 menos la probabilidad especificada para el primer grupo.

Ejemplo. Si se lanza una moneda al aire, la probabilidad de que salga cara es 1/2. Basándose en esta hipótesis, se lanza una moneda al aire 40 veces y se anotan los resultados (cara o cruz). De la prueba binomial, podría deducir que en 3/4 de los lanzamientos salió cara y que el nivel crítico es pequeño (0,0027). Estos resultados indican que no es verosímil que la probabilidad de que salga cara sea 1/2; probablemente la moneda presenta una tendencia a caer por un sentido determinado.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles.

Datos. Las variables de contraste deben ser numéricas y dicotómicas. Para convertir las variables de cadena en variables numéricas, utilice el procedimiento Recodificación automática, disponible en el menú Transformar. Una **variable dicotómica** es una variable que sólo puede tomar dos valores posibles: *sí* o *no*, *verdadero* o *falso*, 0 ó 1, etc. Si las variables no son dicotómicas, debe especificar un punto de corte. El punto de corte asigna a un grupo los casos con valores mayores que el punto de corte y el resto de los casos a otro grupo.

Supuestos. Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Se asume que los datos son una muestra aleatoria.

Resultados de muestra

Figura 34-4
Resultados del procedimiento Prueba binomial

		Categoría	N	Proporción observada	Prop. de prueba	Sig. asintót. (bilateral)
Moneda	Grupo 1	Cara	30	.75	.50	.003 ¹
	Grupo 2	Cruz	10	.25		
	Total		40	1.00		

1. Basado en la aproximación Z.

Para obtener una prueba binomial

- ▶ Elija en los menús:
 - Analizar
 - Pruebas no paramétricas
 - Binomial...

Figura 34-5
Cuadro de diálogo Prueba binomial

Prueba binomial

Contrastar variables:

Clasificación étnica [mino]

Contrastar proporción: .50

Definir la dicotomía:

Obtener de los datos

Punto de corte:

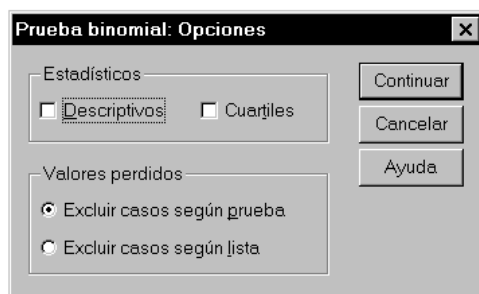
Buttons: Aceptar, Pegar, Restablecer, Cancelar, Ayuda, Exactas..., Opciones...

- Seleccione una o más variables de contraste numéricas.

Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y el control sobre el tratamiento de los datos perdidos.

Prueba binomial: Opciones

Figura 34-6
Cuadro de diálogo Prueba binomial: Opciones



Estadísticos. Puede seleccionar uno o ambos de los estadísticos de resumen siguientes:

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Se excluirán de todos los análisis los casos con valores perdidos de cualquier variable.

Funciones adicionales del comando NPAR TESTS (Prueba binomial)

El lenguaje de comandos de SPSS también permite:

- Seleccionar grupos específicos (y excluir otros) si una variable tiene más de dos categorías (mediante el subcomando BINOMIAL).

- Especificar diferentes probabilidades o puntos de corte para diferentes variables (mediante el subcomando `BINOMIAL`).
- Contrastar la misma variable respecto a diferentes probabilidades o puntos de corte (mediante el subcomando `EXPECTED`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Prueba de rachas

El procedimiento Prueba de rachas contrasta si es aleatorio el orden de aparición de dos valores de una variable. Una racha es una secuencia de observaciones similares. Una muestra con un número excesivamente grande o excesivamente pequeño de rachas sugiere que la muestra no es aleatoria.

Ejemplos. Suponga que se realiza una encuesta a 20 personas para saber si comprarían un producto. Si todas estas personas fueran del mismo sexo, se pondría seriamente en duda la supuesta aleatoriedad de la muestra. La prueba de rachas se puede utilizar para determinar si la muestra fue extraída de manera aleatoria.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles.

Datos. Las variables deben ser numéricas. Para convertir las variables de cadena en variables numéricas, utilice el procedimiento Recodificación automática, disponible en el menú Transformar.

Supuestos. Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Utilice muestras de distribuciones de probabilidad continua.

Figura 34-7
Resultados de la prueba de rachas

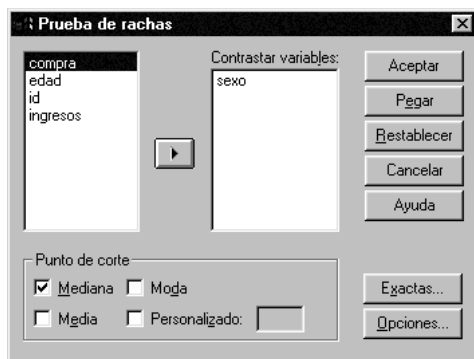
Runs Test	
	Gender
Test Value ¹	1.00
Cases < Test Value	7
Cases >= Test Value	13
Total Cases	20
Number of Runs	15
Z	2.234
Asymptotic Significance (2-tailed)	.025

¹. Median

Para obtener una prueba de rachas

- ▶ Elija en los menús:
Analizar
Pruebas no paramétricas
Rachas...

Figura 34-8
Cuadro de diálogo Prueba de rachas



- ▶ Seleccione una o más variables de contraste numéricas.

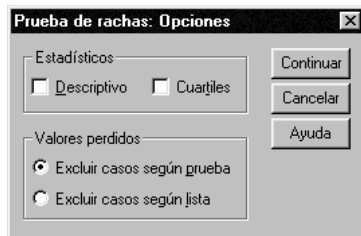
Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y el control sobre el tratamiento de los datos perdidos.

Prueba de rachas: punto de corte

Punto de corte. Especifica un punto de corte para dicotomizar las variables seleccionadas. Puede utilizar como punto de corte los valores observados para la media, la mediana o la moda, o bien un valor especificado. Los casos con valores menores que el punto de corte se asignarán a un grupo y los casos con valores mayores o iguales que el punto de corte se asignarán a otro grupo. Se lleva a cabo una prueba para cada punto de corte seleccionado.

Prueba de rachas: Opciones

Figura 34-9
Cuadro de diálogo Prueba de rachas: Opciones



Estadísticos. Puede seleccionar uno o ambos de los estadísticos de resumen siguientes:

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

Funciones adicionales del comando NPAR TESTS (Prueba de rachas)

El lenguaje de comandos de SPSS también permite:

- Especificar puntos de corte diferentes para las distintas variables (con el subcomando `RUNS`).
- Contrastar la misma variable con distintos puntos de corte personalizados (con el subcomando `RUNS`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Prueba de Kolmogorov-Smirnov para una muestra

El procedimiento Prueba de Kolmogorov-Smirnov para una muestra compara la función de distribución acumulada observada de una variable con una distribución teórica determinada, que puede ser la normal, la uniforme, la de Poisson o la exponencial. La Z de Kolmogorov-Smirnov se calcula a partir de la diferencia mayor (en valor absoluto) entre las funciones de distribución acumuladas teórica y observada. Esta prueba de bondad de ajuste contrasta si las observaciones podrían razonablemente proceder de la distribución especificada.

Ejemplo. Muchas pruebas paramétricas requieren que las variables se distribuyan de forma normal. La prueba de Kolmogorov-Smirnov para una muestra se puede utilizar para comprobar que una variable, por ejemplo *ingresos*, se distribuye normalmente.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles.

Datos. Utilice variables cuantitativas (a nivel de medida de razón o de intervalo).

Supuestos. La prueba de Kolmogorov-Smirnov asume que los parámetros de la distribución de contraste se han especificado previamente. Este procedimiento estima los parámetros a partir de la muestra. La media y la desviación típica de la muestra son los parámetros de una distribución normal, los valores mínimo y máximo de la muestra definen el rango de la distribución uniforme, la media muestral es el parámetro de la distribución de Poisson y la media muestral es el parámetro de la distribución exponencial.

Figura 34-10

Resultados del procedimiento Prueba de Kolmogorov-Smirnov para una muestra

Prueba de Kolmogorov-Smirnov para una muestra

		Ingresos
N		20
Parámetros normales ^{1,2}	Media	56250.00
	Desviación típica	45146.40
Diferencias más extremas	Absoluta	.170
	Positiva	.170
	Negativa	-.164
Z de Kolmogorov-Smirnov		.760
Sig. asintót. (bilateral)		.611

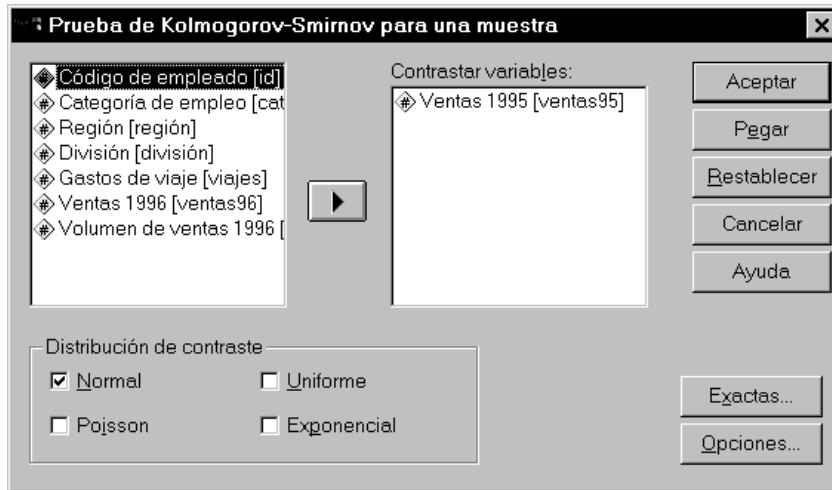
1. La distribución de contraste es la Normal.

2. Se han calculado a partir de los datos.

Para obtener una prueba de Kolmogorov-Smirnov para una muestra

- ▶ Elija en los menús:
 - Analizar
 - Pruebas no paramétricas
 - K-S de 1 muestra...

Figura 34-11
Cuadro de diálogo Prueba de Kolmogorov-Smirnov para una muestra

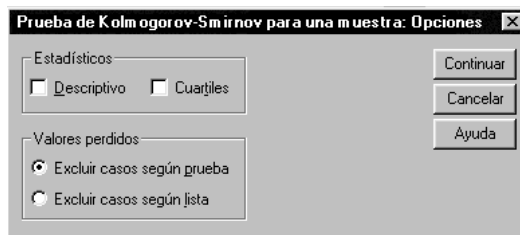


- Seleccione una o más variables de contraste numéricas. Cada variable genera una prueba independiente.

Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y el control sobre el tratamiento de los datos perdidos.

Prueba de Kolmogorov-Smirnov para una muestra: Opciones

Figura 34-12
Cuadro de diálogo Prueba de Kolmogorov-Smirnov para una muestra: Opciones



Estadísticos. Puede seleccionar uno o ambos de los estadísticos de resumen siguientes:

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

Funciones adicionales del comando NPAR TESTS (Prueba de Kolmogorov-Smirnov para una muestra)

El lenguaje de comandos de SPSS también permite:

- Especificar los parámetros de la distribución de contraste (con el subcomando $K-S$).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Pruebas para dos muestras independientes

El procedimiento Pruebas para dos muestras independientes compara dos grupos de casos existentes en una variable.

Ejemplo. Se han desarrollado nuevos correctores dentales diseñados para que sean más cómodos y estéticos, así como para facilitar un progreso más rápido en la realineación de la dentadura. Para averiguar si el nuevo corrector debe llevarse tanto tiempo como el modelo antiguo, se eligen 10 niños al azar para que lleven este último y otros 10 para que usen el nuevo. Mediante la prueba U de Mann-Whitney podría descubrir que, de media, los niños que llevaban el nuevo corrector tenían que llevarlo puesto menos tiempo que los que llevaban el antiguo.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas: U de Mann-Whitney, reacciones extremas de Moses, Z de Kolmogorov-Smirnov, rachas de Wald-Wolfowitz.

Datos. Utilice variables numéricas que puedan ser ordenables.

Supuestos. Utilice muestras independientes y aleatorias. La prueba U de Mann-Whitney requiere que las dos muestras probadas sean similares en la forma.

Figura 34-13

Resultados del procedimiento *Dos muestras independientes*

Rangos				
Días de utilización	Aparato	N	Rango promedio	Suma de rangos
	Tipo Antiguo	10	14.10	141.00
	Tipo Nuevo	10	6.90	69.00
	Total	20		

Estadísticos de contraste ²

	Días de utilización
U de Mann-Whitney	14.000
W de Wilcoxon	69.000
Z	-2.721
Sig. asintót. (bilateral)	.007
Sig. exacta [2*(Sig. unilateral)]	.005 ¹

1. No corregidos para los empates.

2. Variable de agrupación:
Aparato

Para obtener pruebas para dos muestras independientes

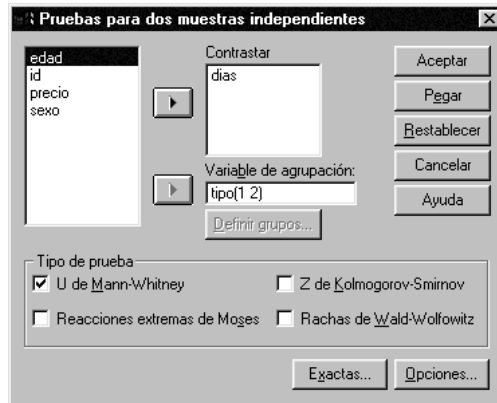
Elija en los menús:

Analizar

Pruebas no paramétricas

2 muestras independientes...

Figura 34-14
Cuadro de diálogo Pruebas para dos muestras independientes



- ▶ Seleccione una o más variables numéricas.
- ▶ Seleccione una variable de agrupación y pulse en Definir grupos para segmentar el archivo en dos grupos o muestras.

Dos muestras independientes: Tipos de pruebas

Tipo de prueba. Hay cuatro pruebas disponibles para contrastar si dos muestras (grupos) independientes proceden de una misma población.

La **prueba U de Mann-Whitney** es la más conocida de las pruebas para dos muestras independientes. Es equivalente a la prueba de la suma de rangos de Wilcoxon y a la prueba de Kruskal-Wallis para dos grupos. La prueba de Mann-Whitney contrasta si dos poblaciones muestreadas son equivalentes en su posición. Las observaciones de ambos grupos se combinan y clasifican, asignándose el rango promedio en caso de producirse empates. El número de empates debe ser pequeño en relación con el número total de observaciones. Si la posición de las poblaciones es idéntica, los rangos deberían mezclarse aleatoriamente entre las dos muestras. A continuación se calcula el número de veces que una puntuación del grupo 1 precede a una puntuación del grupo 2 y el número de veces que una puntuación del grupo 2 precede a una puntuación del grupo 1. El estadístico U de Mann-Whitney es el menor de estos dos números. El estadístico W de la suma de rangos de Wilcoxon, que también se muestra, representa la suma de los rangos de la

muestra más pequeña. Si ambas muestras tienen el mismo número de observaciones, W es la suma de los rangos del primer grupo especificado en el cuadro de diálogo Dos muestras independientes: Definir grupos.

La **prueba Z de Kolmogorov-Smirnov** y la **prueba de rachas de Wald-Wolfowitz** son pruebas más generales que detectan las diferencias entre las posiciones y las formas de las distribuciones. La prueba de Kolmogorov-Smirnov se basa en la diferencia máxima absoluta entre las funciones de distribución acumulada observadas para ambas muestras. Cuando esta diferencia es significativamente grande, se consideran diferentes las dos distribuciones. La prueba de rachas de Wald-Wolfowitz combina y ordena las observaciones de ambos grupos. Si las dos muestras proceden de una misma población, los dos grupos deben dispersarse aleatoriamente en la ordenación de los rangos.

La **prueba de reacciones extremas de Moses** presupone que la variable experimental afectará a algunos sujetos en una dirección y a otros sujetos en la dirección opuesta. Contrasta las respuestas extremas comparándolas con un grupo de control. Esta prueba se centra en la amplitud del grupo de control y supone una medida de la influencia de los valores extremos del grupo experimental en la amplitud al combinarse con el grupo de control. El grupo de control se define en el cuadro Grupo 1 del cuadro de diálogo Dos muestras independientes: Definir grupos. Las observaciones de ambos grupos se combinan y ordenan. La amplitud del grupo de control se calcula como la diferencia entre los rangos de los valores mayor y menor del grupo de control más 1. Debido a que los valores atípicos ocasionales pueden distorsionar fácilmente el rango de la amplitud, de manera automática se recorta de cada extremo un 5% de los casos de control.

Pruebas para dos muestras independientes: Definir grupos

Figura 34-15

Cuadro de diálogo Dos muestras independientes: Definir grupos

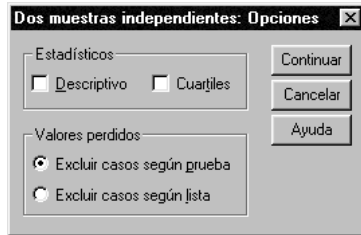


Dos muestras independientes: Definir grupos	
Grupo 1:	1
Grupo 2:	2
	Continuar
	Cancelar
	Ayuda

Para segmentar el archivo en dos grupos o muestras, introduzca un valor entero para el Grupo 1 y otro para el Grupo 2. Los casos con otros valores se excluyen del análisis.

Pruebas para dos muestras independientes: Opciones

Figura 34-16
Cuadro de diálogo *Dos muestras independientes: Opciones*



Estadísticos. Puede seleccionar uno o ambos de los estadísticos de resumen siguientes:

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

Funciones adicionales del comando *NPARTESTS* (Dos muestras independientes)

Con el lenguaje de comandos también podrá:

- Especificar el número de casos que se recortarán en la prueba de Moses (con el subcomando `MOSES`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Pruebas para dos muestras relacionadas

El procedimiento Pruebas para dos muestras relacionadas compara las distribuciones de dos variables.

Ejemplo. En general, cuando una familia vende su casa ¿logra obtener la cantidad que pide inicialmente? Si aplica la prueba de Wilcoxon de los rangos con signo a 10 casas, podría descubrir que siete familias reciben menos cantidad de la solicitada, una recibe más y dos familias reciben el precio solicitado.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas: Wilcoxon de los rangos con signo, signo, McNemar.

Datos. Utilice variables numéricas que puedan ser ordenables.

Supuestos. Aunque no se suponen distribuciones en particular para las dos variables, se supone que la distribución de población de las diferencias emparejadas es simétrica.

Figura 34-17

Resultados del procedimiento Dos muestras relacionadas

		Rangos		
		N	Rango promedio	Suma de rangos
Precio de salida - Precio de venta	Rangos negativos	1 ¹	1.50	1.50
	Rangos positivos	7 ²	4.93	34.50
	Empates	2 ³		
	Total	10		

1. Precio de salida < Precio de venta

2. Precio de salida > Precio de venta

3. Precio de venta = Precio de salida

Estadísticos de contraste ²

	Precio de salida - Precio de venta
Z	-2.313 ¹
Sig. asintót. (bilateral)	.021

1. Basado en los rangos negativos.

2. Prueba de los rangos con signo de Wilcoxon

Para obtener pruebas para dos muestras relacionadas

Elija en los menús:

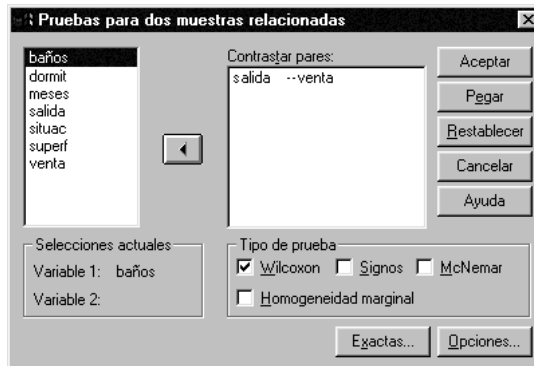
Analizar

Pruebas no paramétricas

2 muestras relacionadas...

Figura 34-18

Cuadro de diálogo Pruebas para dos muestras relacionadas



- ▶ Seleccione uno o más pares de variables, como se indica:
 - Pulse en cada una de las dos variables. La primera variable aparecerá en la sección Selecciones actuales como *Variable 1* y la segunda aparecerá como *Variable 2*.
 - Después de seleccionar una pareja de variables, pulse en el botón de flecha para incluir el par en la lista Contrastar pares. Si lo desea, pueden seleccionar más parejas de variables. Para eliminar un par de variables del análisis, seleccione el par en la lista Contrastar pares y pulse en el botón de flecha.

Dos muestras relacionadas: Tipos de pruebas

Las pruebas de esta sección comparan las distribuciones de dos variables relacionadas. La prueba apropiada depende del tipo de datos.

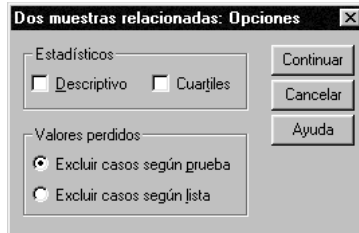
Si los datos son continuos, use la prueba de los signos o la prueba de Wilcoxon de los rangos con signo. La **prueba de los signos** calcula las diferencias entre las dos variables para todos los casos y clasifica las diferencias como positivas, negativas o empatadas. Si las dos variables tienen una distribución similar, el número de diferencias positivas y negativas no difiere de forma significativa. La **prueba de Wilcoxon de los rangos con signo** tiene en cuenta la información del signo de las diferencias y de la magnitud de las diferencias entre los pares. Dado que la prueba de Wilcoxon de los rangos con signo incorpora más información acerca de los datos, es más potente que la prueba de los signos.

Si los datos son binarios, use la **prueba de McNemar**. Esta prueba se utiliza normalmente en una situación de medidas repetidas, en la que la respuesta de cada sujeto se obtiene dos veces, una antes y otra después de que ocurra un evento especificado. La prueba de McNemar determina si la tasa de respuesta inicial (antes del evento) es igual a la tasa de respuesta final (después del evento). Esta prueba es útil para detectar cambios en las respuestas causadas por la intervención experimental en los diseños del tipo antes-después.

Si los datos son categóricos, use la **prueba de homogeneidad marginal**. Se trata de una extensión de la prueba de McNemar a partir de la respuesta binaria a la respuesta multinomial. Contrasta los cambios de respuesta utilizando la distribución de chi-cuadrado y es útil para detectar cambios de respuesta causados por intervención experimental en diseños antes-después. La prueba de homogeneidad marginal sólo está disponible si se ha instalado Pruebas exactas.

Pruebas para dos muestras relacionadas: Opciones

Figura 34-19
Cuadro de diálogo Dos muestras relacionadas: Opciones



Estadísticos. Puede seleccionar uno o ambos de los estadísticos de resumen siguientes:

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

Funciones adicionales del comando *NPARTESTS* (Dos muestras relacionadas)

Con el lenguaje de comandos también podrá:

- Contrastar una variable con todas las variables de una lista.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Pruebas para varias muestras independientes

El procedimiento Pruebas para varias muestras independientes compara dos o más grupos de casos respecto a una variable.

Ejemplo. ¿Es diferente el tiempo medio en que se fundirán las bombillas de 100 vatios de tres marcas distintas? A partir del análisis de varianza de un factor de Kruskal-Wallis, puede comprobar que las tres marcas sí se diferencian en su vida media.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas: H de Kruskal-Wallis, de la mediana.

Datos. Utilice variables numéricas que puedan ser ordenables.

Supuestos. Utilice muestras independientes y aleatorias. La prueba H de Kruskal-Wallis requiere que las muestras comparadas tengan formas similares.

Figura 34-20

Resultados del procedimiento Pruebas para varias muestras independientes

Rangos

			N	Rango promedio
Horas	Marca	Marca A	10	15,20
		Marca B	10	25,50
		Marca C	10	5,80
		Total	30	

Estadísticos de contraste ^{1,2}

	Horas de luz
Chi-cuadrado	25,061
gl	2
Sig. asintót.	.000

1. Prueba de Kruskal-Wallis

2. Variable de agrupación:
Marca

Para obtener pruebas para varias muestras independientes

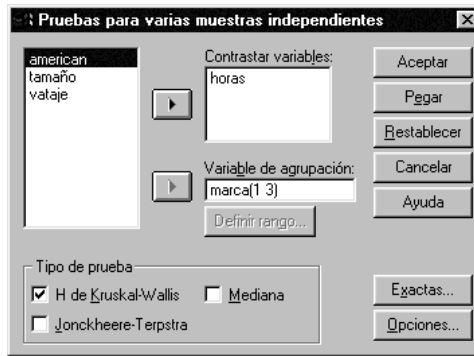
Elija en los menús:

Analizar

Pruebas no paramétricas

K muestras independientes...

Figura 34-21
Cuadro de diálogo Pruebas para varias muestras independientes



- ▶ Seleccione una o más variables numéricas.
- ▶ Seleccione una variable de agrupación y pulse en Definir rango para especificar los valores enteros máximo y mínimo para la variable de agrupación.

Tipos de pruebas para varias muestras independientes

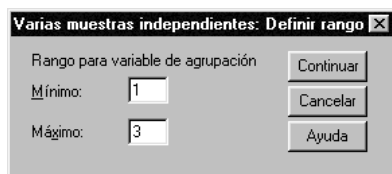
Se hallan disponibles tres pruebas para determinar si varias muestras independientes proceden de la misma población.

La prueba *H* de Kruskal-Wallis, la prueba de la mediana y la prueba de Jonckheere-Terpstra contrastan si varias muestras independientes proceden de la misma población. La prueba **H de Kruskal-Wallis**, una extensión de la prueba *U* de Mann-Whitney, es el análogo no paramétrico del análisis de varianza de un factor y detecta las diferencias en la localización de las distribuciones. La **prueba de la mediana**, que es una prueba más general pero no tan potente, detecta diferencias distribucionales en la localización y en la forma. La prueba *H* de Kruskal-Wallis y la prueba de la mediana suponen que no existe una ordenación *a priori* de las poblaciones *k* de las cuales se extraen las muestras. Cuando *existe* una ordenación natural *a priori* (ascendente o descendente) de las poblaciones *k*, la prueba **Jonckheere-Terpstra** es más potente. Por ejemplo, las *k* poblaciones pueden representar *k* temperaturas ascendentes. Se contrasta la hipótesis de que diferentes temperaturas producen la misma distribución de respuesta, con la hipótesis alternativa de que cuando la temperatura aumenta, la magnitud de la respuesta aumenta. La hipótesis alternativa se encuentra aquí ordenada; por tanto, la prueba de

Jonckheere-Terpstra es la prueba más apropiada. La prueba de Jonckheere-Terpstra estará disponible sólo si ha instalado Pruebas exactas de SPSS.

Varias muestras independientes: Definir rango

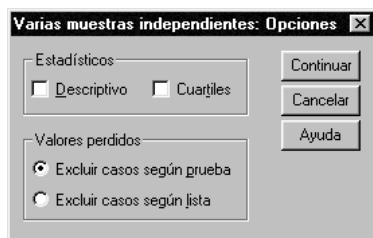
Figura 34-22
Cuadro de diálogo Varias muestras independientes: Definir



Para definir el rango, introduzca valores enteros para el máximo y el mínimo que se correspondan con las categorías mayor y menor de la variable de agrupación. Se excluyen los casos con valores fuera de los límites. Por ejemplo, si indica un valor mínimo de 1 y un valor máximo de 3, únicamente se utilizarán los valores enteros entre 1 y 3. Debe indicar ambos valores y el valor mínimo ha ser menor que el máximo.

Varias muestras independientes: Opciones

Figura 34-23
Cuadro de diálogo Varias muestras independientes: Opciones



Estadísticos. Puede elegir uno o los dos estadísticos de resumen siguientes.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Valores perdidos. Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

Funciones adicionales del comando NPAR TESTS (K muestras independientes)

Con el lenguaje de comandos también podrá:

- Especificar un valor distinto de la mediana observada para la prueba de la mediana (con el subcomando `MEDIAN`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Pruebas para varias muestras relacionadas

El procedimiento Pruebas para varias muestras relacionadas compara las distribuciones de dos o más variables.

Ejemplo. ¿Asocia la gente diferentes niveles de prestigio a doctores, abogados, policías y profesores? Se pide a diez personas que ordenen estas cuatro profesiones por orden de prestigio. La prueba de Friedman indica que la gente efectivamente asocia diferentes niveles de prestigio con estas cuatro profesiones.

Estadísticos. Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas: Friedman, W de Kendall y Q de Cochran.

Datos. Utilice variables numéricas que puedan ser ordenables.

Supuestos. Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Utilice muestras aleatorias y dependientes.

Figura 34-24
Resultados del procedimiento Pruebas para varias muestras relacionadas
 Rangos

	Rango medio
Doctor	1.50
Abogado	2.50
Policía	3.40
Profesor	2.60

Estadísticos de contraste ¹

N	10
Chi-cuadrado	10.920
gl	3
Sig. asintót.	.012

1. Prueba de Friedman

Para obtener pruebas para varias muestras relacionadas

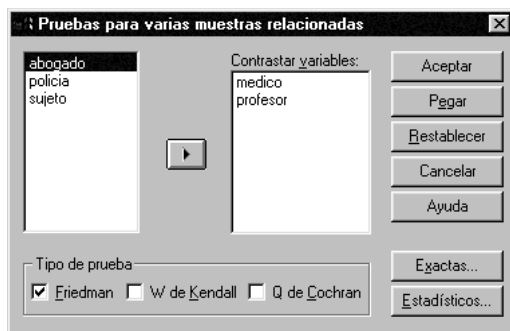
Elija en los menús:

Analizar

Pruebas no paramétricas

K muestras relacionadas...

Figura 34-25
Cuadro de diálogo Pruebas para varias muestras relacionadas



- ▶ Seleccione dos o más variables de contraste numéricas.

Tipos de prueba en el procedimiento Pruebas para varias muestras relacionadas

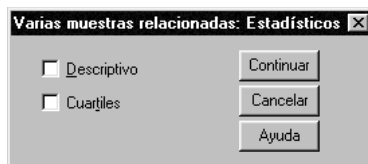
Hay tres pruebas disponibles para comparar las distribuciones de diversas variables relacionadas.

La **prueba de Friedman** es el equivalente no paramétrico de un diseño de medidas repetidas para una muestra o un análisis de varianza de dos factores con una observación por casilla. Friedman contrasta la hipótesis nula de que las k variables relacionadas procedan de la misma población. En cada caso, las k variables se ordenan en un rango de 1 a k . El estadístico de contraste se basa en estos rangos. La **W de Kendall** es una normalización del estadístico de Friedman. La prueba W de Kendall se puede interpretar como el coeficiente de concordancia, que es una medida de acuerdo entre evaluadores. Cada caso es un juez o evaluador y cada variable es un elemento o persona que está siendo evaluada. Para cada variable, se calcula la suma de rangos. La W de Kendall varía entre 0 (no hay acuerdo) y 1 (acuerdo completo). La prueba **Q de Cochran** es idéntica a la prueba de Friedman pero se puede aplicar cuando todas las respuestas son binarias. Es una extensión de la prueba de McNemar para la situación de k muestras. La Q de Cochran contrasta la hipótesis de que diversas variables dicotómicas relacionadas tienen la misma media. Las variables se miden al mismo individuo o a individuos emparejados.

Pruebas para varias muestras relacionadas: Estadísticos

Figura 34-26

Cuadro de diálogo Varias pruebas relacionadas: Estadísticos



- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

Funciones adicionales del comando NPAR TESTS (K pruebas relacionadas)

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Análisis de respuestas múltiples

Se ofrecen dos procedimientos para analizar los conjuntos de categorías múltiples y de dicotomías múltiples. El procedimiento Frecuencias de respuestas múltiples muestra tablas de frecuencias. El procedimiento Tablas de contingencia de respuestas múltiples muestra tablas de contingencia de dos y tres dimensiones. Antes de utilizar cualquiera de estos procedimientos, deberá definir conjuntos de respuestas múltiples.

Ejemplo. Este ejemplo ilustra el uso de elementos de respuestas múltiples en un estudio de investigación de mercado. Los datos son ficticios y no deben interpretarse como reales. Una línea aérea podría hacer una encuesta a los pasajeros que realicen una determinada ruta para evaluar las líneas aéreas de la competencia. En este ejemplo, American Airlines desea conocer el uso que hacen sus pasajeros de otras líneas aéreas en la ruta Chicago-Nueva York y la importancia relativa del horario y el servicio a la hora de seleccionar una línea aérea. El encargado del vuelo proporciona a cada pasajero un breve cuestionario durante el embarque. La primera pregunta dice: Rodee con un círculo todas las líneas aéreas con la que haya volado al menos una vez en los últimos seis meses en este mismo trayecto: American, United, TWA, USAir, Otras. Se trata de una pregunta de respuestas múltiples, ya que el pasajero puede marcar más de una respuesta. Sin embargo, la pregunta no se puede codificar directamente, ya que una variable sólo puede tener un valor para cada caso. Deberá utilizar distintas variables para asignar respuestas a cada pregunta. Existen dos formas de hacerlo. Una consiste en definir una variable para cada una de las opciones (por ejemplo, American, United, TWA, USAir y Otras). Si el pasajero marca United, a la variable *united* se le asignará el código 1; en caso contrario se le asignará 0. Éste es un **método de dicotomías múltiples** de asignación de variables. La otra forma de asignar respuestas es el **método de categorías múltiples**, en el que se estima el número máximo de posibles respuestas a la pregunta y se configura el mismo número de variables, con códigos para especificar la línea aérea utilizada. Examinando una muestra de cuestionarios, podría observarse que ningún usuario ha volado en más de tres líneas aéreas diferentes en esta ruta durante los últimos seis meses. Aún más,

se observará que debido a la liberalización de las líneas aéreas, aparecen otras 10 en la categoría Otras. Con el método de respuestas múltiples, definiría tres variables, cada una codificada como 1 = *american*, 2 = *united*, 3 = *twa*, 4 = *usair*, 5 = *delta* y así sucesivamente. Si un pasajero determinado marca American y TWA, la primera variable tendrá el código 1, la segunda el 3 y la tercera un código de valor perdido. Otro pasajero podría haber marcado American e introducido Delta. Así, la primera variable tendrá el código 1, la segunda el 5 y la tercera un código de valor perdido. Por el contrario, si utiliza el método de dicotomías múltiples, terminará con 14 variables independientes. Aunque cualquiera de los métodos de asignación anteriores es viable para este estudio, el método seleccionado dependerá de la distribución de respuestas.

Definir conjuntos de respuestas múltiples

El procedimiento Definir conjuntos de respuestas múltiples agrupa variables elementales en conjuntos de categorías múltiples y de dicotomías múltiples, para los que se pueden obtener tablas de frecuencias y tablas de contingencia. Se pueden definir hasta 20 conjuntos de respuestas múltiples. Cada conjunto debe tener un nombre exclusivo. Para eliminar un conjunto, resáltelo en la lista de conjuntos de respuestas múltiples y pulse en Borrar. Para cambiar un conjunto, resáltelo en la lista, modifique cualquier característica de la definición del conjunto y pulse en Cambiar.

Las variables elementales se pueden codificar como dicotomías o categorías. Para utilizar variables dicotómicas, seleccione Dicotomías para crear un conjunto de dicotomías múltiples. Introduzca un valor entero en Valor contado. Cada variable que tenga al menos una aparición del valor contado se convierte en una categoría del conjunto de dicotomías múltiples. Seleccione Categorías para crear un conjunto de categorías múltiples con el mismo rango de valores que las variables que lo componen. Introduzca valores enteros para los valores máximo y mínimo del rango para las categorías del conjunto de categorías múltiples. El procedimiento suma cada valor entero distinto en el rango inclusivo para todas las variables que lo componen. Las categorías vacías no se tabulan.

A cada conjunto de respuestas múltiples se le debe asignar un nombre exclusivo de hasta siete caracteres. El procedimiento coloca delante del nombre asignado un signo dólar (\$). No se pueden utilizar los siguientes nombres reservados: *casenum*, *sysmis*, *jdate*, *date*, *time*, *length* y *width*. El nombre del conjunto de respuestas múltiples sólo se encuentra disponible para su uso en los procedimientos de respuestas múltiples. No se puede hacer referencia a nombres de conjuntos de respuestas múltiples en otros

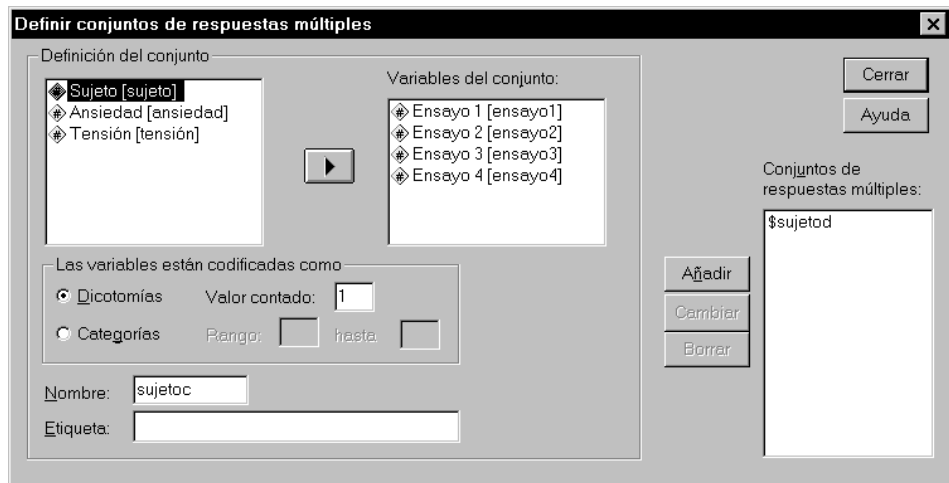
procedimientos. Si lo desea, puede introducir una etiqueta de variable descriptiva para el conjunto de respuestas múltiples. La etiqueta puede tener hasta 40 caracteres.

Para definir conjuntos de respuestas múltiples

- ▶ Elija en los menús:
 - Analizar
 - Respuesta múltiple
 - Definir conjuntos...

Figura 35-1

Cuadro de diálogo Definir conjuntos de respuestas múltiples



- ▶ Seleccione dos o más variables.
- ▶ Si las variables están codificadas como dicotomías, indique qué valor desea contar. Si las variables están codificadas como categorías, defina el rango de las categorías.
- ▶ Introduzca un nombre distintivo para cada conjunto de respuestas múltiples.
- ▶ Pulse Añadir para agregar el conjunto de respuestas múltiples a la lista de conjuntos definidos.

Frecuencias de respuestas múltiples

El procedimiento Frecuencias de respuestas múltiples produce tablas de frecuencias para conjuntos de respuestas múltiples. En primer lugar es necesario definir uno o más conjuntos de respuestas múltiples (véase “Definir conjuntos de respuestas múltiples”).

Para los conjuntos de dicotomías múltiples, los nombres de categorías que se muestran en los resultados proceden de etiquetas de variable definidas para variables elementales del grupo. Si las etiquetas de variable no están definidas, los nombres de las variables se utilizarán como etiquetas. Para los conjuntos de categorías múltiples, las etiquetas de categoría proceden de las etiquetas de valor de la primera variable del grupo. Si las categorías perdidas para la primera variable están presentes para otras variables del grupo, defina una etiqueta de valor para las categorías perdidas.

Valores perdidos. Los casos con valores perdidos se excluyen en base a tabla por tabla. Si lo desea, puede seleccionar una de las opciones siguientes o ambas:

- **Excluir los casos según lista dentro de las dicotomías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de dicotomías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de dicotomías. Por defecto, un caso se considera perdido para un conjunto de dicotomías múltiples si ninguna de sus variables que lo componen contiene el valor contado. Los casos con valores perdidos en algunas variables, pero no en todas, se incluyen en las tabulaciones del grupo si al menos una variable contiene el valor contado.
- **Excluir los casos según lista dentro de las categorías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de categorías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de categorías. Por defecto, un caso se considera perdido para un conjunto de categorías múltiples sólo si ninguno de sus componentes tiene valores válidos dentro del rango definido.

Ejemplo. Cada variable creada a partir de una pregunta de una encuesta es una variable elemental. Para analizar un elemento de respuestas múltiples, deberá combinar las variables en uno o dos tipos de conjuntos de respuestas múltiples: un conjunto de dicotomías múltiples o un conjunto de categorías múltiples. Por ejemplo, si una encuesta sobre líneas aéreas preguntara al encuestado cuál de las tres líneas (American, United, TWA) ha utilizado durante los seis últimos meses y usted utilizara variables dicotómicas y definiera un **conjunto de dicotomías múltiples**, cada una

de las tres variables del conjunto se convertiría en una categoría de la variable de grupo. Las frecuencias y los porcentajes de las tres líneas aéreas se muestran en una tabla de frecuencias. Si observa que ningún encuestado ha mencionado más de dos líneas aéreas, podría crear dos variables, cada una con tres códigos, uno para cada línea aérea. Si define un **conjunto de categorías múltiples**, los valores se tabulan añadiendo los mismos códigos en las variables elementales juntas. El conjunto de valores resultantes es igual a los de cada una de las variables elementales. Por ejemplo, 30 respuestas para United son la suma de las cinco respuestas de United para la línea aérea 1 y las 25 respuestas de United para la línea aérea 2. Las frecuencias y los porcentajes de las tres líneas aéreas se muestran en una tabla de frecuencias.

Estadísticos. Tablas de frecuencias que muestran recuentos, porcentajes de respuestas, porcentajes de casos, número de casos válidos y número de casos perdidos.

Datos. Utilice conjuntos de respuestas múltiples.

Supuestos. Las frecuencias y los porcentajes proporcionan una descripción útil de los datos de cualquier distribución.

Procedimientos relacionados. El procedimiento Definir conjuntos de respuestas múltiples permite definir este tipo de conjuntos.

Figura 35-2

Resultados de las frecuencias de respuestas múltiples

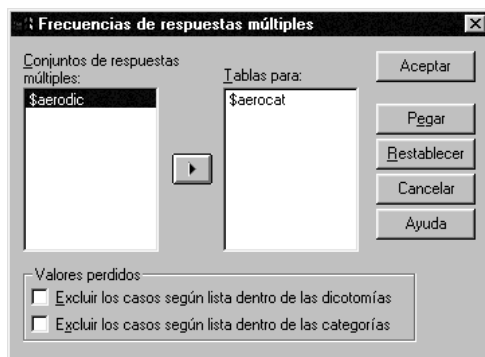
	Name ^a	Recuento		Porcentaje de casos
		N	Porcentaje	
	American	75	67.6%	92.6%
	TWA	6	5.4%	7.4%
	United	30	27.0%	37.0%
	Total	111	100.0%	137.0%

a. Grupo de dicotomía tabulado en el valor 1.

Para obtener frecuencias de respuestas múltiples

- ▶ Elija en los menús:
 - Analizar
 - Respuesta múltiple
 - Frecuencias...

Figura 35-3
Cuadro de diálogo Frecuencias de respuestas múltiples



- Seleccione uno o más conjuntos de respuestas múltiples.

Tablas de contingencia de respuestas múltiples

El procedimiento Tablas de contingencia de respuestas múltiples presenta en forma de tabla de contingencia conjuntos de respuestas múltiples, variables elementales o una combinación. También puede obtener porcentajes de casilla basados en casos o respuestas, modificar la gestión de los valores perdidos u obtener tablas de contingencia emparejadas. Antes debe definir uno o varios conjuntos de respuestas múltiples (véase “Para definir conjuntos de respuestas múltiples”).

Para los conjuntos de dicotomías múltiples, los nombres de categorías que se muestran en los resultados proceden de etiquetas de variable definidas para variables elementales del grupo. Si las etiquetas de variable no están definidas, los nombres de las variables se utilizarán como etiquetas. Para los conjuntos de categorías múltiples, las etiquetas de categoría proceden de las etiquetas de valor de la primera variable del grupo. Si las categorías perdidas para la primera variable están presentes para otras variables del grupo, defina una etiqueta de valor para las categorías perdidas. El procedimiento muestra las etiquetas de categoría por columnas en tres líneas, con un máximo de ocho caracteres por línea. Para evitar la división de palabras, puede invertir los elementos de las filas y las columnas o volver a definir las etiquetas.

Ejemplo. Tanto los conjuntos de categorías múltiples como los conjuntos de dicotomías múltiples se pueden presentar en forma de tabla de contingencia con otras variables de este procedimiento. Un estudio sobre pasajeros de líneas aéreas solicita

a éstos la siguiente información: Marque las líneas aéreas con las que ha volado al menos una vez en los seis últimos meses (American, United, TWA). ¿Qué considera más importante a la hora de seleccionar un vuelo, el horario o el servicio? Seleccione sólo uno. Después de introducir los datos como dicotomías o categorías múltiples y combinarlos en un conjunto, puede presentar en forma de tabla de contingencia las selecciones de línea aérea con la pregunta relativa al servicio o al horario.

Estadísticos. Tabla de contingencia con recuentos de casilla, fila, columna y totales, así como porcentajes de casilla, fila, columna y totales. Los porcentajes de casilla se basan en casos o respuestas.

Datos. Utilice conjuntos de respuestas múltiples o variables categóricas numéricas.

Supuestos. Las frecuencias y los porcentajes proporcionan una útil descripción de los datos de cualquier distribución.

Procedimientos relacionados. El procedimiento Definir conjuntos de respuestas múltiples permite definir este tipo de conjuntos.

Figura 35-4

Resultados de las tablas de contingencia de respuestas múltiples

		Select airline because of		Total	
		Schedule	Service		
Name	American	Recuento	41	34	75
	TWA	Recuento	3	3	6
	United	Recuento	27	3	30
Total		Recuento	44	37	81

Los porcentajes y los totales se basan en los encuestados.

a. Grupo de dicotomía tabulado en el valor 1.

Para obtener tablas de contingencia de respuestas múltiples

- Elija en los menús:

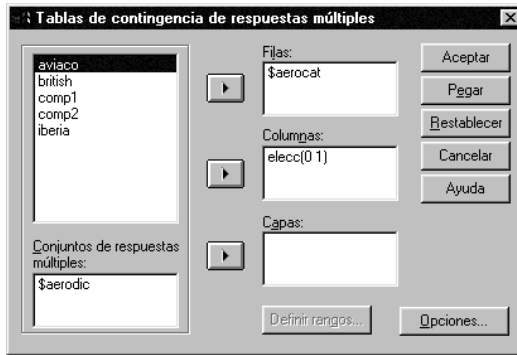
Analizar

Respuesta múltiple

Tablas de contingencia...

Figura 35-5

Cuadro de diálogo Tablas de contingencia de respuestas múltiples



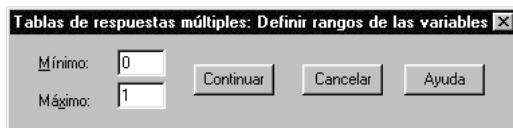
- ▶ Seleccione una o más variables numéricas o conjuntos de respuestas múltiples para cada dimensión de la tabla de contingencia.
- ▶ Defina el rango de cada variable elemental.

Si lo desea, puede obtener una tabla de contingencia de doble clasificación para cada categoría de una variable de control o conjunto de respuestas múltiples. Seleccione uno o varios elementos para la lista Capas.

Tablas de respuestas múltiples: Definir rangos de las variables

Figura 35-6

Cuadro de diálogo Tablas de respuestas múltiples: Definir rangos de las variables

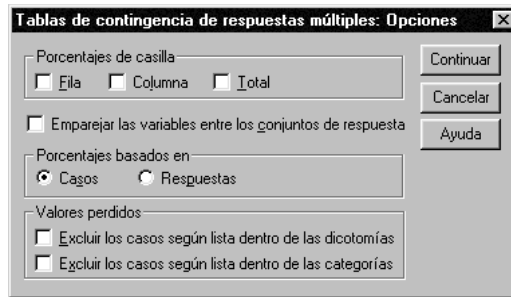


Los rangos de valores deben definirse para cualquier variable elemental de la tabla de contingencia. Introduzca los valores enteros de categoría máximos y mínimos que desee tabular. Las categorías que estén fuera del rango se excluyen del análisis. Se entiende que los valores que estén dentro del rango inclusivo son enteros (los no enteros quedan truncados).

Tablas de contingencia de respuestas múltiples: Opciones

Figura 35-7

Cuadro de diálogo Tablas de contingencia de respuestas múltiples: Opciones



Porcentajes de casilla. Las frecuencias de la casilla siempre se muestran. Puede elegir entre mostrar los porcentajes de fila, los de columna o los de tabla de doble clasificación (totales).

Porcentajes basados en. Los porcentajes de casilla pueden basarse en casos (o encuestados). Esta opción no estará disponible si selecciona la concordancia de variables en conjuntos de categorías múltiples. También se pueden basar en las respuestas. Para los conjuntos de dicotomías múltiples, el número de respuestas es igual al número de valores contados por los casos. Para los conjuntos de categorías múltiples, el número de respuestas es el número de valores del rango definido.

Valores perdidos. Puede elegir una o ambas de las siguientes opciones:

- **Excluir los casos según lista dentro de las dicotomías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de dicotomías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de dicotomías. Por defecto, un caso se considera perdido para un conjunto de dicotomías múltiples si ninguna de sus variables que lo componen contiene el valor contado. Los casos con valores perdidos para algunas variables, pero no todas, se incluyen en las tabulaciones del grupo si al menos una variable contiene el valor contado.
- **Excluir los casos según lista dentro de las categorías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de categorías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de categorías. Por defecto, un caso se considera perdido para

un conjunto de categorías múltiples sólo si ninguno de sus componentes tiene valores válidos dentro del rango definido.

Por defecto, cuando se presentan dos conjuntos de categorías múltiples en forma de tabla de contingencia, el procedimiento tabula cada variable del primer grupo con cada variable del segundo y suma las frecuencias de cada casilla; de esta forma, algunas respuestas pueden aparecer más de una vez en una tabla. Puede seleccionar la opción siguiente:

Emparejar las variables entre los conjuntos de respuesta. Empareja la primera variable del primer grupo con la primera variable del segundo, y así sucesivamente. Si selecciona esta opción, el procedimiento basará los porcentajes de casilla en las respuestas en lugar de hacerlo en los encuestados. El emparejamiento no está disponible para conjuntos de dicotomías múltiples o variables elementales.

Funciones adicionales del comando MULT RESPONSE

El lenguaje de comandos de SPSS también permite:

- Obtener tablas de contingencia de hasta cinco dimensiones (con el subcomando `BY`).
- Cambiar las opciones de formato de los resultados, incluyendo la supresión de etiquetas de valor (con el subcomando `FORMAT`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Informes de los resultados

Los listados de casos y los estadísticos descriptivos son herramientas básicas para estudiar y presentar los datos. Puede obtener listados de casos con el Editor de datos o el procedimiento Resumir, frecuencias y estadísticos descriptivos con el procedimiento Frecuencias, y estadísticos de subpoblación con el procedimiento Medias. Cada uno utiliza un formato diseñado para que la información sea clara. Si desea ver la información con otro formato, las opciones Informe de estadísticos en filas e Informe de estadísticos en columnas le ofrecen el control que precisa para presentar los datos.

Informe de estadísticos en filas

Informe de estadísticos en filas genera informes en los cuales se presentan distintos estadísticos de resumen en filas. También se encuentran disponibles listados de los casos, con o sin estadísticos de resumen.

Ejemplo. Una empresa con una cadena de tiendas registra los datos de sus empleados, incluyendo el salario, el cargo, la tienda y la sección en la que trabaja cada uno. Se podría generar un informe que proporcione los datos individuales de cada empleado (listado) desglosados por tienda y sección (variables de ruptura), con estadísticos de resumen (por ejemplo, el salario medio) por tienda, sección y sección dentro de cada tienda.

Columnas de datos. Muestra una lista de las variables del informe para las que desea obtener el listado de los casos o los estadísticos de resumen y controla el formato de presentación de las columnas de datos.

Romper columnas por. Muestra una lista de las variables de ruptura opcionales que dividen el informe en grupos y controla los estadísticos de resumen y los formatos de presentación de las columnas de ruptura. Si hay varias variables de ruptura, se

creará un grupo distinto para cada una de las categorías de las variables de ruptura dentro de las categorías de la variable de ruptura anterior en la lista. Las variables de ruptura deben ser variables categóricas discretas que dividan los casos en un número limitado de categorías con sentido. Los valores individuales de cada variable de ruptura aparecen ordenados en una columna distinta situada a la izquierda de todas las columnas de datos.

Informe. Controla las características globales del informe, incluyendo los estadísticos de resumen globales, la presentación de los valores perdidos, la numeración de las páginas y los títulos.

Mostrar casos. Muestra los valores reales (o etiquetas de valor) de las variables de la columna de datos para cada caso. Esto genera un informe a modo de listado, que puede ser mucho más largo que un informe de resumen.

Presentación preliminar. Muestra sólo la primera página del informe. Esta opción es útil para ver una presentación preliminar del formato del informe sin tener que procesar el informe completo.

Los datos están ordenados. Para los informes con variables de ruptura, el archivo de datos se debe ordenar por los valores de estas variables antes de generar el informe. Si el archivo de datos ya está ordenado por estos valores, se puede ahorrar tiempo de procesamiento seleccionando esta opción. Esta opción es especialmente útil después de generar la presentación preliminar de un informe.

Resultados de muestra

Figura 36-1
Informe combinado con listado de los casos y estadísticos de resumen

Division	Tenure	Tenure	Grade	Salary—Annual
	in Age	in Company		
Carpeting	27.00	3.67	2.17	\$9,200
	22.00	3.92	3.08	\$10,900
	23.00	3.92	3.08	\$10,900
	24.00	4.00	3.25	\$10,000
	30.00	4.08	3.08	\$10,000
	27.00	4.33	3.17	\$10,000
	33.00	2.67	2.67	\$9,335
	33.00	3.75	3.25	\$10,000
	44.00	4.83	4.33	\$15,690
	36.00	3.83	3.25	\$10,000
	35.00	3.50	3.00	\$15,520
	35.00	6.00	5.33	\$19,500
	Mean	30.75	4.04	3.31
Appliances	21.00	2.67	2.67	\$8,700
	26.00	2.92	2.08	\$8,000
	32.00	2.92	2.92	\$8,900
	33.00	3.42	2.92	\$8,900
	34.00	5.08	4.50	\$15,300
	24.00	3.17	3.17	\$8,975
	42.00	6.50	6.50	\$18,000
	30.00	2.67	2.67	\$7,500
	38.00	5.00	4.42	\$28,300
	Mean	31.11	3.81	3.54

Para obtener un informe de resumen: Estadísticos en filas

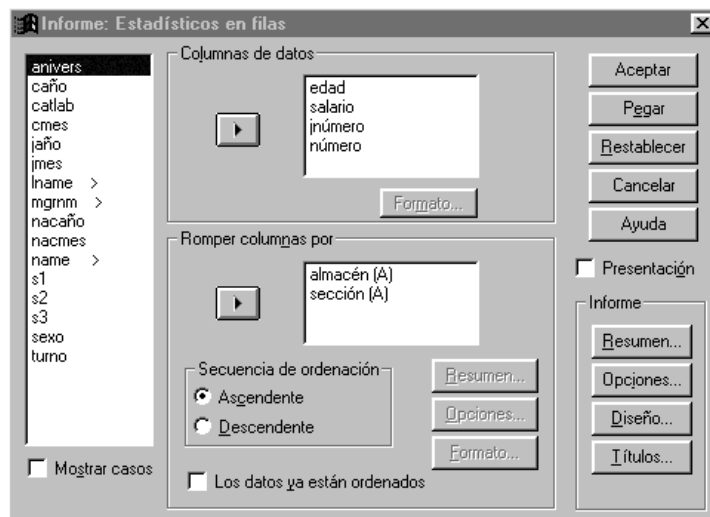
- ▶ Elija en los menús:
 Analizar
 Informes
 Informe de estadísticos en filas...
- ▶ Seleccione una o más variables para las columnas de datos. En el informe se genera una columna para cada variable seleccionada.
- ▶ Para los informes ordenados y mostrados por subgrupos, seleccione una o más variables para Romper columnas por.
- ▶ Para los informes con estadísticos de resumen para los subgrupos definidos por las variables de ruptura, seleccione la variable de ruptura de la lista Romper columnas

por y pulse en Resumen, en la sección Romper columnas por, para especificar las medidas de resumen.

- ▶ Para los informes con estadísticos de resumen globales, pulse en Resumen, en la sección Informe, para especificar las medidas de resumen.

Figura 36-2

Cuadro de diálogo Informe: Estadísticos en filas



Formato de las columnas de datos y de la ruptura de columnas del informe

Los cuadros de diálogo de formato controlan los títulos y el ancho de las columnas, la alineación del texto y la presentación de los valores de los datos o de las etiquetas de valor. El formato de las columnas de datos controla el formato de las columnas de datos situadas en la parte derecha de la página del informe. La opción Formato de ruptura controla el formato de las columnas de ruptura situadas en la parte izquierda.

Figura 36-3
Cuadro de diálogo Informe: Formato de columna de datos

Título de la columna. Para la variable seleccionada, controla el título de la columna. Los títulos largos se ajustan de forma automática dentro de la columna. Utilice la tecla Intro para insertar manualmente líneas de separación donde desee ajustar los títulos.

Posición de valor en la columna. Para la variable seleccionada, controla la alineación de los valores de los datos o de las etiquetas de valor dentro de la columna. La alineación de los valores o de las etiquetas no afecta a la alineación de los encabezados de las columnas. Puede sangrar el contenido de la columna por un número específico de caracteres o centrar el contenido.

Contenido de la columna. Para la variable seleccionada, controla la presentación de los valores de los datos o de las etiquetas de valor definidas. Los valores de los datos siempre se muestran para cualquier valor que no tenga etiquetas de valor definidas. No se encuentra disponible para las columnas de datos en los informes estadísticos en columnas.

Líneas de resumen finales y Líneas de resumen del informe

Los dos cuadros de diálogo Líneas de resumen controlan la presentación de los estadísticos de resumen para los grupos de ruptura y para el informe entero. Líneas de resumen controla los estadísticos de subgrupo para cada categoría definida por las

variables de ruptura. Líneas de resumen finales controla los estadísticos globales que se muestran al final del informe.

Figura 36-4
Cuadro de diálogo Informe: Líneas de resumen

Informe: Líneas de resumen para sección

Suma de valores

Media de los valores

Valor mínimo

Valor máximo

Número de casos

Porcentaje por encima Valor:

Porcentaje por debajo Valor:

Porcentaje dentro Menor: Mayor:

Desviación típica

Curtosis

Varianza

Asimetría

Continuar

Cancelar

Ayuda

Los estadísticos de resumen disponibles son: suma, media, valor mínimo, valor máximo, número de casos, porcentaje de casos por encima y por debajo de un valor especificado, porcentaje de casos dentro de un rango de valores especificado, desviación típica, curtosis, varianza y asimetría.

Opciones de ruptura del informe

Opciones de ruptura controla el espaciado y la paginación de la información de la categoría de ruptura.

Figura 36-5
Cuadro de diálogo Informe: Opciones de ruptura

Informe: Opciones de ruptura para sección

Control de página

Saltar líneas antes de ruptura:

Empezar en página siguiente

Empezar nueva página y restablecer nº de página

Líneas en blanco antes de los estadísticos

Continuar

Cancelar

Ayuda

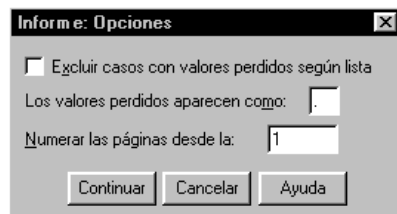
Control de página. Controla el espaciado y la paginación para las categorías de la variable de ruptura seleccionada. Puede especificar un número de líneas en blanco entre las categorías de ruptura o empezar cada categoría de ruptura en una página nueva.

Líneas en blanco antes de los estadísticos. Controla el número de líneas en blanco entre las etiquetas o los datos de la categoría de ruptura y los estadísticos de resumen. Esta opción es especialmente útil para los informes combinados que incluyan tanto el listado de los casos individuales como los estadísticos de resumen para las categorías de ruptura; en estos informes puede insertar un espacio entre el listado de los casos y los estadísticos de resumen.

Opciones del informe

Informe: Opciones controla el tratamiento y la presentación de los valores perdidos y la numeración de las páginas del informe.

Figura 36-6
Cuadro de diálogo Informe: Opciones



Excluir casos con valores perdidos según lista. Elimina (del informe) cualquier caso con valores perdidos para cualquier variable del informe.

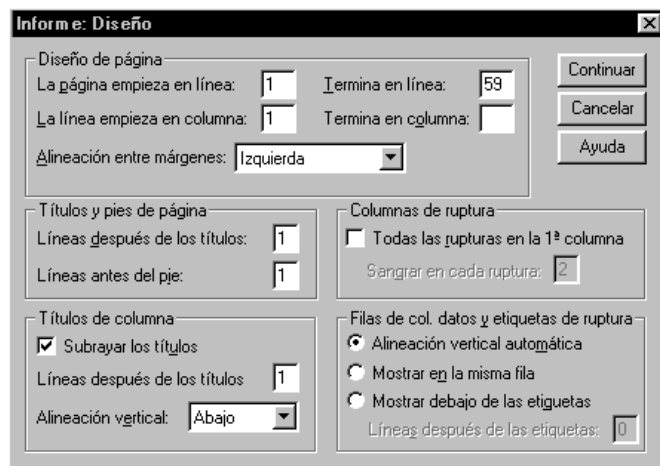
Los valores perdidos aparecen como. Permite especificar el símbolo que representa los valores perdidos en el archivo de datos. Este símbolo sólo puede tener un carácter y se utiliza para representar tanto los valores **perdidos del sistema** como los valores **perdidos definidos por el usuario**.

Numerar las páginas desde la. Permite especificar un número de página para la primera página del informe.

Diseño del informe

Informe: Diseño controla el ancho y alto de cada página del informe, la ubicación del informe dentro de la página y la inserción de etiquetas y líneas en blanco.

Figura 36-7
Cuadro de diálogo Informe: Diseño



Diseño de página. Controla los márgenes de las páginas expresados en líneas (extremos superior e inferior) y caracteres (a la izquierda y a la derecha) y la alineación del informe entre los márgenes.

Títulos y pies de página. Controla el número de líneas que separan los títulos y los pies de página del cuerpo del informe.

Romper columnas por. Controla la presentación de las columnas de ruptura. Si se especifican diversas variables de ruptura, pueden situarse en columnas diferentes o en la primera columna. Si se colocan todas en la primera columna, se generará un informe más estrecho.

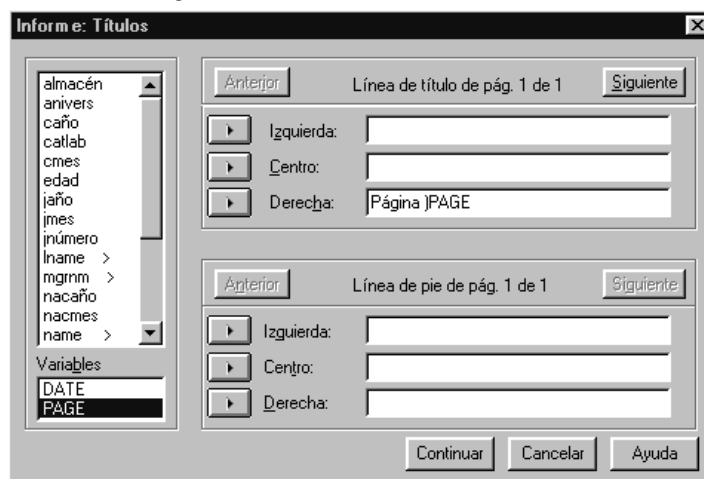
Títulos de columna. Controla la presentación de los títulos de columna, incluyendo el subrayado de títulos, el espacio entre los títulos y el cuerpo del informe y la alineación vertical de los títulos de columna.

Filas de col. datos y etiquetas de ruptura. Controla la ubicación de la información de las columnas de datos (valores de datos o estadísticos de resumen) en relación con las etiquetas de ruptura al principio de cada categoría de ruptura. La primera fila de información puede empezar en la misma línea que la etiqueta de categoría de ruptura o en un número de líneas posterior especificado. Esta sección no se encuentra disponible para los informes de estadísticos en columnas.

Títulos del informe

Informe: Títulos controla el contenido y la ubicación de los títulos y los pies de página del informe. Puede especificar un máximo de diez líneas de títulos de página y otras tantas de pies de página, con componentes justificados a la izquierda, en el centro y a la derecha en cada línea.

Figura 36-8
Cuadro de diálogo Informe: Títulos



Si inserta variables en los títulos o en los pies de página, la etiqueta de valor o el valor de la variable actual aparecerá en el título o en el pie de página. Para los títulos se mostrará la etiqueta de valor correspondiente al valor de la variable al principio de la página; para los pies de página, esta etiqueta se mostrará al final de la página. Si no hay etiqueta de valor, se mostrará el valor real.

Variables especiales. Las variables especiales *DATE* y *PAGE* permiten insertar la fecha actual o el número de página en cualquier línea de un encabezado o pie del informe. Si el archivo de datos contiene variables llamadas *DATE* o *PAGE*, no podrá utilizar estas variables en los títulos ni en los pies del informe.

Informe de estadísticos en columnas

Informe de estadísticos en columnas genera informes de resumen en los que diversos estadísticos de resumen aparecen en columnas distintas.

Ejemplo. Una empresa con una cadena de tiendas registra la información de los empleados, incluyendo el salario, el cargo y la sección en la que trabaja cada uno. Se podría generar un informe que proporcione los estadísticos de salario resumidos (por ejemplo, media, mínimo, máximo) para cada sección.

Columnas de datos. Muestra una lista de las variables del informe para las que se desea obtener estadísticos de resumen y controla el formato de presentación y los estadísticos de resumen mostrados para cada variable.

Romper columnas por. Muestra una lista de las variables de segmentación opcionales que dividen el informe en grupos y controla los formatos de presentación de las columnas de ruptura. Si hay varias variables de ruptura, se creará un grupo distinto para cada una de las categorías de las variables de ruptura dentro de las categorías de la variable de ruptura anterior en la lista. Las variables de ruptura deben ser variables categóricas discretas que dividan los casos en un número limitado de categorías con sentido.

Informe. Controla las características globales del informe, incluyendo la presentación de los valores perdidos, la numeración de las páginas y los títulos.

Presentación preliminar. Muestra sólo la primera página del informe. Esta opción es útil para ver una presentación preliminar del formato del informe sin tener que procesar el informe completo.

Los datos están ordenados. Para los informes con variables de ruptura, el archivo de datos se debe ordenar por los valores de estas variables antes de generar el informe. Si el archivo de datos ya está ordenado por estos valores, se puede ahorrar tiempo de procesamiento seleccionando esta opción. Esta opción es especialmente útil después de generar la presentación preliminar de un informe.

Resultados de muestra

Figura 36-9

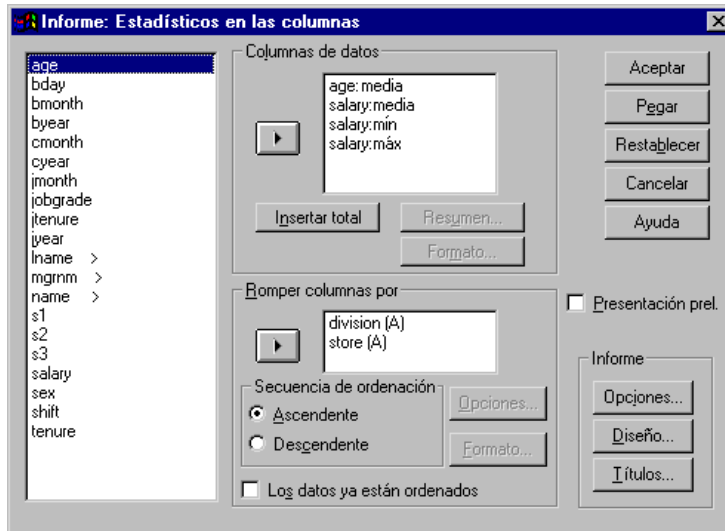
Informe de resumen con estadísticos de resumen en columnas

Division	Mean	Minimum	Maximum	
	Annual	Annual	Annual	Salary
Division	Mean Age	Salary	Salary	Salary
Carpeting	30.75	\$11,754	\$9,200	\$19,500
Appliances	31.11	\$12,508	\$7,500	\$28,300
Furniture	36.87	\$13,255	\$8,975	\$17,050
Hardware	36.20	\$17,580	\$7,450	\$22,500

Para obtener un informe de resumen: Estadísticos en columnas

- ▶ Elija en los menús:
 - Analizar
 - Informes
 - Informe de estadísticos en columnas...
- ▶ Seleccione una o más variables para las columnas de datos. En el informe se genera una columna para cada variable seleccionada.
- ▶ Para cambiar la medida de resumen para una variable, seleccione la variable de la lista de columnas de datos y pulse en Resumen.
- ▶ Para obtener más de una medida de resumen para una variable, seleccione la variable en la lista de origen y desplácela hasta la lista Columnas de datos varias veces, una para cada medida que desee obtener.
- ▶ Para mostrar una columna con la suma, la media, la razón o cualquier otra función de las columnas existentes, pulse en Insertar total. Al hacerlo se situará una variable llamada *total* en la lista Columnas de datos.
- ▶ Para los informes ordenados y mostrados por subgrupos, seleccione una o más variables para Romper columnas por.

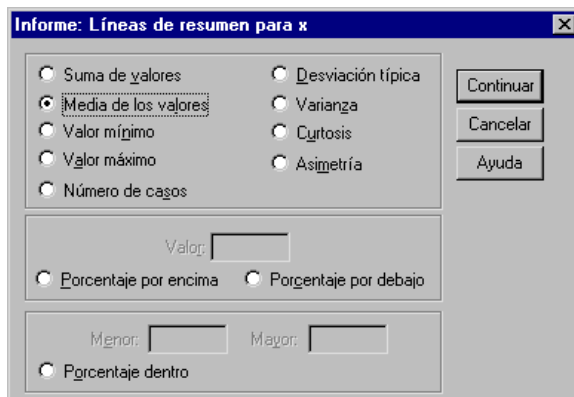
Figura 36-10
Cuadro de diálogo Informe: Estadísticos en las columnas



Función Columna de resumen total

Líneas de resumen controla el estadístico de resumen mostrado para la variable de las columnas de datos seleccionada.

Figura 36-11
Cuadro de diálogo Informe: Líneas de resumen



Los estadísticos de resumen disponibles son: suma, media, mínimo, máximo, número de casos, porcentaje de casos por encima o por debajo de un valor especificado, porcentaje de casos dentro de un rango especificado de valores, desviación típica, varianza, curtosis y asimetría.

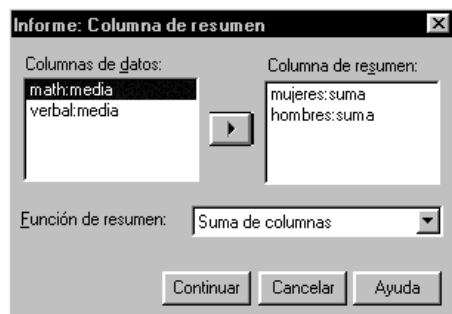
Columna de resumen total

Columna de resumen controla los estadísticos de resumen del total que resumen dos o más columnas de datos.

Los estadísticos de resumen del total son la suma de columnas, la media de columnas, el mínimo, el máximo, la diferencia entre los valores de dos columnas, el cociente de los valores de una columna dividido por los valores de otra y el producto de los valores de las columnas multiplicados entre sí.

Figura 36-12

Cuadro de diálogo Informe: Columna de resumen



Suma de columnas. La columna *total* es la suma de las columnas de la lista Columna de resumen.

Media de columnas. La columna *total* es la media de las columnas de la lista Columna de resumen.

Mínimo de columnas. La columna *total* es el mínimo de las columnas de la lista Columna de resumen.

Máximo de columnas. La columna *total* es el máximo de las columnas de la lista Columna de resumen.

1ª columna – 2ª columna. La columna *total* es la resta de las columnas de la lista Columna de resumen. Esta lista debe contener, exactamente, dos columnas.

1ª columna / 2ª columna. La columna *total* es el cociente de las columnas de la lista Columna de resumen. Esta lista debe contener, exactamente, dos columnas.

% 1ª columna / 2ª columna. La columna *total* es el porcentaje de la primera columna dividido por la segunda columna de la lista Columna de resumen. Esta lista debe contener, exactamente, dos columnas.

Producto de columnas. La columna *total* es el producto de las columnas de la lista Columna de resumen.

Formato de columna del informe

Las opciones de formato de columna de ruptura y de datos para el informe de estadísticos en columnas son las mismas que las descritas para el informe de estadísticos en filas.

Opciones de la ruptura de columnas para los estadísticos en el informe

Opciones de ruptura controla la presentación del subtotal, el espaciado y la paginación para las categorías de ruptura.

Figura 36-13
Cuadro de diálogo Informe: Opciones de ruptura

Informe: Opciones de ruptura para verbal

Subtotal

Mostrar subtotal

Etiqueta:
Subtotal verbal

Continuar

Cancelar

Ayuda

Control de página

Saltar líneas antes de ruptura: 1

Empezar en página siguiente

Empezar nueva página y restablecer nº de página

Líneas en blanco antes del subtotal: 0

Subtotal. Controla los subtotales mostrados para cada categoría de ruptura.

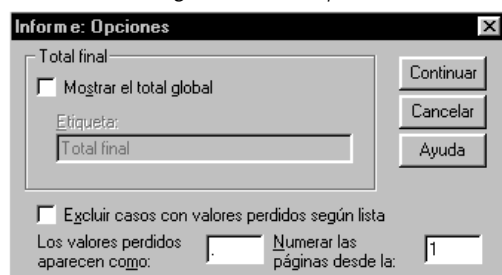
Control de página. Controla el espaciado y la paginación para las categorías de la variable de ruptura seleccionada. Puede especificar un número de líneas en blanco entre las categorías de ruptura o empezar cada categoría de ruptura en una página nueva.

Líneas en blanco antes del subtotal. Controla el número de líneas en blanco entre los datos de las categorías de ruptura y los subtotales.

Opciones de columnas para los estadísticos en el informe

Opciones controla la presentación de los totales finales y de los valores perdidos y la paginación de los informes de estadísticos en columnas.

Figura 36-14
Cuadro de diálogo Informe: Opciones



Total final. Muestra y etiqueta un total global para cada columna que aparece al final de la columna.

Valores perdidos. Permite excluir los valores perdidos del informe o seleccionar un único carácter para indicar estos valores.

Diseño del informe para los estadísticos en columnas

Las opciones de diseño para el informe de estadísticos en columnas son las mismas que las descritas para el informe de estadísticos en filas.

Funciones adicionales del comando REPORT

El lenguaje de comandos de SPSS también permite:

- Mostrar funciones de resumen diferentes en las columnas de una única línea de resumen.
- Insertar líneas de resumen en las columnas de datos para variables que no sean la variable de la columna de datos o para diversas combinaciones (funciones compuestas) de las funciones de resumen.
- Utilizar Mediana, Moda, Frecuencia y Porcentaje como funciones de resumen.
- Controlar de forma más precisa el formato de presentación de los estadísticos de resumen.
- Insertar líneas en blanco en diversos puntos de los informes.
- Insertar líneas en blanco después de cada n-ésimo caso de los informes en formato de listado.

Debido a la complejidad de la sintaxis de `REPORT`, a la hora de construir un nuevo informe con sintaxis puede resultarle útil, para aproximar el informe generado a partir de los cuadros de diálogo, copiar y pegar la sintaxis correspondiente y depurar esa sintaxis para generar exactamente el informe que le interese.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Análisis de fiabilidad

El análisis de fiabilidad permite estudiar las propiedades de las escalas de medición y de los elementos que las constituyen. El procedimiento Análisis de fiabilidad calcula un número de medidas de fiabilidad de escala que se utilizan normalmente y también proporciona información sobre las relaciones entre elementos individuales de la escala. Se pueden utilizar los coeficientes de correlación intraclase para calcular estimaciones de la fiabilidad inter-evaluadores.

Ejemplo. ¿ El cuestionario mide la satisfacción del cliente de manera útil? El análisis de fiabilidad le permitirá determinar el grado en que los elementos del cuestionario se relacionan entre sí, obtener un índice global de la replicabilidad o de la consistencia interna de la escala en su conjunto e identificar elementos problemáticos que deberían ser excluidos de la escala.

Estadísticos. Descriptivos para cada variable y para la escala, estadísticos de resumen comparando los elementos, correlaciones y covarianzas inter-elementos, estimaciones de la fiabilidad, tabla de ANOVA, coeficientes de correlación intraclase, T^{cuadrado} de Hotelling y prueba de aditividad de Tukey.

Modelos. Están disponibles los siguientes modelos de fiabilidad:

- **Alfa (Cronbach).** Es un modelo de consistencia interna, que se basa en la correlación inter-elementos promedio.
- **Dos mitades.** Este modelo divide la escala en dos partes y examina la correlación entre dichas partes.
- **Guttman.** Este modelo calcula los límites inferiores de Guttman para la fiabilidad verdadera.
- **Paralelo.** Este modelo asume que todos los elementos tienen varianzas iguales y varianzas error iguales a través de las réplicas.
- **Paralelo estricto.** Este modelo asume los supuestos del modelo paralelo y también asume que las medias son iguales a través de los elementos.

Datos. Los datos pueden ser dicotómicos, ordinales o de intervalo, pero deben estar codificados numéricamente.

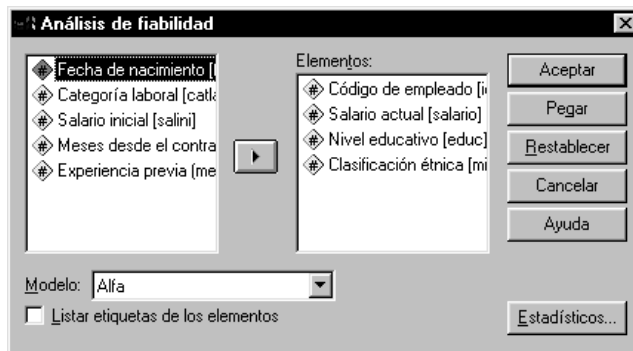
Supuestos. Las observaciones deben ser independientes y los errores no deben estar correlacionados entre los elementos. Cada par de elementos debe tener una distribución normal bivariada. Las escalas deben ser aditivas, de manera que cada elemento esté linealmente relacionado con la puntuación total.

Procedimientos relacionados. Si desea explorar la dimensionalidad de los elementos de la escala (para comprobar si es necesario más de un constructo para explicar el patrón de puntuaciones en los elementos), utilice el Análisis factorial o el Escalamiento multidimensional. Para identificar grupos homogéneos de variables, puede utilizar Análisis de conglomerados jerárquico para agrupar las variables.

Para obtener un análisis de fiabilidad

- ▶ Elija en los menús:
 Analizar
 Escala
 Análisis de fiabilidad...

Figura 37-1
 Cuadro de diálogo Análisis de fiabilidad



- ▶ Seleccione dos o más variables como componentes potenciales de una escala aditiva.
- ▶ Elija un modelo de la lista desplegable Modelo.

Análisis de fiabilidad: Estadísticos

Figura 37-2
Cuadro de diálogo Análisis de fiabilidad: Estadísticos

Puede seleccionar diversos estadísticos para describir la escala y sus elementos. Los estadísticos de los que se informa por defecto incluyen el número de casos, el número de elementos y las estimaciones de la fiabilidad, según se explica a continuación:

- Modelos Alfa: coeficiente alfa. Para datos dicotómicos, éste es equivalente al coeficiente 20 de Kuder-Richardson (KR20).
- Modelos de Dos mitades: correlación entre formas, fiabilidad de dos mitades de Guttman, fiabilidad de Spearman-Brown (longitud igual y desigual) y coeficiente alfa para cada mitad.
- Modelos de Guttman: coeficientes de fiabilidad lambda 1 a lambda 6.
- Modelos de Paralelo y Paralelo estricto: prueba de bondad de ajuste del modelo, estimaciones de la varianza error, varianza común y varianza verdadera, correlación común inter-elementos estimada, fiabilidad estimada y estimación de la fiabilidad insesgada.

Descriptivos para. Genera estadísticos descriptivos para las escalas o los elementos a través de los casos. Las opciones disponibles son: Elemento, Escala y Escala si se elimina el elemento.

- **Escala si se elimina el elemento.** Muestra estadísticos de resumen para comparar cada elemento con la escala compuesta por otros elementos. Los estadísticos incluyen la media de escala y la varianza si el elemento se ha eliminado de la escala, la correlación entre el elemento y la escala compuesta por otros elementos, y alfa de Cronbach si el elemento se ha eliminado de la escala.

Resúmenes. Proporciona estadísticos descriptivos sobre las distribuciones de los elementos a través de todos los elementos de la escala. Las opciones disponibles son: Medias, Varianzas, Covarianzas y Correlaciones.

- **Medias (Fiabilidad).** Estadísticos de resumen para las medias de los elementos. Se muestran el máximo, el mínimo y el promedio de las medias de los elementos, el rango y la varianza de las medias de los elementos, y la razón de la mayor media sobre la menor media de los elementos.
- **Varianzas.** Estadísticos de resumen de las varianzas de los elementos (ítems). Se muestran: la menor, la mayor y el promedio de las varianzas de los ítems, el rango y la varianza de las varianzas de los ítems, y la razón de la mayor varianza sobre la menor varianza de los ítems.
- **Covarianzas.** Estadísticos de resumen para las covarianzas entre los elementos. Se muestran el máximo, el mínimo y el promedio de las covarianzas inter-elementos, el rango y la varianza de las covarianzas inter-elementos, y la razón de la mayor covarianza sobre la menor covarianza inter-elementos.
- **Correlaciones.** Estadísticos de resumen para las correlaciones entre los elementos. Se muestran el máximo, el mínimo y el promedio de las correlaciones inter-elementos, el rango y la varianza de las correlaciones inter-elementos, y la razón de la mayor correlación sobre la menor correlación inter-elementos.

Inter-elementos. Genera las matrices de correlaciones o covarianzas entre los elementos.

Tabla de ANOVA. Produce pruebas de medias iguales. Las opciones disponibles son: Ninguno, Prueba *F*, Chi-cuadrado de Friedman y Chi-cuadrado de Cochran.

- **Prueba F (Fiabilidad).** Muestra la tabla de un análisis de varianza de medidas repetidas.

- **Chi-cuadrado de Friedman.** Muestra el chi-cuadrado de Friedman y el coeficiente de concordancia de Kendall. Esta opción es adecuada para datos que se encuentren en el formato de rangos. La prueba de chi-cuadrado sustituye a la prueba F habitual en la tabla de ANOVA.
- **Chi-cuadrado de Cochran.** Muestra la Q de Cochran. Esta opción es adecuada para datos que son dicotómicos. El estadístico Q sustituye a la F habitual en la tabla de ANOVA.

T-cuadrado de Hotelling. Genera un contraste multivariado sobre la hipótesis nula de que todos los elementos de la escala tienen la misma media.

Prueba de aditividad de Tukey. Genera un contraste sobre el supuesto de que no existe una interacción multiplicativa entre los elementos.

Coefficiente de correlación intraclass. Genera medidas sobre la consistencia o sobre el acuerdo de los valores entre los propios casos.

- **Modelo.** Seleccione el modelo para calcular el coeficiente de correlación intraclass. Los modelos disponibles son: Dos factores, efectos mixtos; Dos factores, efectos aleatorios; y Un factor, efectos aleatorios. Seleccione Dos factores, efectos mixtos, si los efectos de personas son aleatorios y los efectos de elementos son fijos, Dos factores, efectos aleatorios, si los efectos de personas y los efectos de elementos son aleatorios, y Un factor, efectos aleatorios, si los efectos de personas son aleatorios.
- **Tipo.** Seleccione el tipo de índice. Los tipos disponibles son: Los tipos disponibles son: Consistencia y Acuerdo absoluto.
- **Intervalo de confianza.** Especifica el nivel para el intervalo de confianza. El valor por defecto es 95%.
- **Valor de prueba.** Especifica el valor hipotetizado para el coeficiente, para el contraste de hipótesis. Este es el valor con el que se compara el valor observado. El valor por defecto es 0.

Funciones adicionales del comando RELIABILITY

El lenguaje de comandos de SPSS también permite:

- Leer y analizar una matriz de correlaciones.

- Escribir una matriz de correlaciones para su análisis posterior.
- Especificar una división distinta de la de dos mitades iguales, para el método de dos mitades.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Escalamiento multidimensional

El escalamiento multidimensional trata de encontrar la estructura de un conjunto de medidas de distancia entre objetos o casos. Esto se logra asignando las observaciones a posiciones específicas en un espacio conceptual (normalmente de dos o tres dimensiones) de modo que las distancias entre los puntos en el espacio concuerden al máximo con las disimilaridades dadas. En muchos casos, las dimensiones de este espacio conceptual son interpretables y se pueden utilizar para comprender mejor los datos. Si las variables se han medido objetivamente, puede utilizar el escalamiento multidimensional como técnica de reducción de datos (el procedimiento Escalamiento multidimensional permitirá calcular las distancias a partir de los datos multivariados, si es necesario). El escalamiento multidimensional puede también aplicarse a valoraciones subjetivas de disimilaridad entre objetos o conceptos. Además, el procedimiento Escalamiento multidimensional puede tratar datos de disimilaridad procedentes de múltiples fuentes, como podrían ser múltiples evaluadores o múltiples sujetos evaluados por un cuestionario.

Ejemplo. ¿Cómo percibe el público las diferencias entre distintos coches? Si posee datos de las valoraciones de similaridad emitidas por los sujetos sobre las diferentes marcas y modelos de coches, puede utilizar el escalamiento multidimensional para identificar las dimensiones que describan las preferencias de los consumidores. Puede encontrar, por ejemplo, que el precio y el tamaño de un vehículo definen un espacio de dos dimensiones, capaz de explicar las similitudes de las que informan los encuestados.

Estadísticos. Para cada modelo: Matriz de datos, Matriz de datos escalada óptimamente, S-stress (de Young), Stress (de Kruskal), R^2 , Coordenadas de los estímulos, Stress promedio y R^2 para cada estímulo (modelos RMDS). Para modelos de diferencias individuales (INDSCAL): ponderaciones del sujeto e índice de peculiaridad para cada sujeto. Para cada matriz en los modelos de escalamiento multidimensional replicado: stress y R^2 para cada estímulo. Gráficos: coordenadas de

los estímulos (de dos o tres dimensiones), diagrama de dispersión de las disparidades frente a las distancias.

Datos. Si los datos son de disimilaridad, todas las disimilaridades deben ser cuantitativas y deben estar medidas en la misma métrica. Si los datos son datos multivariantes, las variables pueden ser datos cuantitativos, binarios o de recuento. El escalamiento de las variables es un tema importante, ya que las diferencias en el escalamiento pueden afectar a la solución. Si las variables tienen grandes diferencias en el escalamiento (por ejemplo, una variable se mide en dólares y otra en años), debe considerarse el tipificarlas (esto puede llevarse a cabo automáticamente con el propio procedimiento Escalamiento multidimensional).

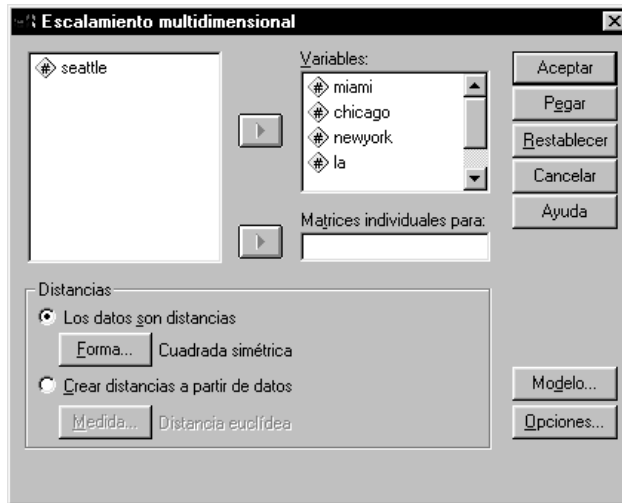
Supuestos. El procedimiento Escalamiento multidimensional está relativamente libre de supuestos distribucionales. Compruebe que selecciona el nivel de medida adecuado (ordinal, de intervalo, o de razón) en Opciones para asegurar que los resultados se calculan correctamente.

Procedimientos relacionados. Si su objetivo es la reducción de los datos, un método alternativo a tener en cuenta es el análisis factorial, sobre todo si las variables son cuantitativas. Si desea identificar grupos de casos similares, considere complementar el análisis de escalamiento multidimensional con un análisis de conglomerados jerárquico o de k -medias.

Para obtener un análisis de escalamiento multidimensional

- ▶ Elija en los menús:
 - Analizar
 - Escala
 - Escalamiento multidimensional...

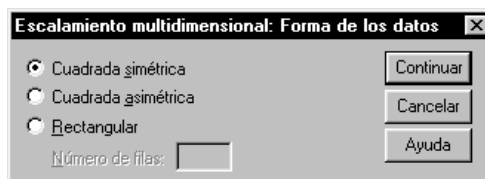
Figura 38-1
Cuadro de diálogo Escalamiento multidimensional



- ▶ En Distancias, seleccione Los datos son distancias o Crear distancias a partir de datos.
- ▶ Si los datos son distancias, debe seleccionar cuatro variables numéricas como mínimo para el análisis y pulsar en Forma para indicar la forma de la matriz de distancias.
- ▶ Si desea que SPSS cree las distancias antes de analizarlas, debe seleccionar al menos una variable numérica y puede pulsar en Medida para especificar el tipo de medida de distancia que desea. Puede crear matrices distintas para cada categoría de una variable de agrupación (la cual puede ser numérica o de cadena) moviendo esa variable a la lista Matrices individuales para.

Escalamiento multidimensional: Forma de los datos

Figura 38-2
Cuadro de diálogo Escalamiento multidimensional: Forma de los datos



Si el archivo de datos de trabajo representa distancias entre un o dos conjuntos de objetos, debe especificar la forma de la matriz de datos para obtener los resultados correctos. Elija una alternativa: Cuadrada simétrica, Cuadrada asimétrica o Rectangular. *Nota:* No puede seleccionar Cuadrada simétrica si el cuadro de diálogo Modelo especifica la condicionalidad de filas.

Escalamiento multidimensional: Crear la medida a partir de los datos

Figura 38-3

Cuadro de diálogo Escalamiento multidimensional: Crear la medida a partir de los datos

El escalamiento multidimensional utiliza datos de disimilaridad para crear una solución de escalamiento. Si los datos son datos multivariantes (los valores de las variables que se han medido), debe crear los datos de disimilaridad para poder calcular una solución de escalamiento multidimensional. Puede especificar los detalles para la creación de las medidas de disimilaridad a partir de los datos.

Medida. Le permite especificar la medida de disimilaridad para el análisis. Seleccione una opción del grupo Medida que se corresponda con el tipo de datos y, a continuación, seleccione una de las medidas de la lista desplegable correspondiente a ese tipo de medida. Las opciones disponibles son:

- **Intervalo.** Distancia euclídea, Distancia euclídea al cuadrado, Chebychev, Bloque, Minkowski o Personalizada.

- **Contar apariciones.** Medida de chi-cuadrado o Medida de phi-cuadrado.
- **Binaria.** Distancia euclídea, Distancia euclídea al cuadrado, Diferencia de tamaño, Diferencia de configuración, Varianza o Lance y Williams.

Crear matriz de distancias. Le permite elegir la unidad de análisis. Las opciones son Entre variables o Entre casos.

Transformar valores. En determinados casos, como cuando las variables se miden en escalas muy distintas, puede que desee tipificar los valores antes de calcular las proximidades (no es aplicable a datos binarios). Seleccione un método de estandarización de la lista desplegable Estandarizar (si no se requiere ninguna estandarización, seleccione Ninguna).

Escalamiento multidimensional: Modelo

Figura 38-4
Cuadro de diálogo Escalamiento multidimensional: Modelo

La estimación correcta de un modelo de escalamiento multidimensional depende de aspectos que atañen a los datos y al modelo en sí.

Nivel de medida. Le permite especificar el nivel de medida de los datos. Las opciones son Ordinal, Intervalo y Razón. Si las variables son ordinales, al seleccionar Desempatar observaciones empatadas se solicitará que sean consideradas como variables continuas, de forma que los empates (valores iguales para casos diferentes) se resuelvan óptimamente.

Condicionabilidad. Le permite especificar qué comparaciones tienen sentido. Las opciones son Matriz, Fila o Incondicional.

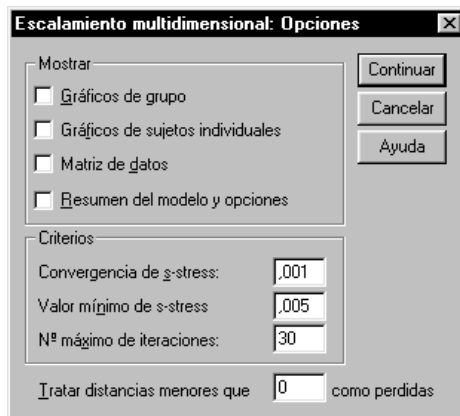
Dimensiones. Le permite especificar la dimensionalidad de la solución o soluciones del escalamiento. Se calcula una solución para cada número del rango especificado. Especifique números enteros entre 1 y 6; se permite un mínimo de 1 sólo si selecciona Distancia euclídea como modelo de escalamiento. Para una solución única, especifique el mismo número para el mínimo y el máximo.

Modelo de escalamiento. Le permite especificar los supuestos bajo los que se realiza el escalamiento. Las opciones disponibles son Distancia euclídea o Distancia euclídea de diferencias individuales (también conocida como INDSCAL). Para el modelo de Distancia euclídea de diferencias individuales, puede seleccionar Permitir ponderaciones negativas de los sujetos, si es adecuado para los datos.

Escalamiento multidimensional: Opciones

Figura 38-5

Cuadro de diálogo Escalamiento multidimensional: Opciones



Puede especificar opciones para el análisis de escalamiento multidimensional:

Mostrar. Le permite seleccionar varios tipos de resultados. Las opciones disponibles son Gráficos de grupo, Gráficos para los sujetos individuales, Matriz de datos y Resumen del modelo y de las opciones.

Criterios. Le permite determinar cuándo debe detenerse la iteración. Para cambiar los valores por defecto, introduzca valores para la Convergencia de s-stress, el Valor mínimo de s-stress y el N° máximo de iteraciones.

Tratar distancias menores que n como perdidas. Las distancias menores que este valor se excluyen del análisis.

Funciones adicionales del comando de ALSCAL

El lenguaje de comandos de SPSS también permite:

- Utilizar tres tipos de modelos adicionales, conocidos como ASCAL, AINDS y GEMSCAL en la literatura del escalamiento multidimensional.
- Realizar transformaciones polinómicas en los datos de intervalo y de razón.
- Analizar similaridades (en lugar de distancias) con los datos ordinales.
- Analizar datos nominales.
- Guardar en archivos varias matrices de ponderación y coordenadas y leerlas posteriormente para el análisis.
- Restringir el desplegamiento multidimensional.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Estadísticos de la razón

El procedimiento Estadísticos de la razón proporciona una amplia lista de estadísticos de resumen para describir la razón entre dos variables de escala.

Se pueden ordenar los resultados por los valores de una variable de agrupación, en orden ascendente o descendente. Se puede eliminar de los resultados el informe de los estadísticos de la razón y almacenar los resultados en un archivo externo.

Ejemplo. ¿Existe una buena uniformidad en la razón entre el precio de tasación y el precio de venta de viviendas en cada una de las cinco regiones? En los resultados, se puede descubrir que la distribución de las razones varía considerablemente entre regiones.

Estadísticos. Mediana, media, media ponderada, intervalos de confianza, coeficiente de dispersión (CDD), coeficiente de variación centrado en la mediana, coeficiente de variación centrado en la media, el diferencial de precio (DRV), desviación típica, desviación absoluta promedio (DAP), rango, valores mínimos y máximos y el índice de concentración calculado dentro de un rango o porcentaje (especificados por el usuario) respecto a la razón mediana.

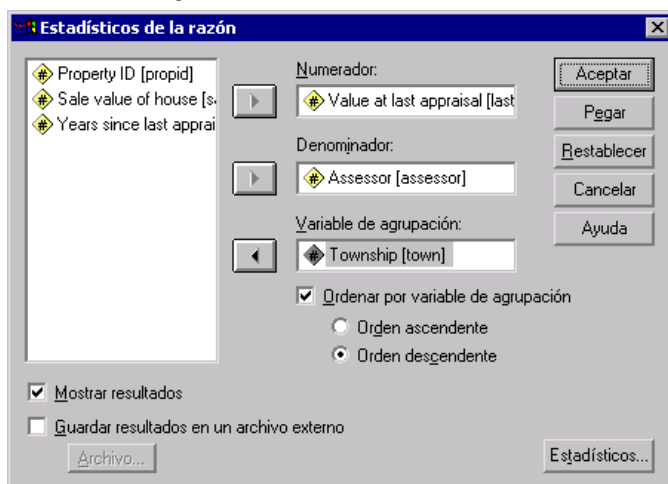
Datos. Utilice códigos numéricos o cadenas cortas para codificar las variables de agrupación (medidas de nivel nominal u ordinal).

Supuestos. Las variables que definen el numerador y el denominador de la razón deben ser variables de escala, que toman valores positivos.

Para obtener estadísticos de la razón

- ▶ Elija en los menús:
 - Analizar
 - Estadísticos descriptivos
 - Razón...

Figura 39-1
Cuadro de diálogo Estadísticos de la razón



- ▶ Seleccione una variable de numerador.
- ▶ Seleccione una variable de denominador.

Si lo desea, tiene la posibilidad de:

- Seleccionar una variable de agrupación y especificar el orden de los grupos en los resultados.
- Elegir si desea mostrar los resultados en el Visor de resultados.
- Seleccionar si desea guardar los resultados en un archivo externo para un uso posterior y especificar el nombre del archivo en el que se van a guardar los resultados.

Estadísticos de la razón

Figura 39-2
Cuadro de diálogo Estadísticos

Tendencia central. Las medidas de tendencia central son estadísticos que describen la distribución de las razones.

- **Mediana.** Un valor tal que el número de razones menores que este valor es igual al número de razones mayores que el mismo.
- **Media.** El resultado de sumar las razones y dividir la suma entre el número total de razones.
- **Media ponderada.** El resultado de dividir la media del numerador entre la media del denominador. También es la media de las razones ponderadas por el denominador.
- **Intervalos de confianza.** Muestra los intervalos de confianza para la media, la mediana y la media ponderada (si se solicita). Especifique un valor mayor o igual que 0 y menor que 100 como nivel de confianza.

Dispersión. Estos son estadísticos que miden cantidad de variación o de dispersión entre los valores observados.

- **DAP.** La desviación absoluta promedio es el resultado de sumar las desviaciones absolutas de las razones respecto a la mediana y dividir el resultado entre el número total de razones.
- **CDD.** El coeficiente de dispersión es el resultado de expresar la desviación absoluta promedio como un porcentaje de la mediana.
- **DRP.** El diferencial relativo al precio, también conocido como el índice de regresibilidad, es el resultado de dividir la media por la media ponderada.
- **CDV centrado en la mediana.** El coeficiente de variación centrado en la mediana es el resultado de expresar la raíz de la media cuadrática de las desviaciones respecto a la mediana como un porcentaje de la mediana.
- **CDV centrado en la media.** El coeficiente de variación centrado en la media es el resultado de expresar la desviación típica como un porcentaje de la media.
- **Desviación típica.** El resultado de sumar las desviaciones cuadráticas de las razones respecto a la media, dividir la suma por el número total de razones menos uno y extraer la raíz cuadrada positiva.
- **Rango.** El resultado de restar la razón mínima de la razón máxima.
- **Mínimo.** La razón menor.
- **Máximo.** La razón mayor.

Índice de concentración. El coeficiente de concentración mide el porcentaje de razones que están dentro de un intervalo. Se puede calcular de dos maneras:

- **Razones dentro del.** En este caso, el intervalo se define de forma explícita especificando los valores superior e inferior del intervalo. Introduzca valores para las proporciones superior e inferior y pulse en **Añadir** para obtener un intervalo.
- **Razones en.** En este caso, el intervalo se define de forma implícita al especificar el porcentaje de la mediana. Introduzca un valor entre 0 y 100 y pulse en **Añadir**. El límite inferior del intervalo será igual a $(1 - 0.01 \times \text{valor}) \times \text{mediana}$, y el límite superior será igual a $(1 + 0.01 \times \text{valor}) \times \text{mediana}$.

Conceptos básicos de la utilidad de gráficos

Se pueden crear gráficos de alta resolución mediante los procedimientos del menú Gráficos y mediante muchos de los procedimientos del menú Analizar. Este capítulo explica los conceptos básicos de la utilidad de gráficos.

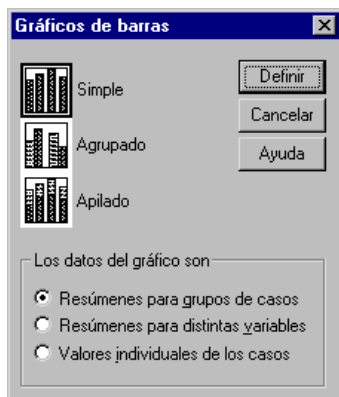
Creación y modificación de gráficos

Antes de crear un gráfico es necesario tener los datos en el Editor de datos. Es posible introducir los datos directamente en el Editor de datos; abrir un archivo de datos previamente guardado o leer una hoja de cálculo, un archivo de datos de texto delimitado por tabuladores o un archivo de base de datos. La selección del Tutorial del menú de ayuda ofrece ejemplos en pantalla de creación y modificación de gráficos. Además, el sistema de ayuda en pantalla incluye información sobre cómo crear y modificar cualquier tipo de gráfico.

Creación de gráficos

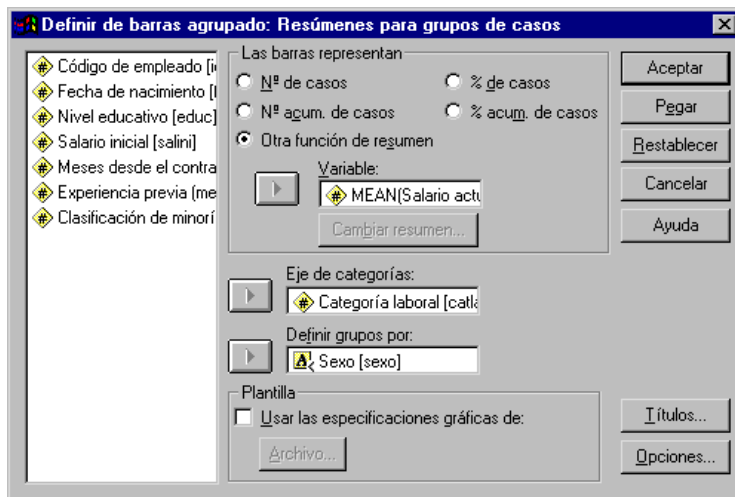
Una vez que los datos se encuentran en el Editor de datos, puede crear un gráfico seleccionando un tipo de gráfico en el menú Gráficos. Aparecerá el cuadro de diálogo del gráfico.

Figura 40-1
Cuadro de diálogo de gráficos



Este cuadro de diálogo contiene iconos para los diferentes tipos de gráficos y una lista de estructuras de datos. Pulse en Definir para abrir un cuadro de diálogo de definición de gráficos como el que se muestra a continuación.

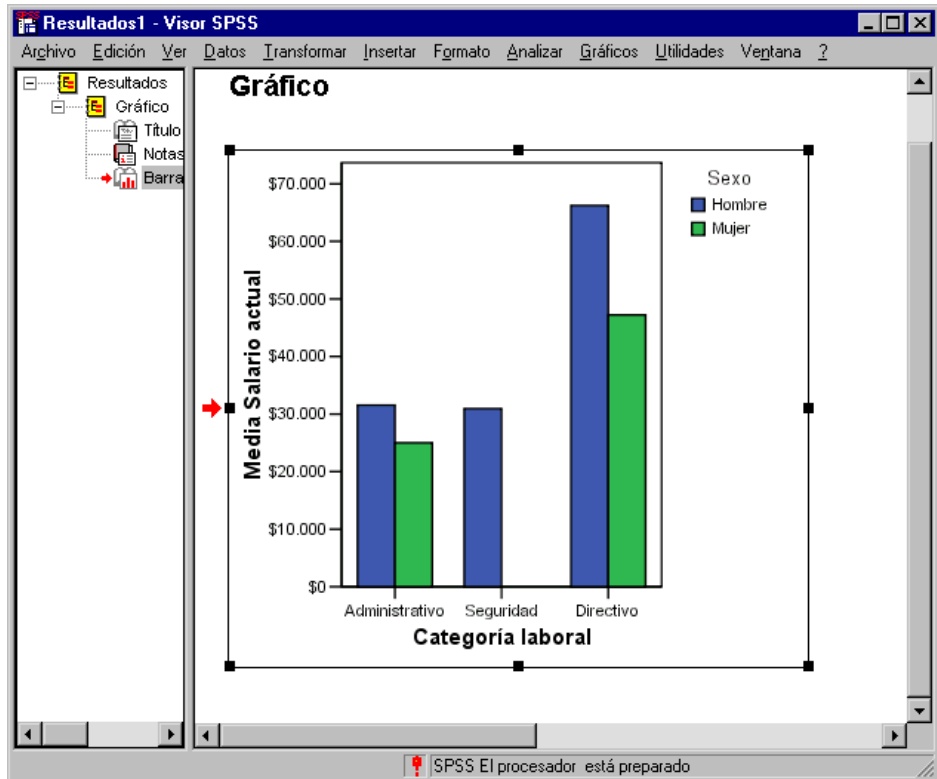
Figura 40-2
Cuadro de diálogo de definición del gráfico



En este cuadro de diálogo, puede seleccionar las variables apropiadas para el gráfico y elegir las opciones deseadas. Si desea obtener información acerca de las diferentes opciones, pulse en Ayuda.

Aparecerá el gráfico en el Visor.

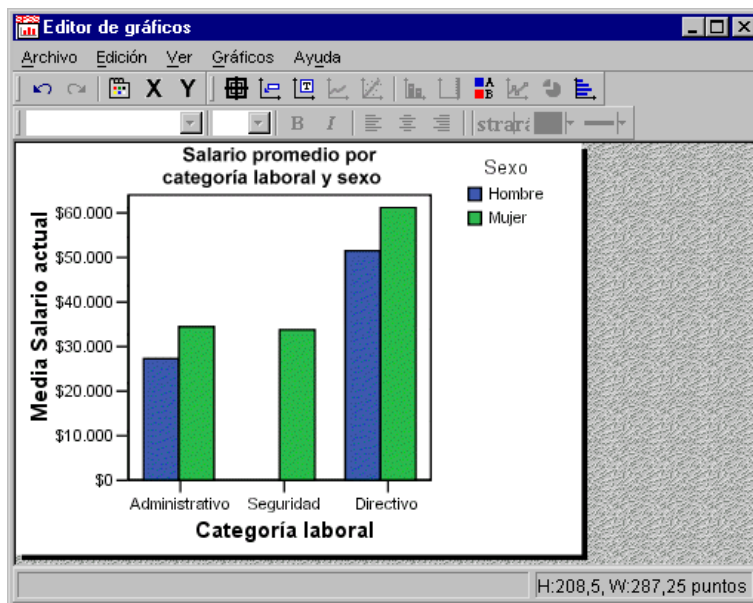
Figura 40-3
Gráfico en el Visor



Modificación de gráficos

Para modificar un gráfico, pulse dos veces en cualquier parte del gráfico que aparece en el Visor. Al hacer esto, aparecerá el gráfico en el Editor de gráficos.

Figura 40-4
Gráfico original en el Editor de gráficos



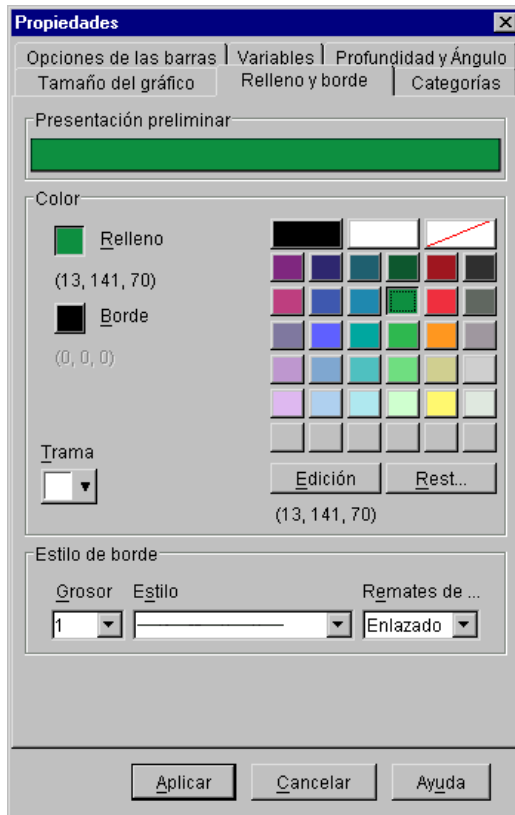
Puede modificar cualquier parte del gráfico o cambiar a otro tipo de gráfico representando los mismos datos. Puede añadir elementos, así como ocultarlos o mostrarlos utilizando los menús del Editor de gráficos.

Para modificar un elemento de un gráfico:

- ▶ Seleccione el elemento que desea modificar.
- ▶ Elija en los menús:
 - Edición
 - Propiedades...

Se abrirá la ventana Propiedades. Las pestañas que aparecen en la ventana Propiedades dependen de la selección realizada. La ayuda en pantalla explica cómo hacer que aparezcan las pestañas que necesita.

Figura 40-5
Ventana Propiedades

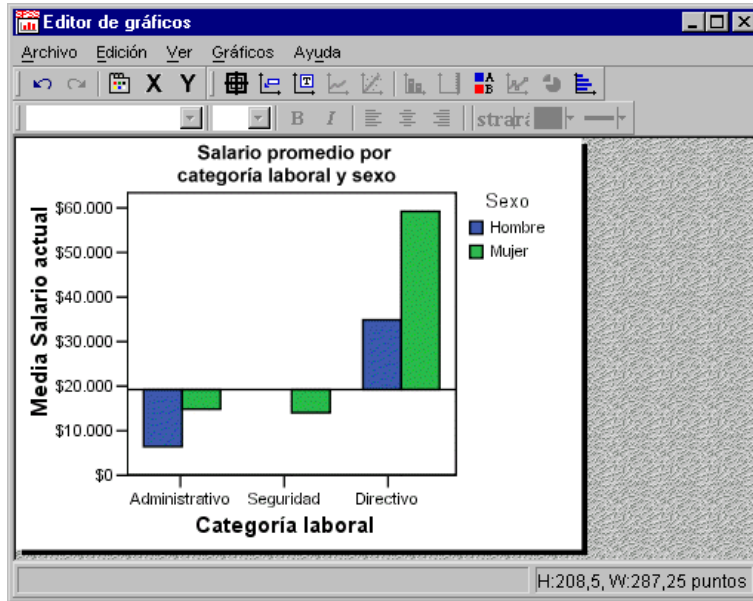


Entre las modificaciones habituales, se incluyen las siguientes:

- Editar el texto del gráfico.
- Cambiar el color y la trama de relleno de las barras.
- Añadir texto al gráfico, como por ejemplo un título o una anotación.
- Cambiar la posición de la línea de origen de las barras.
- Cambiar el borde del marco exterior de transparente a negro.

A continuación, se muestra un gráfico modificado.

Figura 40-6
Gráfico modificado

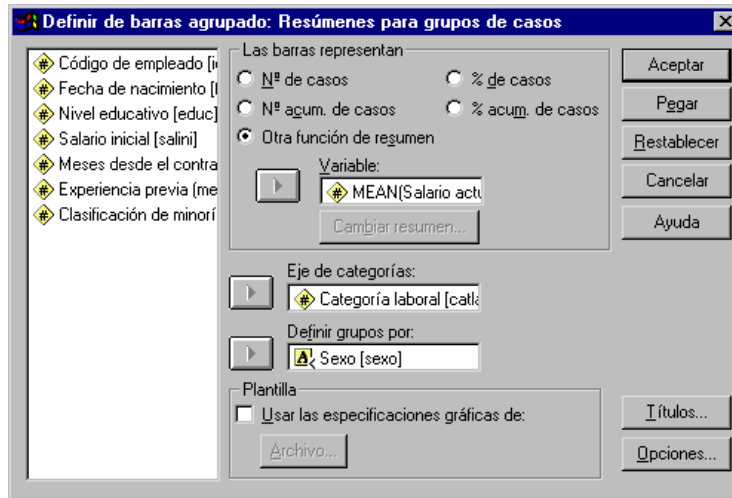


Las modificaciones realizadas al gráfico se guardan al cerrar la ventana del gráfico y el gráfico modificado se muestra en el Visor.

Opciones de definición de gráfico

Al definir un gráfico, el cuadro de diálogo de definición de dicho gráfico en particular normalmente contiene los botones Títulos y Opciones y un grupo Plantilla. Estas opciones globales se encuentran disponibles para la mayoría de los gráficos, independientemente del tipo concreto. No obstante, no están disponibles para gráficos P-P, gráficos Q-Q, gráficos de secuencia o gráficos de serie temporal.

Figura 40-7
Un cuadro de diálogo de definición del gráfico

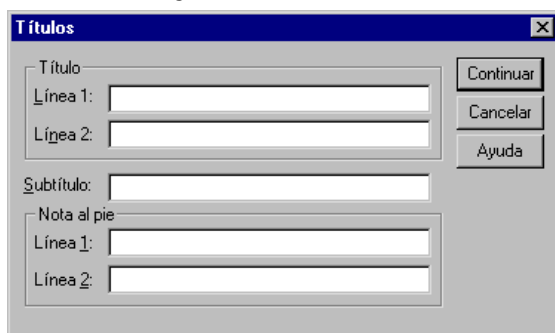


Pulse en **Títulos** para especificar los títulos, los subtítulos y las notas al pie. Puede pulsar en **Opciones** para controlar distintas opciones de gráfico como, por ejemplo, el tratamiento de los valores perdidos o la presentación de barras de error. Las opciones específicas disponibles dependen del tipo de gráfico. Además, puede aplicar una plantilla de atributos previamente seleccionados tanto al definir el gráfico como una vez que se ha creado el gráfico. Las siguientes secciones describen cómo definir estas características en el momento en que se define el gráfico.

Títulos, subtítulos y notas al pie

En cualquier gráfico, puede definir dos líneas de título, una línea de subtítulo y dos líneas de notas al pie como parte de la definición del gráfico original. Para especificar los títulos o las notas al pie al definir un gráfico, pulse en **Títulos** en el cuadro de diálogo de definición del gráfico. Se abrirá el cuadro de diálogo **Títulos**.

Figura 40-8
Cuadro de diálogo Títulos



Cada línea puede tener hasta 72 caracteres. El número de caracteres que realmente aparecerán en el gráfico dependerá de la fuente y el tamaño. La mayoría de los títulos se justifican a la izquierda por defecto y, si son demasiado largos, se recortan a la derecha. Los títulos de los gráficos de sectores, por defecto, se centran y, si son demasiado largos, se recortan por ambos extremos.

Los títulos, los subtítulos y las notas al pie se representan como cuadros de texto en el Editor de gráficos. Puede añadir, eliminar o revisar los cuadros de texto en el Editor de gráficos, así como cambiar su fuente, tamaño y justificación.

Opciones

El cuadro de diálogo Opciones ofrece una serie de opciones para el gráfico que va a crear. Para acceder a este cuadro de diálogo, pulse en Opciones, Opciones categóricas u Opciones de escala en el cuadro de diálogo de definición del gráfico.

La disponibilidad de cada opción depende de la selección que se haya hecho anteriormente o del tipo de gráfico. Las opciones de valores perdidos no están disponibles para los gráficos que utilizan valores de casos individuales ni para los histogramas. La opción de presentación de etiquetas de caso sólo está disponible cuando se trata de un diagrama de dispersión en el que se haya seleccionado una variable para las etiquetas de caso. Las barras de error muestran las opciones disponibles sólo para los gráficos de categorías que muestran medias, medianas, porcentajes o recuentos. La opción de presentación de curva normal y las opciones de intervalo están disponibles sólo para las pirámides de población que muestran la

distribución de una variable de escala. La opción de presentación de la forma del gráfico está disponible sólo para los gráficos de puntos.

Valores perdidos

Si ha seleccionado resúmenes de distintas variables para un gráfico de categorías o si quiere crear un diagrama de dispersión, puede optar por una de las siguientes posibilidades para excluir los casos que tienen valores perdidos:

- **Excluir casos según lista.** Si alguna de las variables del gráfico tiene un valor perdido para un determinado caso, se excluirá el caso completo del gráfico.
- **Excluir casos variable por variable.** Si una variable seleccionada tiene algún valor perdido, los casos que tengan estos valores perdidos se excluirán al analizar dicha variable.

Para ver la diferencia existente entre la exclusión por lista y la exclusión por variables de los valores perdidos, observe las siguientes figuras, que muestran un gráfico de barras para cada una de las dos opciones.

Figura 40-9
Exclusión por lista de los valores perdidos

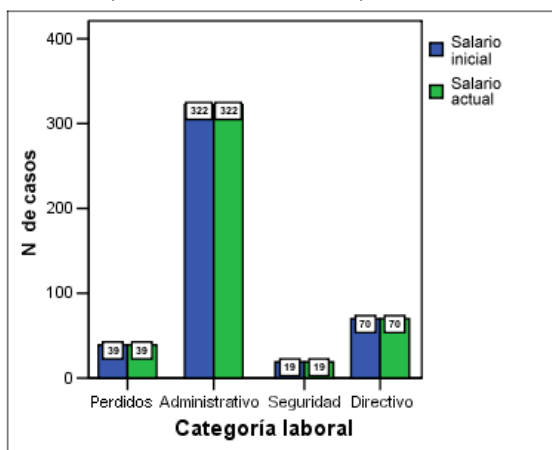
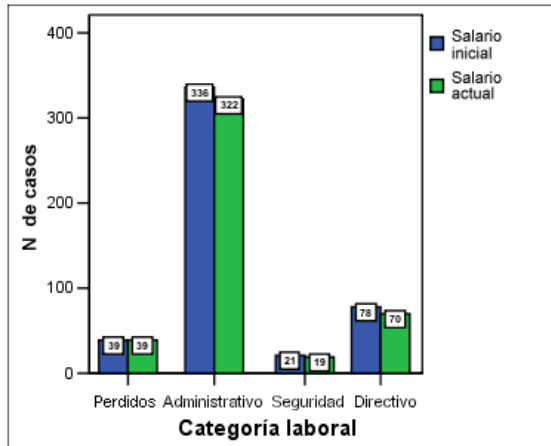


Figura 40-10
Exclusión por variables de los valores perdidos



Estos gráficos se crearon a partir de una versión del archivo *Employee data.sav* que se editó para que tuviese algunos valores perdidos por el sistema (en blanco) para las variables correspondientes al salario actual y a la categoría laboral. En otros casos, se introdujo el valor 0 y se definió como perdido. En ambos gráficos, se ha seleccionado la opción *Mostrar los grupos definidos por los valores perdidos*, que añade la categoría *Perdidos* al resto de las categorías laborales mostradas. En cada gráfico, los valores de la función de resumen, *Número de casos*, se muestran en las etiquetas de las barras.

En ambos gráficos, 26 casos tienen un valor perdido del sistema para la categoría laboral y 13 casos tienen el valor perdido definido por el usuario (0). En el gráfico por lista, el número de casos es el mismo para ambas variables en cada agrupación de barras ya que siempre que había un valor perdido se excluyó el caso de todas las variables. En el gráfico por variables, el número de casos no perdidos de cada variable de una categoría se representa sin tener en cuenta los valores perdidos en las otras variables.

La siguiente opción también está disponible para los valores perdidos:

- Mostrar los grupos definidos por los valores perdidos.** Si hay valores perdidos en los datos de las variables utilizadas para definir las categorías o los subgrupos, los valores perdidos definidos por el usuario (los valores identificados como perdidos por el usuario) y los valores perdidos por el sistema se incluyen juntos en una categoría con la etiqueta *Perdidos*. La categoría “perdidos” aparece en el eje

de categorías o en la leyenda, añadiendo, por ejemplo, una barra adicional, un sector a un gráfico de sectores o una caja adicional a un diagrama de caja. En un diagrama de dispersión, los valores perdidos añaden una categoría “perdidos” al conjunto de marcadores. Si no hay ningún valor perdido, no aparecerá la categoría “perdidos”.

Si selecciona esta opción y quiere suprimir la presentación después de que se dibuje el gráfico, seleccione el gráfico y, a continuación, seleccione Propiedades en el menú Edición. +Utilice la pestaña Categorías para mover las categorías que desea suprimir a la lista Excluidos.

Esta opción no está disponible para un diagrama de dispersión superpuesto ni para gráficos de una única serie en los que los datos se resuman por distintas variables.

Etiquetas de caso

Otra opción controla el estado de las etiquetas de caso si se muestra un diagrama de dispersión en primer lugar.

- **Mostrar el gráfico con las etiquetas de caso.** Si se selecciona esta opción, se mostrarán todas las etiquetas de caso al crear un diagrama de dispersión. Por defecto, esta opción no está seleccionada, es decir, el diagrama de dispersión por defecto se muestra sin etiquetas. Si selecciona esta opción, es posible que las etiquetas de caso se solapen.

Barras de error

Si va a crear un gráfico de categorías que muestre medias, medianas, recuentos o porcentajes, hay otra opción disponible:

- **Mostrar las barras de error.** Esta opción controla la presentación de las barras de error. Para las medias, puede optar por mostrar intervalos de confianza en torno a las medias de categorías, más y menos n veces la desviación típica de la variable o más y menos n veces el típico de la media. Para las medianas, los recuentos y los porcentajes, sólo hay disponibles intervalos de confianza.

Todos los intervalos de confianza son intervalos individuales, con niveles de cobertura (basados en el alfa especificado) que se aplican a categorías individuales y no al conjunto de todas las categorías del gráfico simultáneamente. Para las frecuencias y porcentajes, los intervalos se basan en intervalos binomiales para las proporciones de

cada categoría del total de todas las categorías, con la escala ajustada debidamente para el recuento o la métrica de porcentaje. Observe que los intervalos de confianza para los recuentos, las medianas y los porcentajes suelen ser no simétricos en torno al estadístico.

Curva normal y opciones de intervalo

Si va a crear una pirámide de población y Mostrar la distribución de es una variable de escala, puede optar por mostrar una curva normal o por cambiar el modo en que se agrupan los casos en el gráfico:

- **Mostrar la curva normal.** Superponga a cada mitad de la pirámide de población una curva normal con la misma media y desviación típica que las de los datos.
- **Anclar el primer intervalo.** Especifique el valor de inicio del primer intervalo. Este número debe ser igual o inferior al valor mínimo del conjunto de datos. Por defecto, el primer intervalo incluye el valor de los datos que sea el mínimo. El anclaje se define de modo que los límites del intervalo tengan los valores adecuados. Puede cambiar el valor por defecto para que el primer intervalo incluya valores no presentes en el conjunto de datos. Por ejemplo, puede que desee incluir los valores de 0 a 5 incluso aunque el valor inferior del conjunto de datos sea 6.
- **Tamaño de los intervalos.** Cambiar el tamaño de los intervalos. Se puede especificar el número de intervalos o el ancho de cada intervalo. La anchura también afecta al número de intervalos. Por ejemplo, si el rango del eje es de 0 a 100 y especifica una anchura de 5, habrá 20 intervalos. Cuanto mayor sea el número de intervalos, más preciso será el histograma de cada mitad de la pirámide de población. Sin embargo, un mayor número de intervalos hará que el histograma sea demasiado detallado, ya que el propósito del histograma consiste en resumir los datos. Debe seleccionar un valor que devuelva un histograma que resuma con precisión los datos.

Forma del gráfico

Si va a crear un gráfico de puntos, puede cambiar la forma de dicho gráfico de puntos.

Plantillas gráficas

Puede aplicar muchos de los atributos y elementos de texto de un gráfico a otro. Esto le permite modificar un gráfico, guardar dicho gráfico como una plantilla y utilizar a continuación la plantilla para crear una serie de gráficos similares.

Para utilizar una plantilla al crear un gráfico, seleccione Usar las especificaciones gráficas de (en el grupo Plantilla del cuadro de diálogo de definición del gráfico) y pulse en Archivo. Se abrirá un cuadro de diálogo estándar de selección de archivos.

Para aplicar una plantilla a un gráfico que ya se encuentra en el Editor de gráficos, seleccione las siguientes opciones en los menús:

Archivo

Aplicar plantilla gráfica...

Se abrirá un cuadro de diálogo estándar de selección de archivos. Seleccione el archivo que desea utilizar como plantilla. Si está creando un nuevo gráfico, el nombre del archivo seleccionado aparecerá en el grupo Plantilla cuando regrese al cuadro de diálogo de definición del gráfico.

Las plantillas se utilizan para tomar el formato de un gráfico y aplicarlo al nuevo gráfico que se está generando. En general, se aplicará automáticamente toda la información acerca del formato del antiguo gráfico que sea posible aplicar al nuevo formato. Por ejemplo, si el antiguo gráfico es un gráfico de barras agrupadas en el que los colores de las barras son el amarillo y el verde, si el nuevo gráfico es un gráfico de líneas múltiples, las líneas serán amarillas y verdes. Si el antiguo gráfico es un gráfico de barras simple con sombras verticales y el nuevo gráfico es un gráfico de líneas simple, las líneas no tendrán sombras verticales, ya que las sombras verticales no se aplican a los gráficos de líneas. Si hay títulos en el gráfico de plantilla pero no los hay en el nuevo gráfico, se tomarán dichos títulos del gráfico de plantilla. Si hay títulos definidos en el nuevo gráfico, se utilizarán en lugar de los títulos del gráfico de plantilla.

Para crear una plantilla gráfica

- ▶ Cree un gráfico.
- ▶ Edite el gráfico para que contenga los atributos que desea tener en la plantilla.

- ▶ En los menús del Editor de gráficos, seleccione:
 - Archivo
 - Guardar plantilla gráfica...
- ▶ En el cuadro de diálogo Guardar plantilla gráfica, especifique las características del gráfico que desea guardar en la plantilla. La ayuda en pantalla describe con detalle las opciones disponibles.
- ▶ Pulse en Continuar.
- ▶ Escriba el nombre del archivo y la ubicación de la nueva plantilla. La extensión de la plantilla es *.sgt*.

Curvas COR

Este procedimiento es un método útil para evaluar la realización de esquemas de clasificación en los que exista una variable con dos categorías por las que se clasifiquen los sujetos.

Ejemplo. Un banco tiene interés en clasificar a sus clientes dependiendo de si se retrasarán o no en el pago de sus préstamos; por tanto, se desarrollan métodos especiales para tomar estas decisiones. Las curvas COR se pueden utilizar para evaluar el grado de acierto de estos métodos.

Estadísticos. Es un área situada bajo la curva COR con un intervalo de confianza y puntos de coordenadas de la curva COR. Gráficos: Curva COR.

Métodos. Se puede calcular la estimación del área situado bajo la curva COR de forma paramétrica o no paramétrica mediante un modelo exponencial binegativo.

Datos. Las variables de contraste son cuantitativas. Suele estar constituida por probabilidades, resultantes de un análisis discriminante o de una regresión logística, o por puntuaciones atribuidas en una escala arbitraria que indican el «grado de convicción» que tiene un evaluador de que el sujeto pueda pertenecer a una u otra categoría. La variable de estado puede ser de cualquier tipo e indicar la categoría real a la que pertenece un sujeto. El valor de la variable de estado indica la categoría que se debe considerar *positiva*.

Supuestos. Se considera que los números ascendentes de la escala del evaluador representan la creciente convicción de que el sujeto pertenece a una categoría. Por el contrario, los números descendentes representan la creciente convicción de que el sujeto pertenece a la otra categoría. El usuario deberá elegir qué dirección es *positiva*. También se considera que se conoce la categoría *real* a la que pertenece el sujeto.

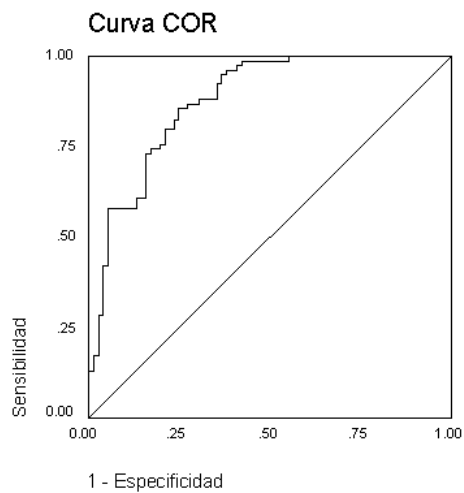
Figura 41-1
Resultados de curvas COR

**Resumen por proceso
de casos**

REAL	N válido (por lista)
Positivo ¹	74
Negativo	76

Un resultado mayor en las variables resultantes del contraste indica una mayor evidencia de que el estado real sea positivo.

1. El estado positivo real es 1.00



Zona bajo la curva

Variable(s) resultantes del contraste: PROBS

Zona	Error típico ¹	Sig. asintótica ²	95% de confianza para los intervalos de la asintótica	
			Límite inferior	Límite superior
.877	.028	.000	.823	.931

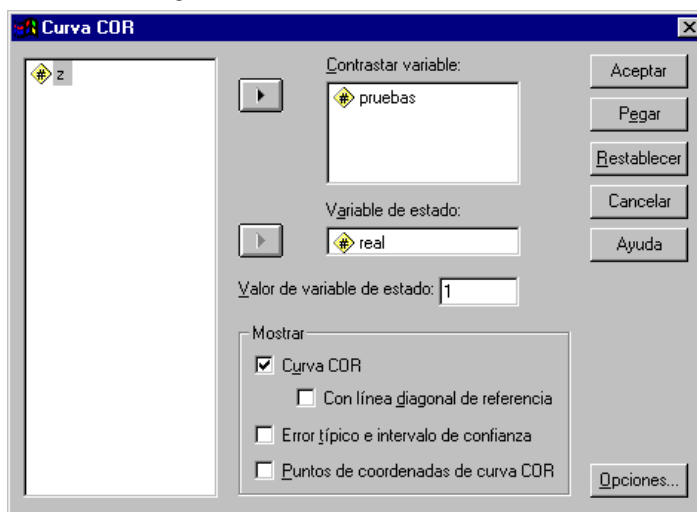
1. Bajo el supuesto no paramétrico

2. Hipótesis nula: zona real = 0.5

Para obtener una curva COR

- ▶ Elija en los menús:
Gráficos
Curva COR...

Figura 41-2
Cuadro de diálogo Curva COR



- ▶ Seleccione una o más variables de probabilidad de contraste.

- ▶ Elija una variable de estado.
- ▶ Identifique el valor *positivo* para la variable de estado.

Curvas COR: Opciones

Figura 41-3
Cuadro de diálogo Curva COR: Opciones

Curva COR: Opciones

Clasificación

Incluir valor de corte para clasificación positiva

Excluir valor de corte para clasificación positiva

Dirección de la prueba

Un resultado mayor en la prueba indica una prueba más positiva

Un resultado menor en la prueba indica una prueba más positiva

Parámetros para el error típico del área

Supuestos distribucionales: No paramétrica

Nivel de confianza: 95 %

Valores perdidos

Excluir tanto valores perdidos por el usuario como valores perdidos por el sistema

Los valores perdidos por el usuario serán considerados válidos

Continuar

Cancelar

Ayuda

Puede seleccionar las opciones siguientes para su análisis:

Clasificación. Permite especificar si se debe incluir o excluir el valor del punto de corte al realizar una clasificación *positiva*. Esta opción no afecta a los resultados.

Dirección de la prueba. Permite especificar la dirección de la escala según la categoría *positiva*.

Parámetros para el error típico del área. Permite especificar el método de estimación del error típico del área situada bajo la curva. Los métodos disponibles son el no paramétrico y el exponencial binnegativo. También se puede establecer el nivel para el intervalo de confianza. El rango disponible es entre el 50,1% y el 99,9%.

Valores perdidos. Permite especificar el tratamiento que reciben los valores perdidos.

Utilidades

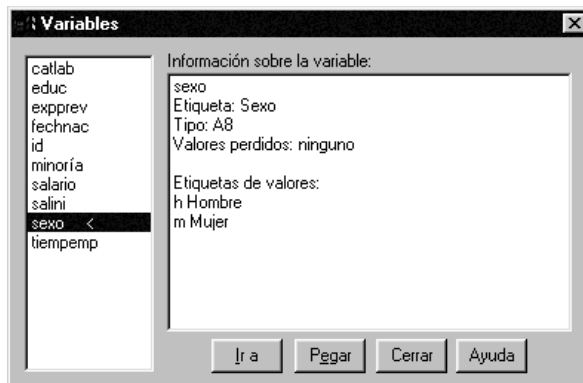
En este capítulo se describen las funciones del menú Utilidades así como las opciones para posibilidad de reordenar las listas de variables de destino usando los menús de sistema de Windows.

Información sobre la variable

El cuadro de diálogo Variables muestra información sobre la definición de la variable seleccionada actualmente, incluyendo:

- Formato de datos
- Etiqueta de variable
- Valores definidos por el usuario como perdidos
- Etiquetas de valor

Figura 42-1
Cuadro de diálogo Variables



Ir a. Se dirige a la variable seleccionada en el Editor de datos.

Pegar. Pega las variables seleccionadas en la posición del cursor en la ventana de sintaxis designada.

Para modificar la definición de una variable, utilice la Vista de variables en el Editor de datos.

Para obtener información sobre la variable

- ▶ Elija en los menús:
Utilidades
Variables...
- ▶ Seleccione la variable cuya información de definición desee mostrar.

Comentarios del archivo de datos

Puede incluir comentarios descriptivos en un archivo de datos. Para los archivos de datos con formato SPSS, estos comentarios se guardan con el archivo de datos.

Para añadir, modificar, eliminar o visualizar los comentarios del archivo de datos:

- ▶ Elija en los menús:
Utilidades
Comentarios del archivo de datos
- ▶ Para mostrar los comentarios en el Visor, seleccione Mostrar comentarios en resultados.

Los comentarios admiten cualquier longitud, aunque están limitados a 80 bytes (por regla general, 80 caracteres en idiomas de un solo byte) por línea; las líneas se dividen automáticamente en 80 caracteres. Los comentarios se muestran en la misma fuente que los resultados de texto para reflejar de forma precisa el modo en que aparecen al mostrarlos en el Visor.

Se añade de forma automática una anotación de fecha (la fecha actual entre paréntesis) al final de la lista de comentarios siempre que se añaden o modifican los comentarios. Esto puede dar lugar a cierta ambigüedad por lo que respecta a las fechas asociadas a los comentarios si modifica un comentario existente o introduce un comentario nuevo entre los comentarios existentes.

Conjuntos de variables

Puede restringir las variables que aparecen en las listas de variables de origen de un cuadro de diálogo, definiendo y utilizando los conjuntos de variables. Es especialmente útil en archivos de datos con un amplio número de variables. Los conjuntos de variables pequeños hacen que la búsqueda y la selección de variables para los análisis sea más fácil y pueden incluso mejorar el rendimiento. Si el archivo de datos contiene un elevado número de variables y cuadros de diálogo que se abren con lentitud, restringir las listas de origen de un cuadro de diálogo a subconjuntos de variables más pequeños reduciría la cantidad de tiempo empleado en abrirlos.

Definir conjuntos de variables

El cuadro de diálogo Definir conjuntos de variables crea subconjuntos de variables que se muestran en las listas de origen de los cuadros de diálogo.

Figura 42-2

Cuadro de diálogo Definir conjuntos de variables



Nombre del conjunto. Los nombres de los conjuntos pueden tener hasta 12 caracteres. Puede utilizar cualquier carácter, incluso los espacios en blanco. Los nombres de los conjuntos no distinguen mayúsculas de minúsculas.

Variables del conjunto. El conjunto puede estar compuesto de cualquier combinación de variables numéricas, de cadena corta y cadena larga. El orden de las variables del conjunto no tiene ningún efecto en el orden de presentación de las variables en las listas de origen del cuadro de diálogo. Una variable puede pertenecer a varios conjuntos.

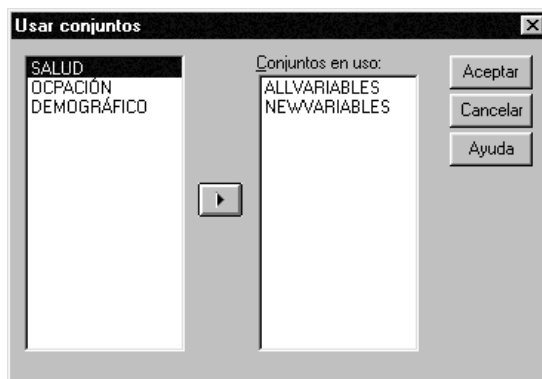
Para definir conjuntos de variables

- ▶ Elija en los menús:
Utilidades
Definir conjuntos...
- ▶ Seleccione las variables que desee incluir en el conjunto.
- ▶ Introduzca un nombre para el conjunto (hasta 12 caracteres).
- ▶ Pulse en Añadir conjunto.

Usar conjuntos

El cuadro de diálogo Usar conjuntos restringe las variables mostradas en las listas de origen de los cuadros de diálogo a los conjuntos seleccionados que haya definido.

Figura 42-3
Cuadro de diálogo Usar conjuntos



Conjuntos en uso. Muestra los conjuntos empleados en la creación de listas de variables de origen en los cuadros de diálogo. Las variables aparecen ordenadas en las listas de origen por nombre o por archivo. El orden de conjuntos y de variables en un conjunto no tiene ningún efecto en el orden de las variables de la lista de origen. Por defecto, se utilizan dos conjuntos definidos por el sistema:

ALLVARIABLES. Este conjunto contiene todas las variables del archivo de datos, incluidas las nuevas variables creadas durante una sesión.

NEWVARIABLES. Este conjunto contiene sólo las nuevas variables creadas durante la sesión.

Puede eliminar estos conjuntos de la lista y seleccionar otros, si bien debe haber al menos un conjunto en la lista. Si no elimina el conjunto *ALLVARIABLES* de la lista Conjuntos en uso, cualquier otro conjunto que incluya no tendrá efecto alguno.

Para restringir las listas de origen de los cuadros de diálogo a los conjuntos de variables definidos

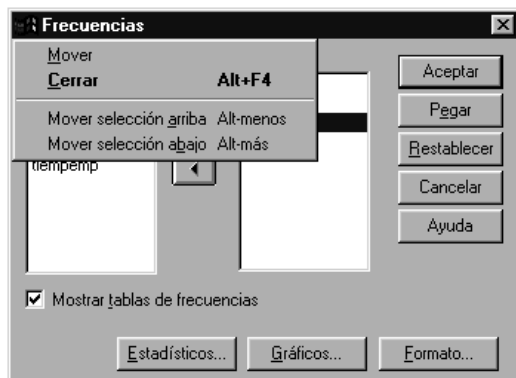
- ▶ Elija en los menús:
 - Utilidades
 - Usar conjuntos...
- ▶ Seleccione los conjuntos de variables definidos que contengan las variables que desee incluir en las listas de origen de los cuadros de diálogo.

Reordenación de listas de variables de destino

Las variables aparecen en las listas de destino de un cuadro de diálogo en el orden en que son seleccionadas en la lista de origen. Si desea modificar el orden de las variables de una lista de destino (pero no anular toda la selección ni volver a seleccionarlás en el nuevo orden), puede moverlas hacia arriba o hacia abajo en la lista de destino mediante el menú de sistema. situado en la esquina superior izquierda del cuadro de diálogo. Para acceder al mismo, pulse en la parte izquierda de la barra de título de dicho cuadro de diálogo.

Figura 42-4

Para acceder al mismo, pulse en la esquina superior izquierda del cuadro de diálogo).
Menú de sistema de Windows con reordenación de la lista de destino



Mover selección arriba. Mover selección arriba.

Mover selección abajo. Mueve las variables seleccionadas una posición hacia abajo en la lista de destino.

Puede mover múltiples variables simultáneamente si son contiguas (es decir, si están agrupadas unas junto a otras). No es posible mover grupos de variables no contiguas.

Opciones

Controla una gran variedad de opciones, entre ellas:

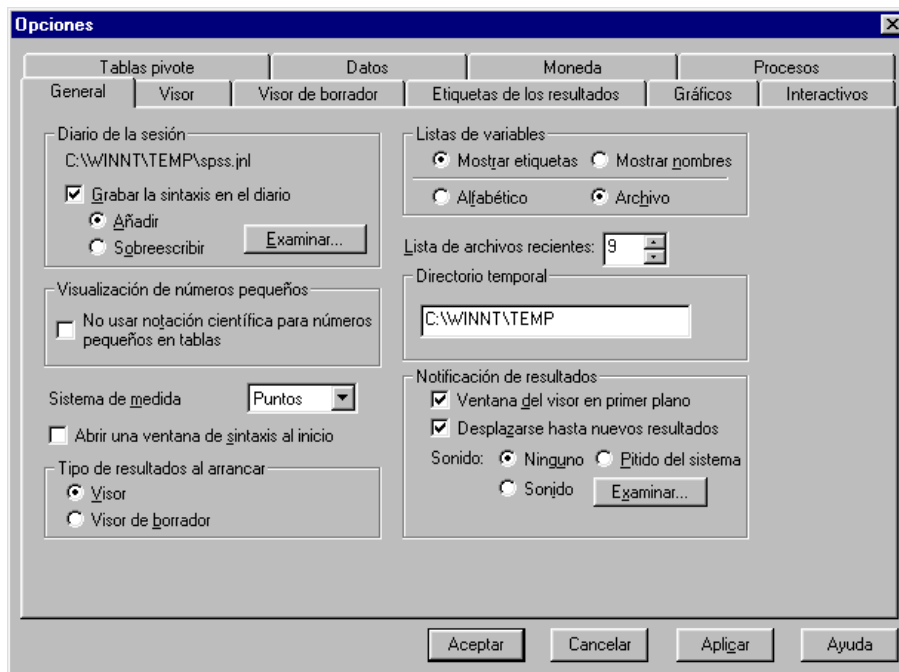
- El diario de la sesión, el cual guarda un registro de todos los comandos ejecutados en cada sesión.
- El orden en que aparecen las variables en las listas de origen de los cuadros de diálogo.
- Los elementos mostrados y ocultos en los nuevos resultados.
- Aspecto de tabla para las nuevas tablas y Aspecto de gráfico para los nuevos gráficos interactivos.
- Formatos de moneda personalizados.
- Archivos de autoproceso y funciones de autoproceso para personalizar los resultados.

Para modificar la configuración de las opciones

- ▶ Elija en los menús:
 - Edición
 - Opciones...
- ▶ Pulse en las pestañas de las selecciones que desee cambiar.
- ▶ Cambie las selecciones.
- ▶ Pulse en Aceptar o Aplicar.

General: Opciones

Figura 43-1
Pestaña general del cuadro de diálogo Opciones



Listas de variables. Controla la presentación de las variables en los cuadros de lista de los cuadros de diálogo. Es posible mostrar los nombres o las etiquetas de las variables. Estos nombres o etiquetas pueden presentarse por orden alfabético o por orden de archivo, que es el orden en el que figuran en el archivo de datos (y en el que se muestran en la ventana Editor de datos). El orden de presentación afecta sólo a las listas de variables de origen. Las listas de variables de destino siempre reflejan el orden en el que las variables han sido seleccionadas.

Diario de la sesión. Archivo de diario que registra todos los comandos ejecutados en una sesión. Incluye comandos introducidos y ejecutados en ventanas de sintaxis y comandos generados por elecciones de cuadros de diálogo. Puede editar este archivo y volver a utilizar los comandos en otras sesiones. Puede activar o desactivar el registro de sesión, añadir o sustituir el archivo diario y seleccionar el nombre y la ubicación del

mismo. También es posible copiar la sintaxis de comandos de este archivo y guardarla en un archivo de sintaxis para utilizarla con la unidad de producción automatizada.

Directorio temporal. Controla la ubicación de los archivos temporales creados durante una sesión. En el análisis en modo distribuido (disponible con la versión de servidor), esto no afecta a la ubicación de los archivos de datos temporales. En el modo distribuido, la ubicación de los archivos de datos temporales está controlada por la variable de entorno *SPSSTMPDIR*, la cual se puede definir sólo en el ordenador que ejecuta la versión del servidor del software. Si necesita cambiar la ubicación del directorio temporal, póngase en contacto con el administrador del sistema.

Lista de archivos recientes. Controla el número de archivos utilizados recientemente que aparecen en el menú Archivo.

Abrir una ventana de sintaxis al inicio. Las ventanas de sintaxis son ventanas de archivos de texto que sirven para introducir, editar y ejecutar comandos. Si utiliza con frecuencia la sintaxis de comandos, seleccione esta opción para abrir automáticamente una ventana de sintaxis al principio de cada sesión. Esta opción es especialmente útil para los usuarios avanzados que prefieren trabajar con la sintaxis de comandos en vez de con los cuadros de diálogo. Esta opción no se encuentra disponible en la versión para estudiantes.

No usar notación científica para números pequeños en tablas. Suprime la presentación de la notación científica para valores decimales pequeños en el resultado. Los valores decimales muy pequeños se muestran como 0 (o 0,000).

Tipo de visor al inicio. Determina el tipo de Visor que debe utilizarse y el formato de los resultados. El Visor genera tablas pivote interactivas y gráficos interactivos. El Visor de borrador convierte las tablas pivote en resultados de texto, y los gráficos en metarchivos.

Sistema de medida. Sistema de medida utilizado (puntos, pulgadas o centímetros) para especificar atributos tales como los márgenes de casillas de las tablas pivote, los anchos de casilla y el espacio entre las tablas para la impresión.

Idioma. Determina el idioma que se utiliza para los resultados. No se aplica a resultados en formato texto, gráficos interactivos ni mapas (disponibles con la opción Mapas). Esta lista de idiomas disponibles depende de los archivos de idioma que estén instalados actualmente. Al instalar SPSS, se instala también automáticamente un número de idiomas. Para acceder a archivos de idioma adicionales, vaya a <http://www.spss.com/tech/downloads/base.htm>.

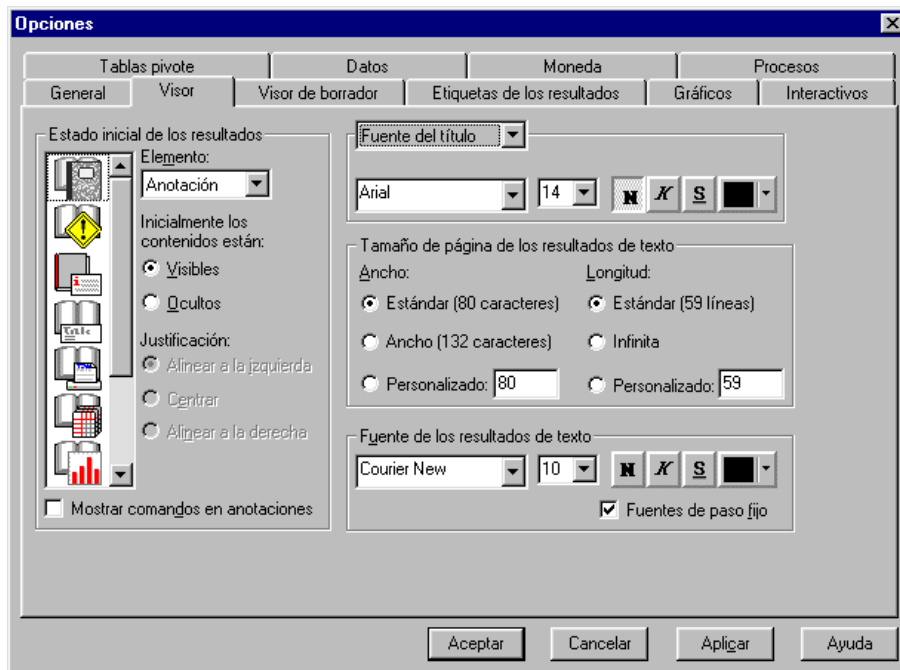
Nota: Es posible que, en los procesos personalizados que se basan en cadenas de texto de un idioma específico, estas cadenas no se ejecuten correctamente cuando se cambie el idioma de los resultados. Si desea obtener más información, consulte “Procesos: Opciones” en p. 633.

Notificación. Determina cómo debe notificar el programa al usuario que ha finalizado la ejecución de un procedimiento y que los resultados están disponibles en el Visor.

Opciones del Visor

Las opciones de salida de presentación de resultados en el Visor sólo afectan a los resultados obtenidos tras el cambio de la configuración. A los resultados que ya se muestran en el Visor no les afectan los cambios en la configuración.

Figura 43-2
Pestaña Visor del cuadro de diálogo Opciones



Estado inicial de los resultados. Controla los elementos que se muestran y se ocultan automáticamente cada vez que se ejecuta un procedimiento, además de la alineación inicial de los elementos. Puede controlar la presentación de los siguientes elementos: registro, advertencias, notas, títulos, tablas pivote, gráficos y resultados de texto (los resultados no se muestran en las tablas pivote). También se puede activar o desactivar la muestra de los comandos en el registro. Se pueden copiar comandos de sintaxis del registro y guardarlos en un archivo de sintaxis para utilizarlos con la utilidad de producción automática.

Nota: Todos los elementos de resultados aparecen alineados a la izquierda en el Visor. Las selecciones de justificación sólo afectarán a la alineación de los resultados impresos. Los elementos con alineación centrada y a la derecha se identifican por un pequeño símbolo situado encima y a la izquierda del elemento.

Fuente del título. Controla el estilo, el tamaño y el color de la fuente de los nuevos títulos de resultados.

Fuente del título de página. Controla el estilo, el tamaño y el color de la fuente de los nuevos títulos de página y de los títulos de página generados mediante la sintaxis de comandos `TITLE` y `SUBTITLE` o creados mediante Nuevo título de página en el menú Insertar.

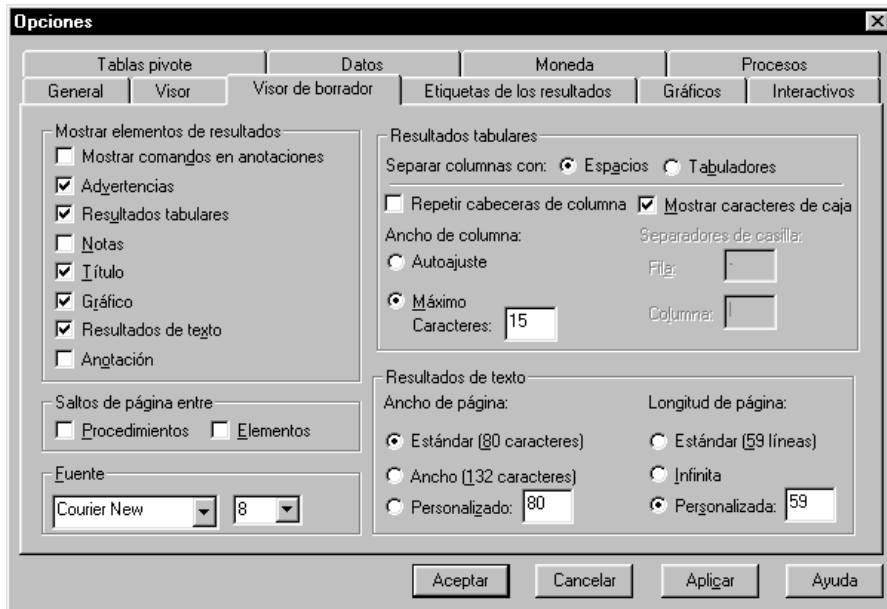
Tamaño de página de los resultados de texto. En los resultados de texto, controla el ancho de página (expresado en número de caracteres) y el largo de página (expresado en número de líneas). En algunos procedimientos, algunos estadísticos se muestran sólo en formato ancho.

Fuentes de los resultados de texto. Fuente utilizada para los resultados de texto. Los resultados de texto se han diseñado para utilizarlos con fuentes de paso fijo. Si selecciona una fuente no monoespaciada, los resultados tabulares no se alinearán adecuadamente.

Opciones del Visor de borrador

Las opciones de presentación de resultados del Visor de borrador sólo afectan a los resultados producidos tras el cambio de las especificaciones. Los cambios en la configuración no afectan al resultado mostrado en el Visor de borrador.

Figura 43-3
Pestaña Visor de borrador del cuadro diálogo Opciones



Mostrar elementos de resultados. Determina los elementos que deben mostrarse automáticamente cada vez que se ejecuta un procedimiento. Puede controlar la presentación de los siguientes elementos: el registro, las advertencias, las notas, los títulos, los resultados tabulados y los resultados de texto (resultados separados por espacios). También se puede activar o desactivar la presentación de los comandos en el registro. Se pueden copiar comandos de sintaxis del registro y guardarlos en un archivo de sintaxis para utilizarlos con la utilidad de producción automática.

Saltos de página entre. Inserta saltos de página entre los resultados de diferentes procedimientos y entre elementos de resultados individuales.

Fuente. Fuente utilizada en los nuevos resultados. Sólo se encuentran disponibles las fuentes de paso fijo (monoespaciadas) porque los resultados de texto separados por espacios no pueden alinearse correctamente si se emplea una fuente proporcional.

Resultados tabulares. Controla las opciones de los resultados de las tablas pivote convertidos a resultados tabulares. Las especificaciones de ancho de columna y separadores de columna sólo se encuentran disponibles si se selecciona Espacios

como separador de columnas. En los resultados tabulares separados por espacios se eliminan por defecto todas las líneas de separación y el ancho de cada columna queda determinado por la etiqueta o al valor más largo contenido en la columna. Para limitar el ancho de las columnas y ajustar las etiquetas largas, especifique un número de caracteres para el ancho de la columna.

Nota: Sin embargo, en el Visor de borrador, los resultados separados por tabuladores no se alinearán correctamente. Este formato resulta de utilidad para copiar y pegar resultados en aplicaciones de procesamiento de texto en las que se pueda utilizar cualquier fuente (y no sólo las de paso fijo) y se puedan establecer tabuladores para alinear correctamente el resultado.

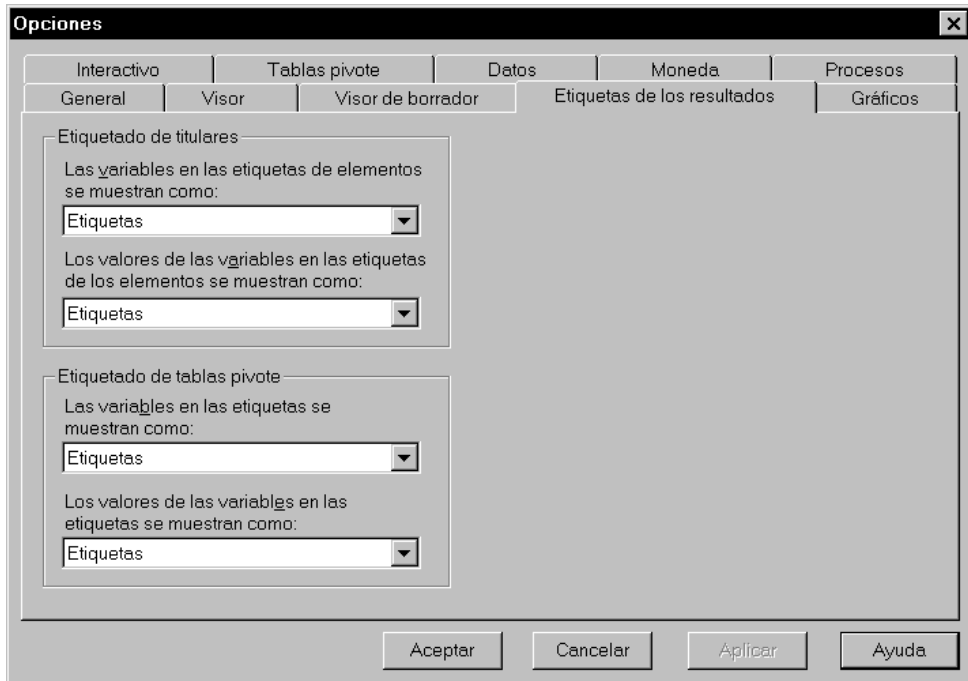
Resultados de texto. En los resultados de texto que no sean resultados de tablas pivote convertidos, controla el ancho de página (expresado en número de caracteres) y la longitud de página (expresada en número de líneas). En algunos procedimientos, algunos estadísticos se muestran sólo en formato ancho.

Etiqueta de los resultados: Opciones

Las opciones de Etiqueta de los resultados controlan la presentación de la información sobre variables y valores de datos en las tablas pivote y en los titulares. Puede visualizar nombres de variable, etiquetas de variable definidas y valores de datos reales, etiquetas de valor definidas o una combinación de estas opciones.

Etiquetas descriptivas de variable y de valor (Vista de variables del Editor de datos, columnas de *Etiqueta* y *Valores*) a menudo facilitan la interpretación de los resultados. Sin embargo, las etiquetas largas pueden crear dificultades en algunas tablas.

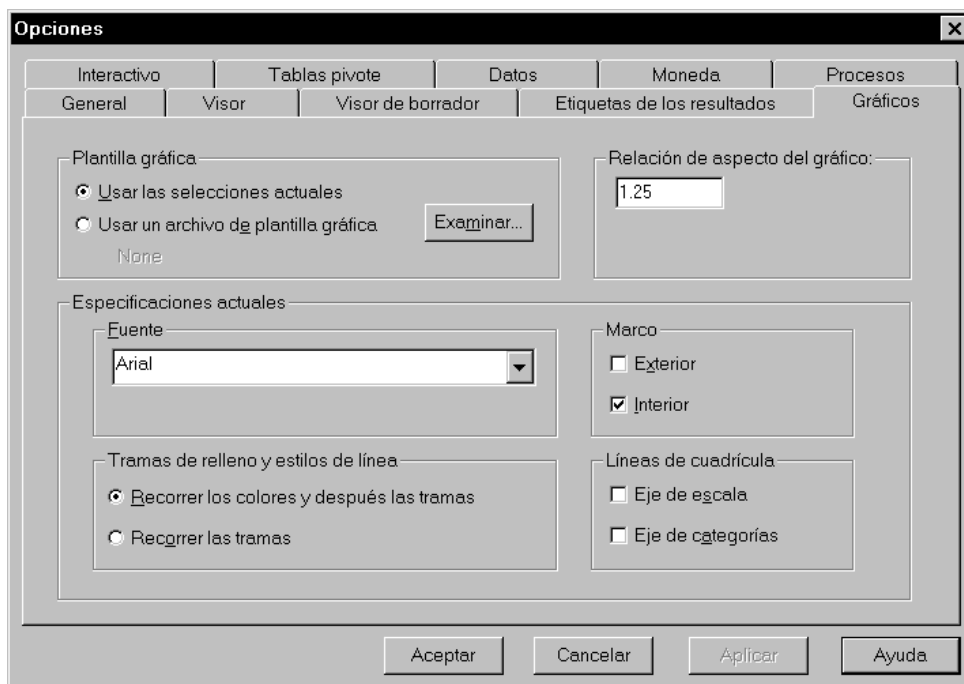
Figura 43-4
Pestaña Etiquetas de los resultados del cuadro de diálogo Opciones



Las opciones de etiquetas de los resultados afectan sólo a los nuevos resultados que se producen después de cambiar las selecciones. A los resultados que ya se muestran en el Visor no les afectan los cambios en la configuración. Esta configuración sólo afecta a los resultados de las tablas pivote. Los resultados de texto no se verán afectados por estas selecciones.

Opciones de gráfico

Figura 43-5
Pestaña Gráficos del cuadro de diálogo Opciones



Plantilla gráfica. Los gráficos nuevos pueden utilizar tanto las opciones seleccionadas aquí, como las opciones de un archivo de plantilla gráfica. Pulse en Examinar para seleccionar un archivo de plantilla gráfica. Para crear un archivo de plantilla gráfica, cree un gráfico con los atributos que desee y guárdelo como una plantilla (seleccione Guardar plantilla gráfica en el menú Archivo).

Relación de aspecto de los gráficos. La relación ancho-alto del marco exterior de los nuevos gráficos. Puede especificar una relación ancho-alto entre los valores 0,1 y 10,0. Los valores inferiores a 1 generan gráficos que son más altos que anchos. Los valores mayores que 1 producen gráficos que son más anchos que altos. Un valor de 1 produce un gráfico cuadrado. Una vez creado un gráfico, no es posible cambiar su relación de aspecto.

Fuente. La fuente utilizada para todo el texto en los nuevos gráficos.

Preferencia de ciclos de estilo. La asignación inicial de colores y tramas para nuevos gráficos. Mostrar sucesivamente los colores y, a continuación, las tramas utiliza la paleta por defecto de colores y luego cambia el estilo de línea o el símbolo de marcador o añade una trama de relleno según sea necesario. Mostrar sucesivamente sólo los colores utiliza únicamente colores para diferenciar los elementos gráficos y no utiliza tramas. Mostrar sucesivamente sólo las tramas sólo utiliza estilos de línea, símbolos de marcador o tramas de relleno para diferenciar los elementos gráficos y no utiliza color.

Marco. Controla la presentación de los marcos interno y externo en los nuevos gráficos.

Líneas de cuadrícula. Controla la presentación de las líneas de cuadrícula de los ejes de categorías y escala en los nuevos gráficos.

Ciclos de estilo. Personaliza los colores, estilos de línea, símbolos de marcador y tramas de relleno para los gráficos nuevos. Puede cambiar el orden de los colores o tramas utilizados al crear un gráfico nuevo.

Nota: Estas opciones no afectan a los gráficos interactivos (menú Gráficos, submenú Interactivo).

Colores de los elementos de datos

Especifique el orden en que se deben utilizar los colores para los elementos de datos (por ejemplo, barras y marcadores) en el nuevo gráfico. Los colores se utilizan siempre que se selecciona una opción que incluye *color* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de barras agrupadas con dos grupos y selecciona Mostrar sucesivamente los colores y, a continuación, las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros colores de la lista Gráficos agrupados se utilizan como colores de barra en el nuevo gráfico.

Para cambiar el orden en que se utilizan los colores:

- Seleccione Gráficos simples y, a continuación, seleccione un color utilizado para los gráficos sin categorías.

- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de color de los gráficos con categorías. Para cambiar el color de una categoría, selecciónela y elija un color en la paleta para dicha categoría.

Si lo desea, tiene la posibilidad de:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.
- Editar un color mediante la selección de su casilla y pulsando en Editar.

Líneas de los elementos de datos

Especifique el orden en que se deben utilizar los estilos para los elementos de datos de línea en el nuevo gráfico. Los estilos de línea se utilizan siempre que en el gráfico hay elementos de datos de línea y se selecciona una opción que incluye *tramas* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de líneas con dos grupos y selecciona Mostrar sucesivamente sólo las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros estilos de la lista Gráficos agrupados se utilizan como tramas de línea en el nuevo gráfico.

Para cambiar el orden en que se utilizan los estilos de línea:

- ▶ Seleccione Gráficos simples y, a continuación, seleccione un estilo de línea utilizado para los gráficos de líneas sin categorías.
- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de trama de los gráficos de líneas con categorías. Para cambiar el estilo de línea de una categoría, selecciónela y elija un estilo de línea en la paleta para dicha categoría.

Si lo desea, tiene la posibilidad de:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.

- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.

Marcadores de los elementos de datos

Especifique el orden en que se deben utilizar los símbolos para los elementos de datos de marcadores en el nuevo gráfico. Los estilos de marcadores se utilizan siempre que en el gráfico hay elementos de datos de marcadores y se selecciona una opción que incluye *tramas* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de diagramas de dispersión con dos grupos y selecciona Mostrar sucesivamente sólo las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros símbolos de la lista Gráficos agrupados se utilizan como marcadores en el nuevo gráfico.

Para cambiar el orden en que se utilizan los estilos de marcadores:

- ▶ Seleccione Gráficos simples y, a continuación, seleccione un símbolo de marcador utilizado para los gráficos sin categorías.
- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de trama de los gráficos con categorías. Para cambiar el símbolo de marcador de una categoría, selecciónela y elija un símbolo en la paleta para dicha categoría.

Si lo desea, tiene la posibilidad de:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.

Rellenos de los elementos de datos

Especifique el orden en que se deben utilizar los estilos de relleno para los elementos de datos de la barra y área en el nuevo gráfico. Los estilos de relleno se utilizan siempre que en el gráfico hay elementos de datos de barra o área y se selecciona una

opción que incluye *tramas* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de barras agrupadas con dos grupos y selecciona Mostrar sucesivamente sólo las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros estilos de la lista Gráficos agrupados se utilizan como tramas de relleno de barra en el nuevo gráfico.

Para cambiar el orden en que se utilizan los estilos de relleno:

- ▶ Seleccione Gráficos simples y, a continuación, seleccione una trama de relleno utilizada para los gráficos sin categorías.
- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de trama de los gráficos con categorías. Para cambiar la trama de relleno de una categoría, selecciónela y elija una trama de relleno en la paleta para dicha categoría.

Si lo desea, tiene la posibilidad de:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.

Opciones de gráficos interactivos

Figura 43-6
Pestaña Interactivo del cuadro de diálogo Opciones



Para los gráficos interactivos (menú Gráficos, submenú Interactivo) se encuentran disponibles las siguientes opciones:

Aspecto de gráfico. Seleccione un aspecto de gráfico en la lista de archivos y pulse en Aceptar o Aplicar. Por defecto, la lista muestra los aspectos de gráfico guardados en el directorio *Looks* del directorio donde se instaló el programa. Puede utilizar uno de los aspectos de gráfico que se incluyen en SPSS, o bien crear uno propio en el Editor de gráficos interactivos (menú Formato, Aspectos de gráfico con un gráfico activado).

- **Directorio.** Permite seleccionar un directorio de aspecto de gráfico. Utilice Examinar para añadir directorios a la lista.
- **Examinar.** Permite seleccionar un aspecto de gráfico de otro directorio.

Datos guardados con el gráfico. Controla la información que se guarda con los gráficos interactivos cuando dejan de estar anexados al archivo de datos que los creó (por ejemplo, si abre un archivo del Visor guardado en una sesión anterior).

El almacenamiento de los datos con el gráfico permite utilizar la mayoría de las funciones interactivas disponibles para los gráficos anexados al archivo de datos que los creó (exceptuando la adición de variables no incluidas en el gráfico original). Sin embargo, esta opción puede aumentar considerablemente el tamaño de los archivos del Visor, especialmente el de los archivos de datos de gran tamaño.

Resolución de impresión. Determina la resolución de impresión de los gráficos interactivos. En la mayoría de los casos, Metarchivo de vectores imprimirá más rápido y proporcionará los mejores resultados. Para mapas de bits, los gráficos de baja resolución se imprimen con mayor rapidez; los de alta resolución son de mejor calidad.

Unidades de medida. Sistema de medida utilizado (puntos, pulgadas o centímetros) para especificar atributos tales como el tamaño de la región de datos de un gráfico.

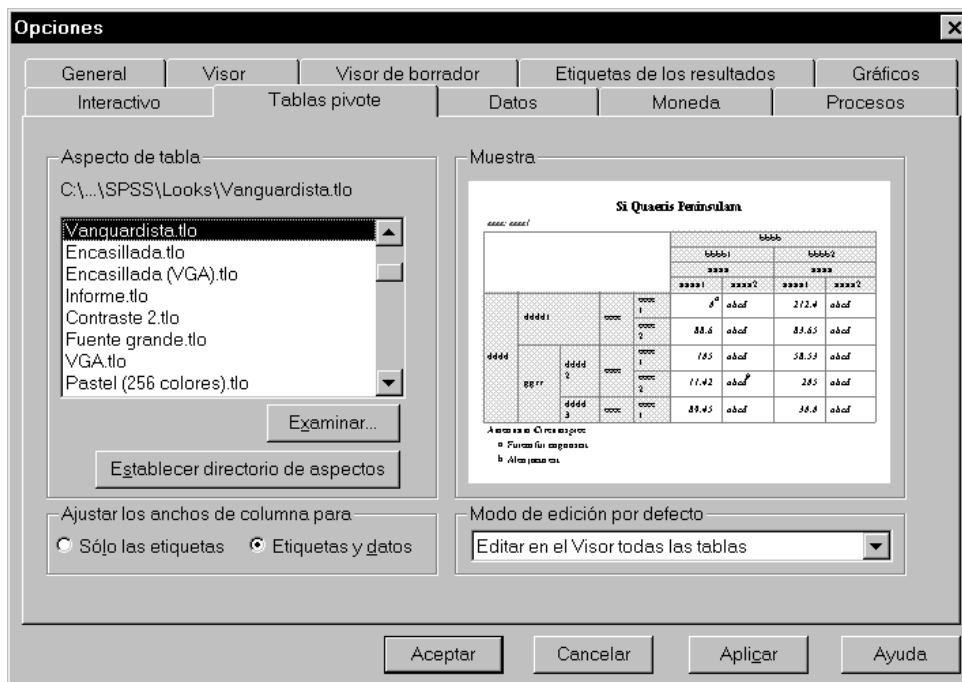
Lectura de archivos de datos anteriores a 8.0. Para los archivos de datos creados en versiones anteriores de SPSS es posible especificar el número mínimo de valores de datos que deben considerarse en una variable numérica para clasificarla como variable de escala o categórica. Las variables con un número de valores únicos inferior al especificado se clasificarán como categóricas. Toda variable con etiquetas de valor definidas se clasifica como categórica independientemente del número de valores únicos que posea.

Nota: Estas opciones sólo afectan a los gráficos interactivos (menú Gráficos, submenú Interactivo).

Tabla pivote: Opciones

Las opciones de Tabla pivote establecen el aspecto de tabla por defecto, utilizado para los resultados de nuevas tablas pivote. Aspectos de tabla puede controlar una variedad de atributos de tabla pivote, incluyendo la presentación y el ancho de las líneas de cuadrícula; el estilo, tamaño y color de la fuente; así como los colores de fondo.

Figura 43-7
Pestaña Tablas pivote del cuadro de diálogo Opciones



Aspecto de tabla. Seleccione un aspecto de tabla en la lista de archivos y pulse en Aceptar o Aplicar. Por defecto, la lista muestra los aspectos de tabla guardados en el directorio *Looks* del directorio donde se instaló el programa. Puede utilizar uno de los aspectos de tabla que se incluye en SPSS, o bien crear uno propio en el editor de tablas pivote (menú Formato, Aspectos de tabla).

- **Examinar.** Permite seleccionar un aspecto de tabla de otro directorio.
- **Establecer directorio de aspectos.** Le permite cambiar el directorio de aspectos por defecto.

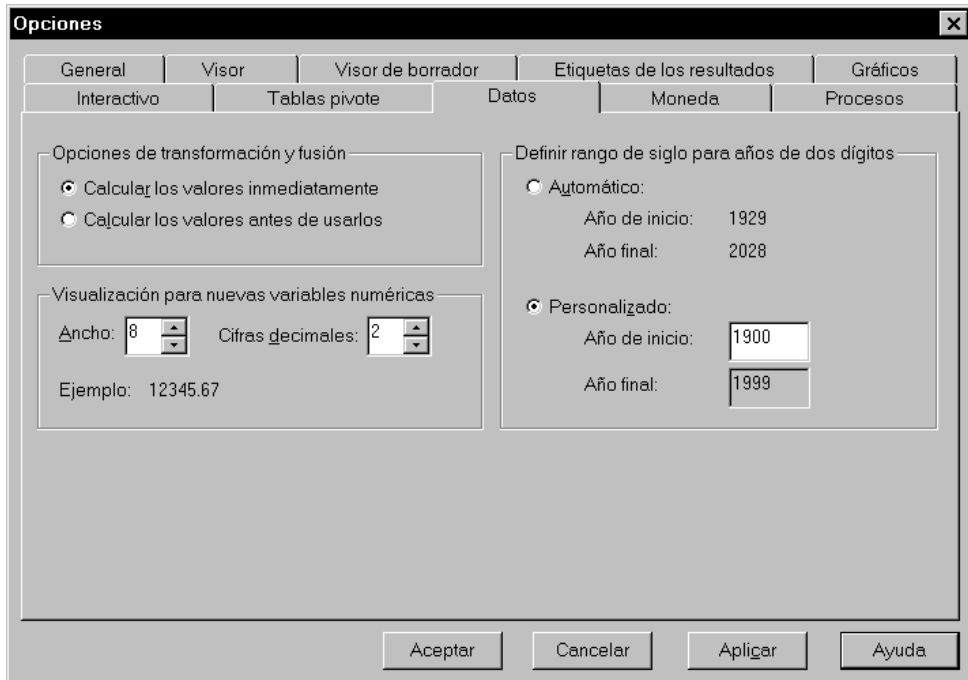
Ajustar los anchos de columna para Controla el ajuste automático de los anchos de columna en las tablas pivote.

- **Sólo las etiquetas.** Ajusta el ancho de columna al ancho de la etiqueta de la columna. Así se generan tablas más compactas, pero no se mostrarán los valores de los datos más anchos que la etiqueta (los asteriscos indican valores demasiado anchos para ser mostrados).
- **Etiquetas y Datos.** Ajusta el ancho de columna al más ancho de entre la etiqueta de columna y el mayor de los valores de los datos. Así se generan tablas más anchas, pero se asegura que se mostrarán todos los valores.

Modo de edición por defecto. Controla la activación de las tablas pivote en la ventana del Visor o en una ventana independiente. Por defecto, cuando se pulsa dos veces en una tabla pivote, ésta se activa en la ventana del Visor. Puede elegir entre activar las tablas pivote en una ventana independiente o seleccionar una opción de tamaño que abra las tablas pivote más pequeñas en la ventana del Visor y las más grandes en una ventana independiente.

Datos: Opciones

Figura 43-8
Pestaña Datos del cuadro de diálogo Opciones



Opciones de transformación y fusión. Cada vez que SPSS ejecuta un comando, lee el archivo de datos. Algunas transformaciones de datos (tales como Calcular y Recodificar) no requieren una lectura diferente de los datos; esto permite postergar su ejecución hasta que el programa lea los datos para ejecutar otro comando, como puede ser un procedimiento estadístico. Para los archivos de datos largos, seleccione Calcular los valores antes de usarlos para retrasar la ejecución y ganar tiempo de procesamiento.

Formato de presentación para las nuevas variables numéricas. Controla la presentación por defecto del ancho y el número de posiciones decimales de las nuevas variables numéricas. No existe formato de presentación por defecto para las nuevas variables de cadena. Si un valor es demasiado largo para el formato de presentación especificado, primero se redondean las posiciones decimales y después los valores se convierten a notación científica. Los formatos de presentación no afectan a los valores de datos

internos. Por ejemplo, el valor 123456,78 se puede redondear a 123457 para la presentación, pero se utilizará el valor original sin redondear en cualquier cálculo.

Definir rango de siglo para años de dos dígitos. Define el rango de años para las variables con formato de fecha introducidas o mostradas con un año de dos dígitos (por ejemplo, 10/28/86, 29-OCT-87). La opción de rango automático se basa en el año actual; es decir, comienza 69 años antes del actual y finaliza 30 años después (sumando el año en curso hace un total de 100 años). En el rango personalizado, el año final se establece de forma automática en función del valor introducido en el año inicial.

Generador de números aleatorios. Hay dos generadores de números aleatorios disponibles:

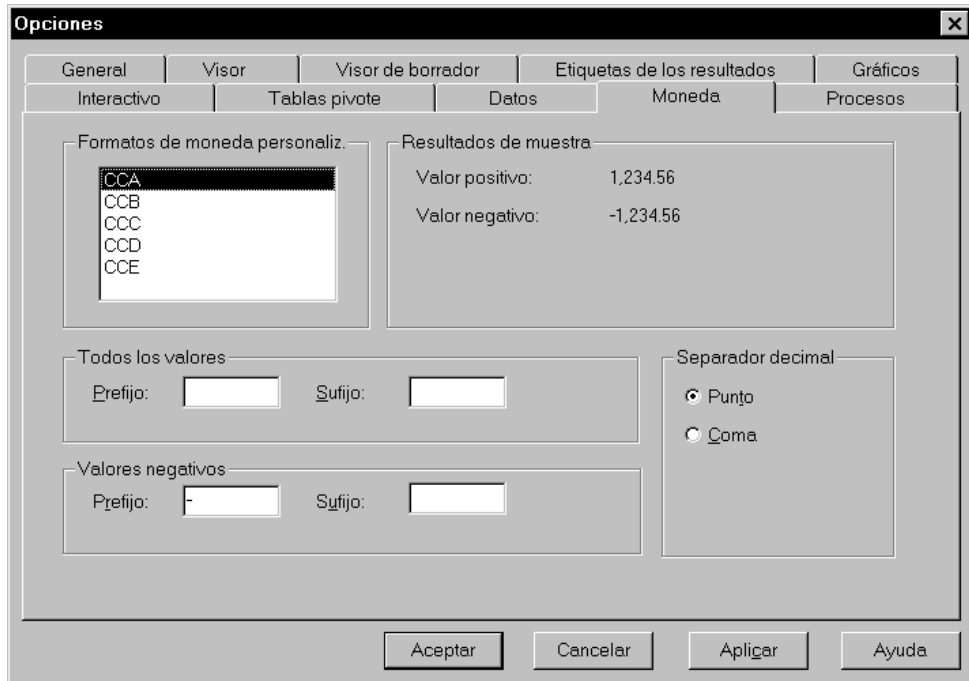
- **Compatible con SPSS 12.** El generador de números aleatorios utilizado en SPSS 12 y versiones anteriores. Utilice este generador de números aleatorios si necesita reproducir los resultados aleatorizados generados por versiones previas basadas en una semilla de aleatorización especificada.
- **Tornado de Mersenne.** Un generador de números aleatorios nuevo que es más fiable en los procesos de simulación. Utilice este generador de números aleatorios si no es necesario reproducir resultados aleatorizados correspondientes a SPSS 12 o anteriores.

Moneda: Opciones

Puede crear hasta cinco formatos de presentación de moneda personalizados que pueden incluir caracteres de prefijo y sufijo especiales además de un tratamiento especial para los valores negativos.

Los nombres de los cinco formatos de moneda personalizados son MPA, MPB, MPC, CCD y MPE. No se pueden cambiar los nombres de los formatos ni añadir otros nuevos. Para modificar un formato de moneda personalizado, seleccione el nombre del formato de la lista de origen y realice los cambios que desee.

Figura 43-9
Pestaña Moneda del cuadro de diálogo Opciones



Los prefijos, los sufijos y los separadores decimales definidos para los formatos monetarios personalizados sólo afectan a la presentación en la pantalla. No es posible introducir valores en el Editor de datos utilizando caracteres de moneda personalizados.

Para crear formatos de moneda personalizados

- ▶ Pulse en la pestaña Moneda.
- ▶ Seleccione uno de los formatos de moneda de la lista (MPA, MPB, MPC, MPD y MPE).
- ▶ Introduzca el prefijo, el sufijo y los valores indicadores de decimales.
- ▶ Pulse en Aceptar o Aplicar.

Procesos: Opciones

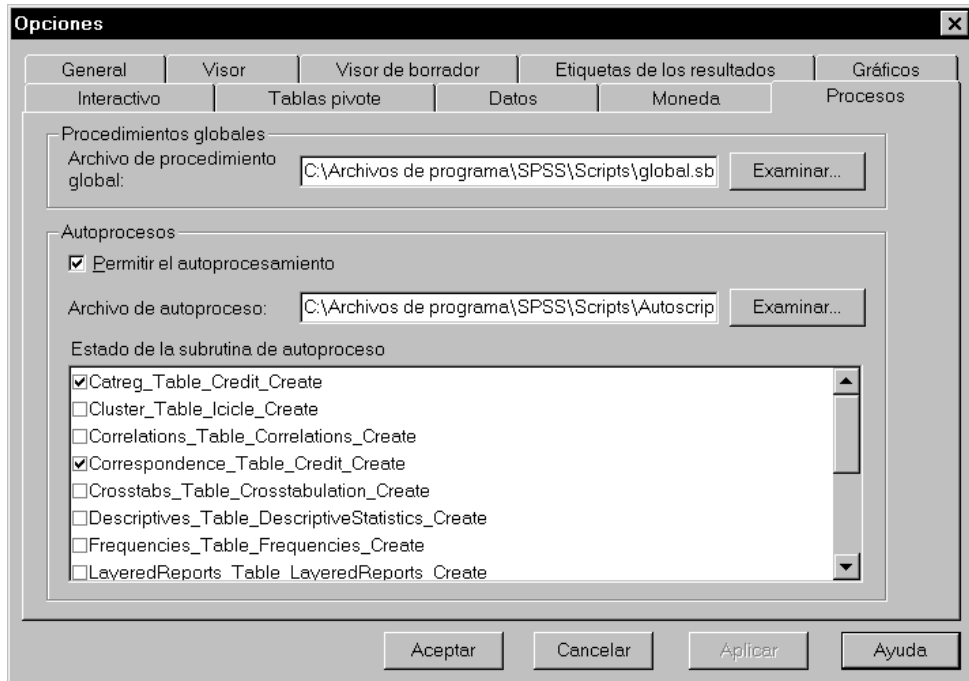
Utilice la pestaña Procesos para especificar el archivo de procedimientos global y el archivo de autoproceto, y para seleccionar las subrutinas de autoproceto que desee utilizar. Puede utilizar procesos para automatizar muchas funciones, incluyendo la personalización de tablas pivote.

Procedimientos globales. Un archivo de procedimientos globales es una biblioteca de subrutinas de proceso y funciones que pueden ser invocadas por archivos de proceso, incluyendo los archivos de autoproceto.

Nota: los procedimientos globales que acompañan al programa son los que aparecen seleccionados por defecto. Muchos de los procesos que se ofrecen utilizan funciones y subrutinas de este archivo de procedimientos global y no funcionarán si se especifica un archivo de procedimientos diferente.

Autoprocetos. Un archivo de autoproceto es un conjunto de subrutinas de proceso que se ejecutan automáticamente cada vez que se ejecutan procedimientos que crean ciertos tipos de objetos de resultados.

Figura 43-10
Pestaña Procesos del cuadro de diálogo Opciones



Muestra todas las subrutinas del archivo de autoprocesos actual, permitiendo activar y desactivar subrutinas individuales.

Para especificar las opciones de autoprocesos y procedimientos globales

- ▶ Pulse en la pestaña Procesos.
- ▶ Seleccione Permitir el autoprocesamiento.
- ▶ Seleccione las subrutinas de autoproceso que desea activar.

También se puede especificar un archivo de autoproceso o un archivo de procedimiento global diferente.

Personalización de menús y barras de herramientas

Editor de menús

Puede personalizar los menús utilizando el Editor de menús. Con el Editor de menús es posible:

- Añadir elementos de menú que ejecuten procesos personalizados.
- Añadir elementos de menú que ejecuten archivos de sintaxis de comandos.
- Añadir elementos de menú que ejecuten otras aplicaciones y envíen los datos automáticamente a otras aplicaciones.

Puede enviar datos a otras aplicaciones en los siguientes formatos: SPSS, Excel 4.0, Lotus 1-2-3 versión 3, SYLK, delimitado por tabuladores y dBASE IV.

Para añadir elementos a los menús

- ▶ Elija en los menús:
Utilidades
Editor de menús...
- ▶ En el cuadro de diálogo Editor de menús, pulse dos veces en el menú donde desee añadir un nuevo elemento.
- ▶ Seleccione el elemento de menú sobre el que desea que aparezca el nuevo elemento.
- ▶ Pulse en Insertar elemento para insertar un nuevo elemento de menú.

- ▶ Seleccione el tipo de archivo para el nuevo elemento (archivo de proceso, archivo de sintaxis de comandos o aplicación externa).
- ▶ Pulse en Examinar para seleccionar un archivo que sea anexionado al elemento de menú.

Figura 44-1
Cuadro de diálogo Editor de menús



También se pueden añadir menús completamente nuevos y separadores entre los elementos de menú.

Si lo desea, puede enviar automáticamente el contenido del Editor de datos a otra aplicación cuando seleccione esa aplicación en los menús.

Personalización de las barras de herramientas

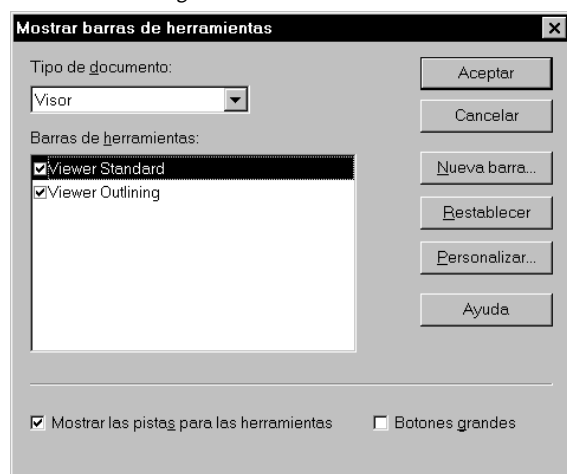
Puede personalizar las barras de herramientas y crear nuevas barras de herramientas. En las barras de herramientas puede incluirse cualquier herramienta disponible, incluso la de cualquier acción de menú. Además pueden contener herramientas personalizadas que ejecutan otras aplicaciones, que ejecutan archivos de sintaxis de comandos o archivos de procesos.

Mostrar barras de herramientas

Utilice Mostrar barras de herramientas para mostrar u ocultar, personalizar y crear nuevas barras de herramientas. En las barras de herramientas puede incluirse cualquier herramienta disponible, incluso la de cualquier acción de menú. Además pueden contener herramientas personalizadas que ejecutan otras aplicaciones, que ejecutan archivos de sintaxis de comandos o archivos de procesos.

Figura 44-2

Cuadro de diálogo Mostrar barras de herramientas



Para personalizar las barras de herramientas

- ▶ Elija en los menús:
Ver
Barras de herramientas...
- ▶ Seleccione la barra de herramientas que desea personalizar y pulse en Personalizar o pulse en Nueva barra para crear una nueva barra de herramientas.
- ▶ Para las barras de herramientas nuevas, introduzca un nombre para la barra de herramientas, seleccione las ventanas en las que desea que aparezca y pulse en Personalizar.

- ▶ Seleccione un elemento en la lista Categorías para que se visualicen las herramientas disponibles en esa categoría.
- ▶ Arrastre y suelte las herramientas que desee en la barra de herramientas que aparece en el cuadro de diálogo.
- ▶ Para eliminar una herramienta de la barra de herramientas, arrástrela a cualquier punto fuera de la barra de herramientas que aparece en el cuadro de diálogo.

Para crear una herramienta personalizada que abra un archivo, ejecute un archivo de sintaxis de comandos o ejecute un proceso:

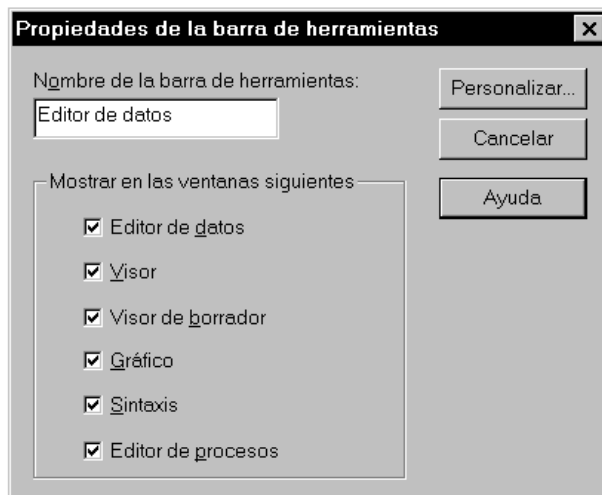
- ▶ Pulse en Nueva herramienta en el cuadro de diálogo Personalizar barra de herramientas.
- ▶ Introduzca una etiqueta descriptiva para la herramienta.
- ▶ Seleccione la acción que desee realizar con la herramienta (abrir un archivo, ejecutar un archivo de sintaxis de comandos o ejecutar un proceso).
- ▶ Pulse en Examinar para seleccionar un archivo o una aplicación para asociarlos a la herramienta.

Las nuevas herramientas se muestran en la categoría Personales , que además contiene los elementos de menú definidos por el usuario.

Propiedades de la barra de herramientas

Utilice Propiedades de la barra de herramientas para seleccionar los tipos de ventana en los que desee que aparezca la barra de herramientas seleccionada. Este cuadro de diálogo también se utiliza para crear nombres para nuevas barras de herramientas.

Figura 44-3
Cuadro de diálogo Propiedades de la barra de herramientas



Para establecer las propiedades de la barra de herramientas

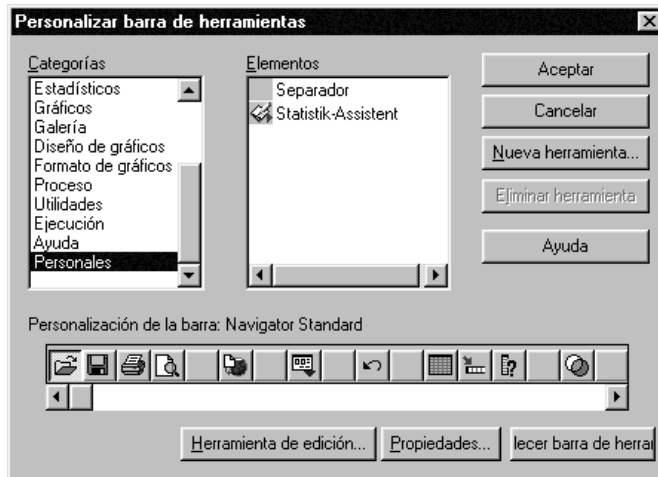
- ▶ Elija en los menús:
Ver
Barras de herramientas...
- ▶ Para las barras de herramientas existentes, pulse en Personalizar y, a continuación, en Propiedades en el cuadro de diálogo Personalizar barra de herramientas.
- ▶ En el caso de barras de herramientas nuevas, pulse en Nueva herramienta.
- ▶ Seleccione los tipos de ventana en los que desee que aparezca la barra de herramientas. Para las barras de herramientas nuevas, introduzca además un nombre.

Personalizar barra de herramientas

Utilice el cuadro de diálogo Personalizar barra de herramientas para personalizar las barras de herramientas existentes y para crear nuevas barras. En las barras de herramientas puede incluirse cualquier herramienta disponible, incluso la de

cualquier acción de menú. Además pueden contener herramientas personalizadas que ejecutan otras aplicaciones, que ejecutan archivos de sintaxis de comandos o archivos de procesos.

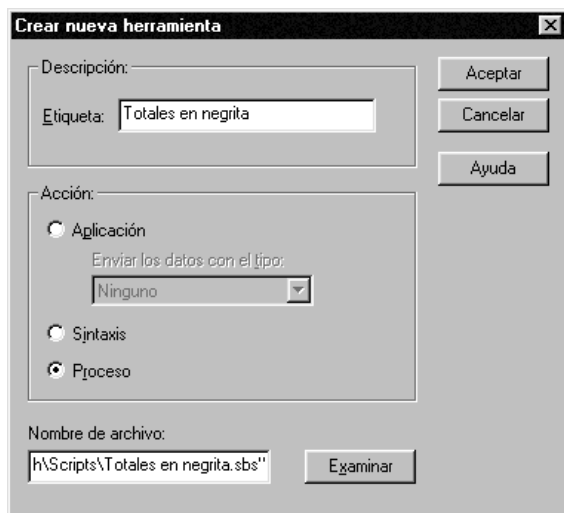
Figura 44-4
Cuadro de diálogo Personalizar barra de herramientas



Crear nueva herramienta

Utilice el cuadro de diálogo Crear nueva herramienta para crear herramientas personalizadas para ejecutar otras aplicaciones, ejecutar archivos de sintaxis de comandos y ejecutar archivos de procesos.

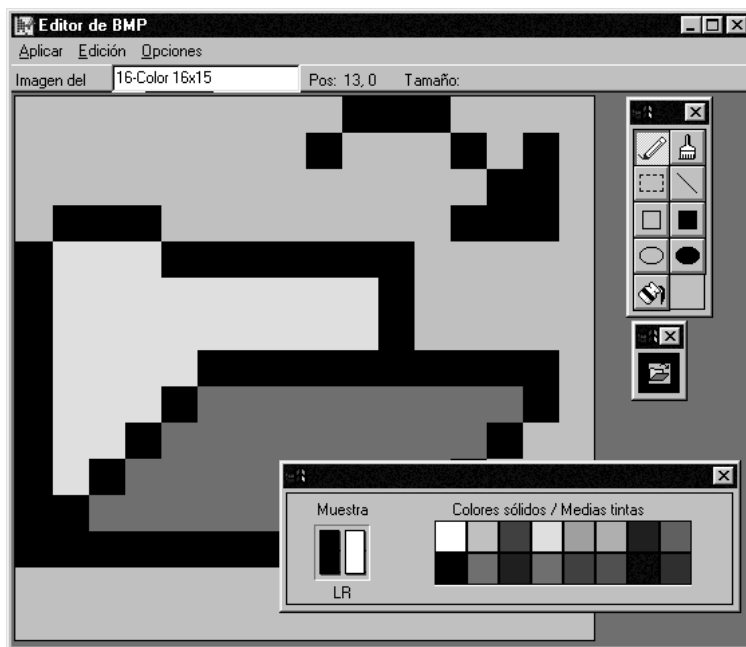
Figura 44-5
Cuadro de diálogo Crear nueva herramienta



Editor de mapa de bits de barras de herramientas

Utilice el Editor de mapa de bits para crear iconos personalizados para los botones de la barra de herramientas. Esto es especialmente útil cuando se crean herramientas personalizadas para ejecutar procesos, sintaxis y otras aplicaciones.

Figura 44-6
Editor de mapa de bits



Para editar los mapas de bits de las barras de herramientas

- ▶ Elija en los menús:
 - Ver
 - Barras de herramientas...
- ▶ Seleccione la herramienta que desee personalizar y pulse en Personalizar.
- ▶ Pulse en la herramienta que tenga el icono de mapa de bits que desea editar en la barra de herramientas de ejemplo.
- ▶ Pulse en Editar herramienta.
- ▶ Utilice la caja de herramientas y la paleta de colores para modificar el mapa de bits o crear un nuevo icono de mapa de bits.

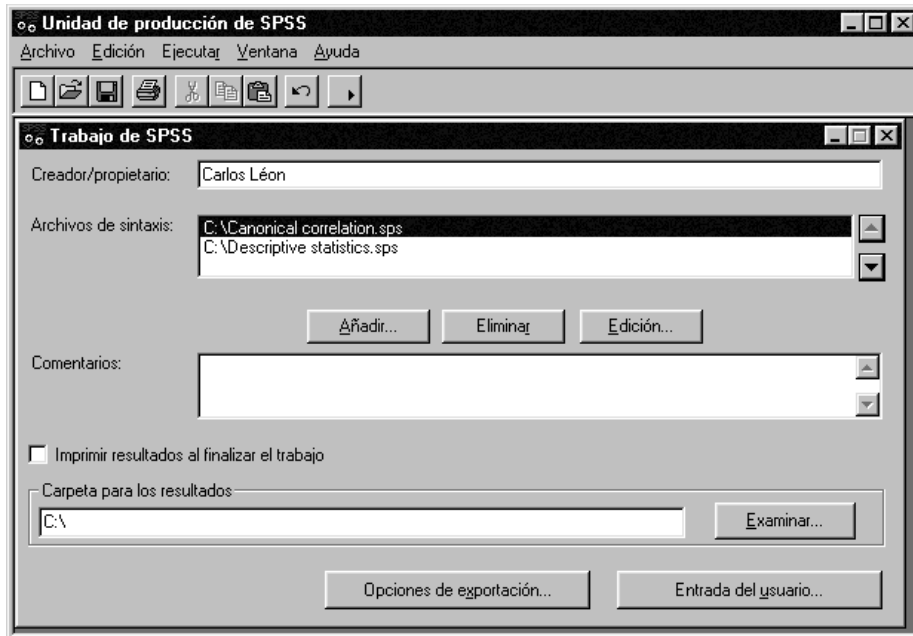
Unidad de producción

La Unidad de producción ofrece la posibilidad ejecutar el programa en modo automático. El programa se ejecuta por su cuenta y finaliza tras ejecutar el último comando, de esta forma podrá realizar mientras tanto otras tareas. Modo de producción es útil si se ejecuta a menudo el mismo conjunto de análisis que consumen mucho tiempo, tales como los informes semanales.

La unidad de producción utiliza archivos de sintaxis de comandos para indicarle al programa qué hacer. Un archivo de sintaxis de comandos es un archivo de sólo texto que contiene sintaxis de comandos. Puede utilizar cualquier editor de textos para crearlo. También puede generar la sintaxis de comandos pegando las selecciones del cuadro de diálogo en una ventana de sintaxis o editando el archivo de diario.

Después de crear los archivos de sintaxis y de incluirlos en un trabajo de producción, puede revisarlos y editarlos en la propia Unidad de producción.

Figura 45-1
Unidad de producción



Formato de entrada de sintaxis. Controla la forma de las reglas de sintaxis utilizadas para el trabajo:

- **Interactivo.** Cada comando debe finalizar con un punto. Los puntos pueden aparecer en cualquier parte del comando y los comandos pueden continuar en varias líneas, aunque un punto como último carácter no en blanco en una línea se interpreta como el final del comando. Las líneas de continuación y los comandos nuevos pueden empezar en cualquier parte de una línea nueva. Se trata de reglas “interactivas” que surten efecto al seleccionar y ejecutar comandos en una ventana de sintaxis.
- **Lote.** Cada comando debe comenzar al principio de una línea nueva (sin espacios en blanco antes del inicio del comando) y las líneas de continuación se deben sangrar como mínimo un espacio. Si desea sangrar comandos nuevos, puede utilizar un signo más o un punto como primer carácter al inicio de la línea y, a continuación, sangrar el comando. El punto del final del comando es opcional.

Este ajuste es compatible con las reglas de sintaxis para los archivos de comandos incluidos con el comando `INCLUDE`.

Comportamiento de errores de sintaxis. Controla el tratamiento de errores en el trabajo:

- **Continuar.** Los errores del trabajo no detienen automáticamente el procesamiento de los comandos. Los comandos de los archivos de trabajo de producción se tratan como parte de la secuencia de comandos normal y el procesamiento de comandos continúa del modo normal.
- **Detener.** El procesamiento de comandos se detiene si se detecta un primer error en el archivo de trabajo de producción. Esto es compatible con el comportamiento de los archivos de comandos incluidos con el comando `INCLUDE`.

Resultados del trabajo de producción. Cada ejecución de producción crea un archivo de resultados con el mismo nombre que el trabajo de producción y la extensión `.spo`. Por ejemplo, un archivo de trabajo de producción llamado `trabprod.spp` creará un archivo de resultados llamado `trabprod.spo`. El archivo de resultados es un documento del Visor.

Tipos de resultados. Los resultados del Visor generan tablas pivote y gráficos interactivos de alta resolución. Los resultados del Visor de borrador producen resultados de texto e imágenes de metarchivo de gráficos. Los resultados de texto se pueden editar en el Visor de borrador, pero no los gráficos.

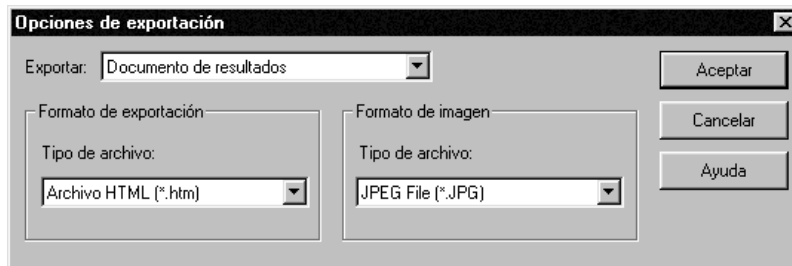
Uso de la Unidad de producción

- ▶ Elabore un archivo de sintaxis de comandos.
- ▶ Arranque la Unidad de producción desde el menú Inicio.
- ▶ Especifique los archivos de sintaxis que desee utilizar en el trabajo de producción. Pulse en **Añadir** para seleccionarlos.
- ▶ Guarde el archivo de trabajo de producción.
- ▶ Ejecute el archivo. Para ello, pulse en el botón **Ejecutar** de la barra de herramientas o bien elija en los menús:
 - Ejecutar
 - Production Job (Trabajo de producción)

Opciones de exportación

Exportar resultados guarda las tablas pivote y los resultados de texto en formato HTML, de texto, Word/RTF y Excel y guarda los gráficos en una amplia variedad de formatos comunes utilizados por otras aplicaciones.

Figura 45-2
Cuadro de diálogo de opciones de exportación



Exportar

Esta lista desplegable especifica lo que se desea exportar.

Documento de salida. Exporta cualquier combinación de tablas pivote, resultados de texto y gráficos.

- Para los formatos HTML y de texto, los gráficos se exportan en el formato de exportación de gráficos seleccionado actualmente. Si se trata de documentos HTML, los gráficos se incrustan mediante referencias y deben exportarse a un formato adecuado para incluirlos en los documentos HTML. En el caso de documentos de texto se inserta una línea en el archivo de texto para cada gráfico, indicando el nombre de archivo del gráfico exportado.
- Para los formatos Word o RTF, los gráficos se exportan en formato de metarchivo de Windows y se incrustan en el documento de Word.
- No se incluyen gráficos en los documentos de Excel.

Documento de resultados (sin gráficos). Exporta tablas pivote y resultados de texto. Se ignoran todos los gráficos que se muestran en el Visor.

Sólo gráficos. Sólo exporta gráficos. Para documentos HTML y de texto, los formatos de exportación son: Metarchivo mejorado (EMF), metarchivo de Windows (WMF), mapa de bits de Windows (BMP), PostScript encapsulado (EPS), JPEG, TIFF, PNG

y PICT de Macintosh. Para documentos de Word o RTF, los gráficos siempre se exportan en formato de metarchivo de Windows (WMF).

Formato de exportación

Para los documentos de resultados, las opciones disponibles son HTML, texto, Word/RTF y Excel; para los formatos HTML y de texto, los gráficos se exportan en el formato de gráfico seleccionado actualmente. Para Sólo gráficos, seleccione un formato de exportación del gráfico en la lista desplegable. Para exportar documentos de resultados, tablas pivote y texto, siga las instrucciones siguientes:

- **Archivo HTML (*.htm).** Las tablas pivote se exportan como tablas HTML. Los resultados de texto se exportan como formato previo de HTML.
- **Archivo de texto (*.txt).** Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Todos los resultados de texto se exportan en formato separado por espacios.
- **Archivo de Excel (*.xls).** Las filas, columnas y casillas de tablas pivote se exportan como filas, columnas y casillas de Excel, con todos los atributos de formato intactos (por ejemplo, bordes de casilla, estilos de fuente, colores de fondo, etc.). Los resultados de texto se exportan con todos los atributos de fuente intactos. Cada línea del resultado de texto constituye una fila del archivo de Excel y se incluye todo su contenido en una sola casilla.
- **Archivo de Word/RTF (*.doc).** Las tablas pivote se exportan como tablas de Word, con todos los atributos de formato intactos (por ejemplo, bordes de casilla, estilos de fuente, colores de fondo, etc.). Los resultados de texto se exportan en formato RTF. Los resultados de texto siempre se muestran en SPSS con un tipo de letra de paso fijo (monoespaciadas) y se exportan con los mismos atributos de fuente. Es necesario utilizar una fuente de paso fijo para que los resultados de texto separado por espacios queden correctamente alineados; por defecto, la mayoría de los navegadores de Web utilizan una fuente de paso fijo para el texto con formato previo.

Formato de imagen

Formato de imagen controla el formato de exportación de los gráficos. Los formatos disponibles son: Metarchivo mejorado (EMF), metarchivo de Windows (WMF), mapa de bits de Windows (BMP), PostScript encapsulado (EPS), JPEG, TIFF, PNG o PICT de Macintosh.

Los nombres de los gráficos exportados se basan en el nombre de archivo del trabajo de producción, un número secuencial y la extensión del formato seleccionado. Por ejemplo, si el trabajo de producción *prodjob.spp* exporta gráficos al formato de metarchivo de Windows, los nombres del gráfico serían *prodjob1.wmf*, *prodjob2.wmf*, *prodjob3.wmf* y así sucesivamente.

Opciones de texto e imágenes

Las opciones de exportación de texto (por ejemplo, separado por tabuladores o por espacios) y las opciones de exportación de gráficos (como el color, el tamaño y la resolución) están predefinidas en SPSS y no se pueden cambiar en la Unidad de producción. Utilice la opción Exportar del menú Archivo en SPSS para modificarlas.

Exportación del Visor de borrador

El Visor de borrador sólo permite exportar los resultados en formato de sólo texto. No se pueden exportar los gráficos de los resultados.

Entradas del usuario

Los símbolos de macro definidos en un archivo de trabajo de producción y utilizados en un archivo de sintaxis de comandos simplifican tareas tales como la ejecución del mismo análisis con diversos archivos de datos o del mismo conjunto de comandos con distintos conjuntos de variables. Por ejemplo, podría definir el símbolo de macro *@archdat* para que se solicite el nombre de un archivo de datos cada vez que se ejecute un trabajo de producción que utilice la cadena *@archdat* en vez de un nombre de archivo.

Figura 45-3
Cuadro de diálogo de entradas del usuario

Símbolo de macro	Entrada	Por defecto	¿Desea poner el valor entre comillas?
@datafie	nombre y situación de los datos	c:\spss\Datos de empleados.sav	Sí
@varlist	lista de los variables	catlab sexo raza	No

@datafie

Aceptar Cancelar Ayuda

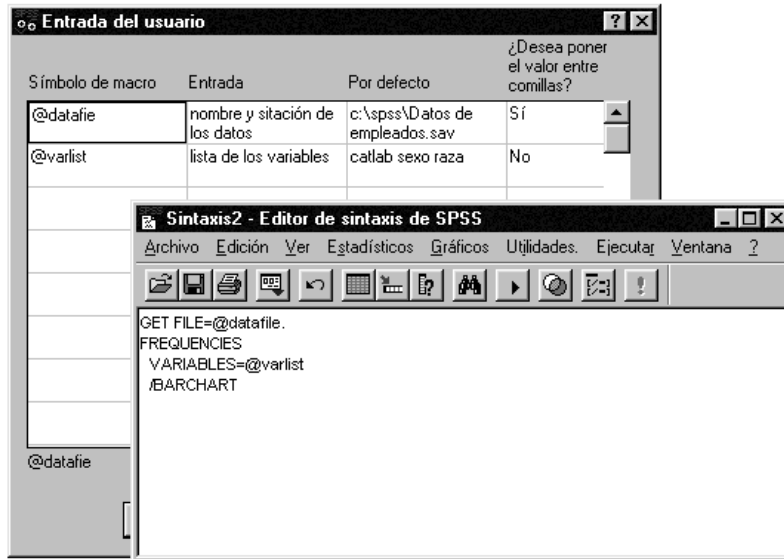
Macro symbol (Símbolo de macro). Nombre de macro utilizado en el archivo de sintaxis de comandos para invocar la macro que solicita información al usuario. El nombre de símbolo de macro debe comenzar por @.

Prompt (Entrada). Etiqueta descriptiva que se muestra cuando el trabajo de producción solicita al usuario que introduzca información. Por ejemplo, podría utilizar la frase “¿Qué archivo de datos desea utilizar?” para identificar un campo que requiera el nombre de un archivo de datos.

Default (Por defecto). Valor que utiliza por defecto el trabajo de producción si no se introduce un valor nuevo. Este valor se muestra cuando el trabajo de producción solicita información al usuario. Puede sustituir o modificar el valor en el momento de la ejecución.

Enclose Value in Quotes (¿Desea poner el valor entre comillas?). Introduzca Y o Yes (Sí) si desea que el valor aparezca entre comillas. En caso contrario, deje el campo en blanco o introduzca N o No. Por ejemplo, para especificar un nombre de archivo debe escribir Yes (Sí) ya que las especificaciones de nombres de archivos deben ir entre comillas.

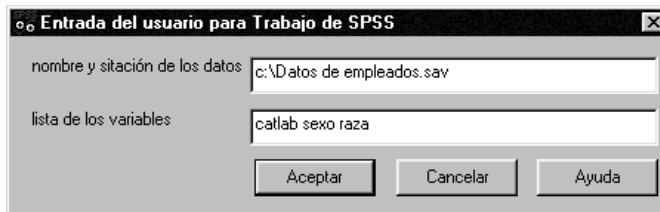
Figura 45-4
 Entradas de macro en un archivo de sintaxis de comandos



Solicitud de entrada del usuario en los trabajos de producción con macros

La Unidad de producción solicita valores al usuario cuando se ejecuta un trabajo de producción que contiene símbolos de macro que han sido definidos. Se pueden sustituir o modificar los valores mostrados por defecto. Posteriormente, estos valores se sustituyen a los símbolos de macro en todos los archivos de sintaxis de comandos asociados al trabajo de producción.

Figura 45-5
 Cuadro de diálogo de solicitud de entrada del usuario para las macros de producción

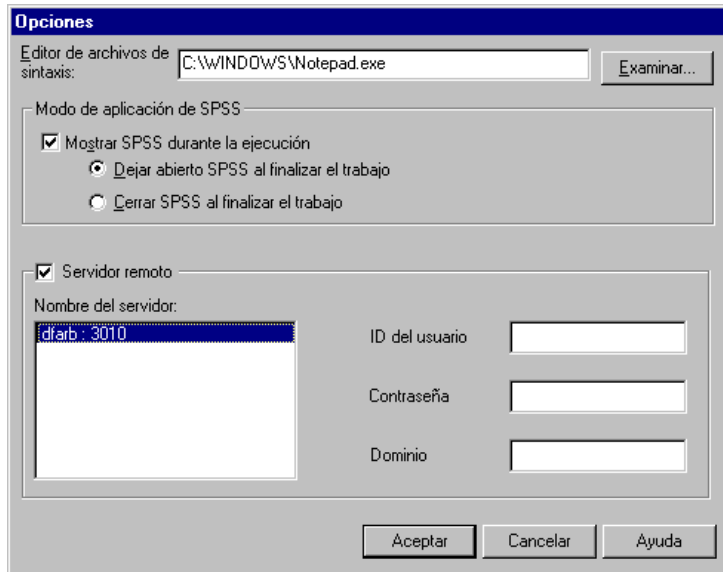


Opciones de producción

Las opciones de producción permiten:

- Especificar un editor de textos por defecto para los archivos de sintaxis a los que se accede mediante el botón Edit (Editar) del cuadro de diálogo principal.
- Ejecutar el trabajo de producción como un proceso invisible en segundo plano o ver los resultados generados a medida que se ejecuta el trabajo.
- Especifique un servidor remoto, un nombre de dominio, un ID de usuario y una contraseña para el análisis en modo distribuido (sólo si tiene acceso de red a la versión de SPSS del servidor). Si no especifica estos ajustes, se utilizarán los ajustes por defecto del cuadro de diálogo Acceso al servidor SPSS. Sólo se pueden seleccionar servidores remotos que se hayan definido previamente en el cuadro de diálogo Añadir servidor en SPSS (menú Archivo, Cambiar servidor, Añadir).

Figura 45-6
Cuadro de diálogo Opciones



Para cambiar las opciones de producción

Elija en los menús de la unidad de producción:

Edición
Opciones...

Control del formato en los trabajos de producción

Una serie de ajustes de SPSS le ayudarán a conseguir el mejor formato posible para las tablas pivote creadas en los trabajos de producción:

Aspectos de tabla. Si edita y guarda Aspectos de tabla (en el menú Formato de la tabla pivote activada), podrá controlar muchos de los atributos de las tablas pivote. Puede especificar tamaños y estilos de fuente, colores y bordes. Para asegurarse de que las tablas anchas no se dividen en varias páginas, seleccione Reducir tabla ancha para caber en la página en la pestaña Impresión del cuadro de diálogo Propiedades de tabla.

Etiquetas de los resultados. Las opciones de Etiquetas de los resultados (menú Edición, Opciones, pestaña Etiquetas de los resultados) controlan la presentación de la información de las variables y los valores de los datos en las tablas pivote. Puede mostrar los nombres de variable o las etiquetas de variable definidas, los valores actuales de los datos o las etiquetas de valor definidas. Generalmente, las etiquetas descriptivas de variable y de valor suelen facilitar la interpretación de los resultados, pero en algunas tablas puede que las etiquetas largas estorben.

Ancho de columna. Las opciones de las tablas pivote (menú Edición, Opciones, pestaña Tablas pivote) determinan el Aspecto de tabla por defecto y el ajuste automático de los anchos de columna en las tablas pivote.

- Sólo las etiquetas ajusta la anchura de la columna a la de su etiqueta. Así se generan tablas más compactas, pero no se mostrarán los valores de los datos más anchos que la etiqueta (los asteriscos indican valores demasiado anchos para ser mostrados).
- Etiquetas y datos ajusta la anchura de la columna al elemento más ancho entre la etiqueta de la columna y el mayor de los valores de los datos. Así se generan tablas más anchas, pero se asegura que se mostrarán todos los valores.

Los trabajos de producción utilizan el Aspecto de tabla actual y las opciones vigentes. Puede establecer el Aspecto de tabla y la configuración de las opciones antes de ejecutar el trabajo de producción, o utilizar los comandos SET dentro de los archivos de sintaxis para controlar estas particularidades. El empleo de los comandos SET en los archivos de sintaxis permite utilizar varios aspectos de tabla y configuraciones de las opciones en un mismo trabajo.

Para crear un aspecto de tabla por defecto personalizado

- ▶ Active una tabla pivote (pulse dos veces en cualquier punto de la tabla).
- ▶ Elija en los menús:
 - Formato
 - Aspectos de tabla...
- ▶ Seleccione un aspecto de tabla en la lista y pulse en Editar aspecto.
- ▶ Ajuste las propiedades de la tabla para los atributos que desee.
- ▶ Pulse en Guardar aspecto o en Guardar como para guardar el aspecto de tabla y, después, en Aceptar.
- ▶ Elija en los menús:
 - Edición
 - Opciones...
- ▶ Pulse en la pestaña Tablas pivote.
- ▶ Elija el aspecto de tabla en la lista y pulse en Aceptar.

Para establecer las opciones de los trabajos de producción

- ▶ Elija en los menús:
 - Edición
 - Opciones...
- ▶ Seleccione las opciones que desee.
- ▶ Pulse en Aceptar.

Puede definir el Aspecto de tabla por defecto, las opciones de las etiquetas de los resultados y el ajuste automático del ancho de columna en el cuadro de diálogo Opciones. Los ajustes realizados en Opciones se guardarán con el programa. Cuando se ejecuta un trabajo de producción, se aplican los ajustes de Opciones vigentes la última vez que se ejecutó el programa.

Control del formato de las tablas pivote con la sintaxis de comandos

SET TLOOK. Controla el Aspecto de tabla por defecto para las nuevas tablas pivote, como en:

- `SET TLOOK = 'c:\trabsprod\mitabla.tlo'`.

SET TVARS. Controla la presentación de los nombres y las etiquetas de variable en las nuevas tablas pivote.

- `SET TVARS = LABELS` muestra las etiquetas de variable.
- `SET TVARS = NAMES` muestra los nombres de las variables.
- `SET TVARS = BOTH` muestra tanto los nombres como las etiquetas de variable.

SET ONUMBER. Controla la presentación de los valores de los datos o las etiquetas de valor en las nuevas tablas pivote.

- `SET ONUMBER = LABELS` muestra las etiquetas de valor.
- `SET ONUMBER = VALUES` muestra los valores de datos.
- `SET ONUMBER = BOTH` muestra los valores de datos y las etiquetas de valor.

SET TFIT. Controla el ajuste automático del ancho de columna para las nuevas tablas pivote.

- `SET TFIT = LABELS` ajusta el ancho de columna al ancho de la etiqueta de la columna.
- `SET TFIT = BOTH` ajusta el ancho de columna al ancho de etiqueta de columna o al mayor valor de los datos, el que sea más ancho.

Ejecución de trabajos de producción desde una línea de comandos

Los modificadores de la línea de comandos permiten programar los trabajos de producción para que se ejecuten cada cierto tiempo con utilidades de programación como la que se encuentra disponible en Microsoft Plus!. Puede ejecutar trabajos de producción desde una línea de comandos con los siguientes modificadores:

- r. Ejecuta el trabajo de producción. Si el trabajo de producción contiene entradas del usuario, debe suministrar la información solicitada antes de ejecutar el trabajo.
- s. Ejecuta el trabajo de producción y suprime cualquier entrada del usuario o advertencia. Se utilizan automáticamente los valores de entrada del usuario por defecto.

Análisis distribuidos. Si dispone de acceso de red a la versión de SPSS en el servidor, también podrá utilizar los siguientes modificadores para ejecutar la unidad de producción en los análisis en modo distribuido:

- x. Nombre o dirección IP del servidor remoto.
- n. Número de puerto.
- d. Nombre del dominio.
- u. ID del usuario para acceder al servidor de remoto.
- p. Contraseña para acceder al servidor remoto.

Si especifica algún modificador de las líneas de comando cuando realiza un análisis distribuido, debe especificar también todos los modificadores de línea de comando en el análisis distribuido (-x, -n, -d, -u y -p).

Debe indicar la ruta completa tanto para la unidad de producción (*spssprod.exe*) como para el trabajo de producción, y ambos deben ir entre comillas, como en:

```
c:\archivos de programa\spss\spssprod.exe c:\spss\datastrabajo\trabprod.spp -s -r
```

En aquellos casos en los que los modificadores de línea de comandos requieren especificaciones adicionales, detrás del modificador debe ir un signo igual y, a continuación, la especificación. Si la especificación contiene espacios (como un nombre de servidor con dos palabras), el valor debe ir entre comillas o apóstrofes, como por ejemplo:

```
-x=HAL 9000 -u=palabra secreta
```

Servidor por defecto. Si dispone de acceso de red a la versión de SPSS en el servidor, el servidor por defecto y la información relacionada (si no se especifica en los modificadores de la línea de comandos) es el servidor especificado por defecto en el cuadro de diálogo Acceso al servidor SPSS. Si no se especificó ningún servidor por defecto, el trabajo se ejecutará en el modo local.

Si desea ejecutar un trabajo de producción en modo local pero el ordenador local no es el servidor por defecto, especifique cadenas entre comillas nulas para todos los modificadores de la línea de comandos en el análisis distribuido, como por ejemplo:

```
c:\archivos de programa\spss\spssprod.exe  
c:\spss\datastrabajo\trabprod.spp -x= -n= -d= -u= -p=
```

Ejecución de varios trabajos de producción

Si utiliza un archivo por lotes (.bat) o una unidad similar para ejecutar varios trabajos de producción, utilice el comando `Start` de Windows con el modificador `/wait` para controlar la ejecución de cada trabajo e impedir así que los trabajos posteriores se inicien antes de que finalice el trabajo anterior como ocurre en:

```
cd \program files\spss  
start /wait spssprod.exe prodjob1.spp -s  
start /wait spssprod.exe prodjob2.spp -s
```

Publicar en Web

La opción Publicar en Web exporta los resultados al servidor web SmartViewer para su publicación. Las tablas e informes publicados en SmartViewer se pueden visualizar y manipular a través de la Web, en tiempo real utilizando un explorador estándar.

- Las tablas pivote se publican como tablas dinámicas que se pueden manipular a través de la Web para obtener diferentes vistas de los datos.
- Los gráficos se publican como archivos gráficos JPEG o PNG.
- Los resultados de texto se publican como formato previo de HTML. (Por defecto, la mayoría de los navegadores Web utilizan un tipo de letra de anchura fija para mostrar el texto con formato.)

Publicar. Permite especificar los resultados que se desea publicar:

- **Documento de salida.** Publica todo el documento de resultados, incluyendo los elementos ocultos o contraídos.
- **Documento de salida (sin notas).** Publica todo excepto la tabla de notas que se genera de forma automática para cada procedimiento.
- **Sólo tablas.** Excluye los gráficos. Se publicarán todas las tablas pivote y todas las tablas de texto.
- **Sólo tablas (sin notas).** Excluya los gráficos y las tablas de notas.
- **Sólo gráficos.** Sólo publica los gráficos del documento.
- **Nada.** Desactiva la publicación en Web. Ya que todos los ajustes se guardan con el trabajo de producción (archivo *.spp*), los resultados se publicarán cada vez que ejecute el trabajo de producción a no ser que seleccione Nada. Esta opción desactiva la publicación pero no la generación de otros tipos de salidas (archivos del Visor, archivos HTML) especificadas en el trabajo de producción.

Publicar tablas como. Controla el modo de publicación de las tablas pivote:

- **Interactivo.** Las tablas son objetos dinámicos que se pueden manipular desde la Web para obtener diferentes vistas de los datos.
- **Estático.** Las tablas son objetos estáticos y no se pueden manipular tras su publicación.

Configurar. Abre la página “Configurar publicación automática” del servidor web SmartViewer en una ventana del navegador. Se requiere cuando se crea un nuevo trabajo de producción para publicar en la Web.

También es necesario una identificación de usuario y una contraseña para acceder al servidor web SmartViewer. Cuando cree un nuevo trabajo de producción para publicar en la Web, se le solicitará una identificación de usuario y una contraseña. Esta información se almacena en el trabajo de producción con un formato encriptado.

Nota: Publicar en web sólo está disponible para sitios que tengan instalado el servidor web SmartViewer y es necesario un complemento para activar la función de publicación. Póngase en contacto con su administrador del sistema o administrador del servidor web para obtener instrucciones sobre la descarga del complemento. Si SmartViewer no está disponible en su sitio web, utilice Exportar resultados para guardar los resultados en formato HTML.

Inicio de sesión del servidor Web de SmartViewer

La publicación en el servidor web SmartViewer requiere un nombre de usuario correcto del servidor web de SmartViewer (identificación de usuario) y una contraseña.

Póngase en contacto con su administrador del sistema o administrador del servidor web para obtener más información.

Utilidad de procesamiento de SPSS

La utilidad de procesamiento permite automatizar tareas, entre las que se incluyen:

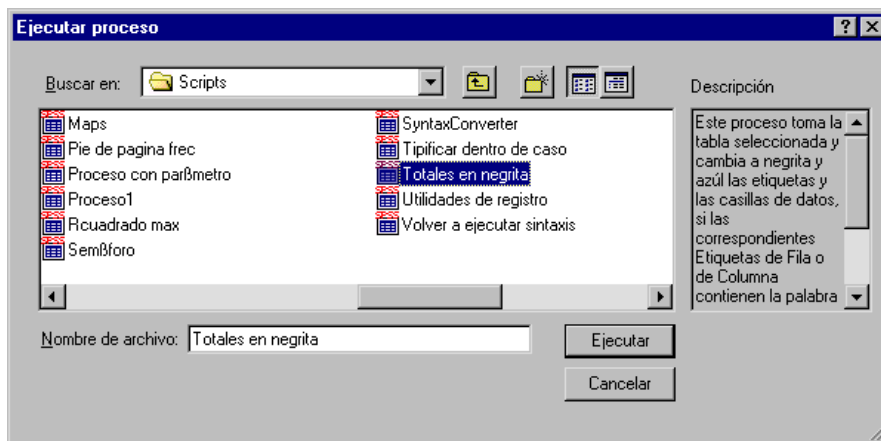
- Personalizar automáticamente los resultados del Visor.
- Abrir y guardar archivos de datos.
- Mostrar los cuadros de diálogo y trabajar con ellos.
- Ejecutar transformaciones de datos y procedimientos estadísticos utilizando la sintaxis de comandos.
- Exportar gráficos como archivos de gráficos en diversos formatos.

Con el programa se incluyen varios procesos, entre ellos algunos autoprocesos que se ejecutan automáticamente siempre que se produce un tipo concreto de resultado. Puede utilizar estos procesos tal y como son o personalizarlos según sus necesidades. Si desea crear sus propios procesos, puede empezar por elegir entre un número de procesos de iniciación.

Para ejecutar un proceso

- ▶ Elija en los menús:
 - Utilidades
 - Ejecutar proceso...

Figura 46-1
Cuadro de diálogo Ejecutar proceso



- ▶ Seleccione la carpeta de procesos *Scripts*.
- ▶ Seleccione el proceso que desee.

Si desea obtener más información, consulte “Personalización de menús y barras de herramientas” en Capítulo 44 en p. 635.

Procesos incluidos en SPSS

El programa incluye los siguientes procesos:

Ejecutar casos no seleccionados. Repite un análisis factorial o discriminante utilizando los casos no seleccionados en un análisis anterior. Antes de ejecutar el proceso se debe seleccionar una tabla Notas producida por una ejecución anterior de Factor o de Discriminante.

Cambiar sig a p. Cambia *Sig.* a $p=$ en las etiquetas de columna de cualquier tabla pivote. Se debe seleccionar la tabla antes de ejecutar el proceso.

Limpiar navegador. Elimina todas las tablas Notas de un documento de resultados. Antes de ejecutar el proceso el documento debe estar abierto en la ventana del Visor designada.

Pie de página frec. Inserta los estadísticos mostrados en una tabla de estadísticos de frecuencias como notas al pie en la tabla de frecuencias correspondiente para cada variable. Antes de ejecutar el proceso se debe seleccionar la tabla de estadísticos de frecuencia.

Totales en negrita. Aplica el formato negrita y el color azul a cualquier fila, columna o capa de datos con la etiqueta *Total* en una tabla pivote. Se debe seleccionar la tabla antes de ejecutar el proceso.

Informe de medias. Extrae información de una tabla Medias y escribe los resultados en varios archivos de resultados ASCII. Antes de ejecutar el proceso se debe seleccionar la tabla Medias.

Eliminar etiquetas. Elimina todas las etiquetas de fila y columna de la tabla pivote seleccionada. Se debe seleccionar la tabla antes de ejecutar el proceso.

Volver a ejecutar sintaxis. Vuelva a enviar el comando encontrado en la tabla Notas seleccionada utilizando el archivo de datos activo. Si no existe ningún archivo de datos abierto, el proceso intentará leer el archivo de datos utilizado originalmente. Se debe seleccionar la tabla Notas antes de ejecutar el proceso.

R cuadrado max. En una tabla de resumen del modelo de regresión, aplica el formato de negrita y el color azul a la fila correspondiente al modelo que maximiza la R cuadrado corregida. Se debe seleccionar la tabla de resumen del modelo antes de ejecutar el proceso.

Si desea obtener más información, consulte “Opciones” en Capítulo 43 en p. 613.

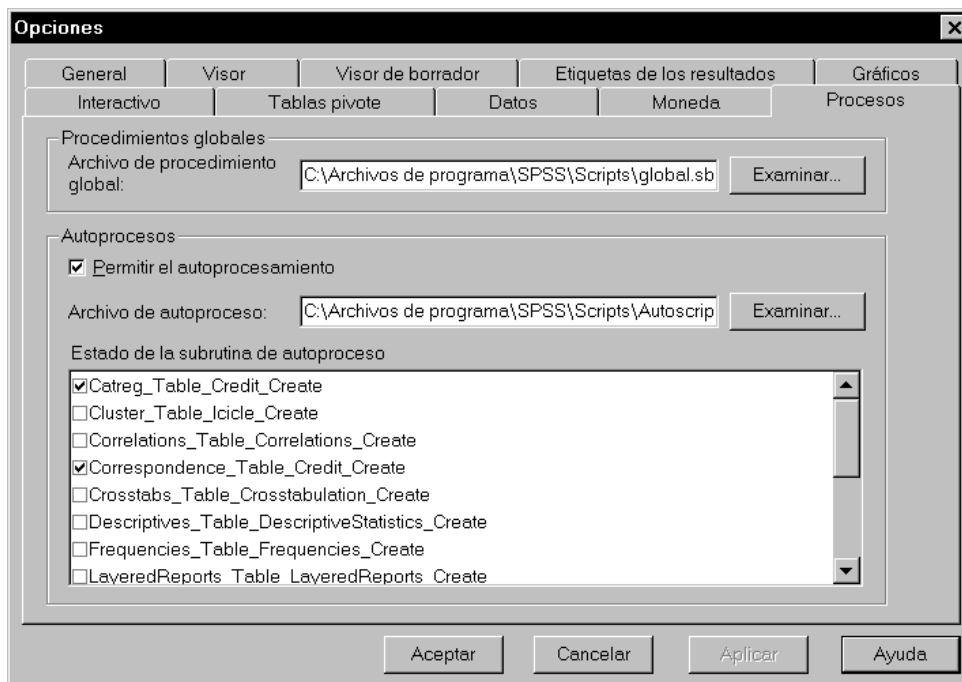
Nota: Esta lista puede estar incompleta.

Autoprocesos

Los autoprosesos se ejecutan automáticamente al ser activados cuando se crea una pieza específica de resultados con un procedimiento dado. Por ejemplo, uno de los autoprosesos elimina automáticamente la diagonal superior y resalta los coeficientes de correlación inferiores a una significación determinada siempre que se genera una tabla de correlaciones mediante el procedimiento Correlaciones bivariadas.

La pestaña Procesos del cuadro de diálogo Opciones (menú Edición) muestra los autoprosesos que están disponibles en el sistema y permite activar o desactivar procesos individuales.

Figura 46-2
Pestaña Procesos del cuadro de diálogo Opciones



Los autoprocesos son específicos para un procedimiento y un tipo de resultados concreto. Un autoproceso que da formato a las tablas de ANOVA producidas por ANOVA de un factor no se activa con las tablas de ANOVA generadas por otros procedimientos estadísticos (aunque puede utilizar procedimientos globales para crear distintos autoprocesos para estas otras tablas que comparten buena parte del mismo código). Sin embargo, puede disponer de un autoproceso diferente para cada tipo de resultado producido por el mismo procedimiento. Por ejemplo, Frecuencias genera tanto una tabla de frecuencias como una tabla de estadísticos y puede tener un autoproceso diferente para cada uno.

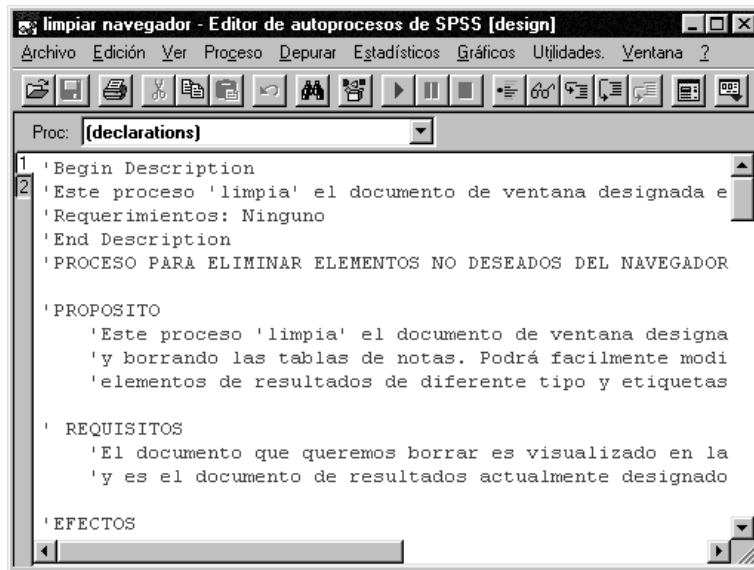
Si desea obtener más información, consulte “Opciones” en Capítulo 43 en p. 613.

Creación y edición de procesos

Puede personalizar muchos de los procesos incluidos con el software para sus necesidades específicas. Por ejemplo, existe un proceso que elimina todas las tablas Notas del documento de resultados designado. Puede modificar fácilmente este proceso para eliminar elementos de resultados de cualquier tipo y etiqueta que desee.

Figura 46-3

Modificación de un proceso en la ventana de proceso

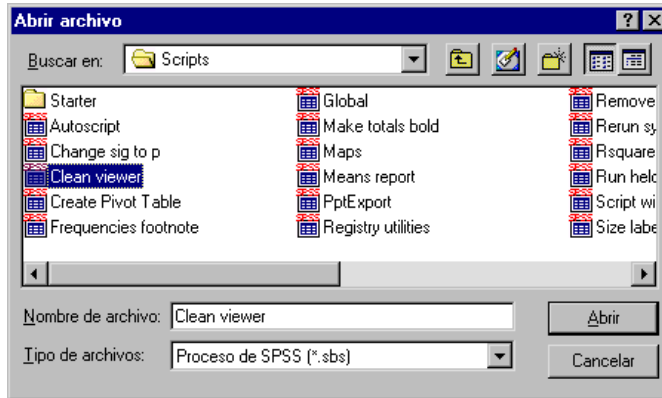


Si prefiere crear sus propios procesos, puede comenzar eligiendo entre varios procesos iniciales.

Para editar un proceso

- ▶ Elija en los menús:
 - Archivo
 - Abrir
 - Proceso...

Figura 46-4
Apertura de un archivo de autoproceto



- ▶ Seleccione la carpeta de procesos *Scripts*.
- ▶ En Tipo de archivos, seleccione Proceso de SPSS (*.sbs).
- ▶ Seleccione el proceso que desee.

Si elige varios procesos, cada uno se abrirá en su propia ventana.

Ventana de proceso

La ventana de proceso es un entorno de programación completo que utiliza el lenguaje Sax BASIC e incluye un editor de cuadros de diálogo, un Visor de objetos, funciones de depuración y ayuda sensible al contexto.

Figura 46-5
Ventana de proceso

```

1 Sub SelectAndRemoveOutputItem(intType As Integer, Optional
2 'This procedure iterates through output items and deletes
  'items that match the specified search type and label.

  'Variable declarations

  Dim objOutputDoc As ISpssOutputDoc
  Dim objItems As ISpssItems
  Dim objItem As ISpssItem
  'By convention, object variable names begin With "obj
  'OutputDoc, ISpssItems, And ISpssItem are the names o
  'For example, the first declaration above declares an
  '"objOutputDoc" and assigns it to the OutputDoc Objec
  'class. Below, that variable is set to the designated
  'in order to access the items in that output document

```

- Según mueva el cursor, el nombre del procedimiento actual aparecerá en la parte superior de la ventana.
- Los términos en color azul son palabras reservadas en BASIC (por ejemplo, Sub, End Sub y Dim). Puede acceder a la ayuda sensible al contexto para estos términos pulsando en ellos y, a continuación, en F1.
- Los términos en color magenta son objetos, propiedades o métodos de SPSS. También puede pulsar en ellos y, después, en F1 para obtener ayuda, pero sólo cuando aparezcan en instrucciones válidas y en color magenta. Si pulsa en el nombre de un objeto en un comentario, accederá a la ayuda sobre el lenguaje Sax BASIC en lugar de a la ayuda sobre los objetos de SPSS.
- Los comentarios se muestran en color verde.
- Pulse F2 en cualquier momento para abrir el Visor de objetos, que muestra objetos, propiedades y métodos de SPSS.

Propiedades del Editor de procesos (Ventana de proceso)

Los elementos de código que aparecen en la ventana de proceso están codificados por colores para que sea más fácil distinguirlos. Por defecto, los comentarios están en verde, los términos Sax BASIC en azul y los nombres de los objetos, las propiedades y los métodos válidos en magenta. Puede especificar otros colores para estos elementos y cambiar el tamaño y la fuente de todo el texto.

Para definir las propiedades del Editor de procesos

- ▶ Elija en los menús:
Proceso
Propiedades del editor...

Figura 46-6

Cuadro de diálogo Propiedades del editor

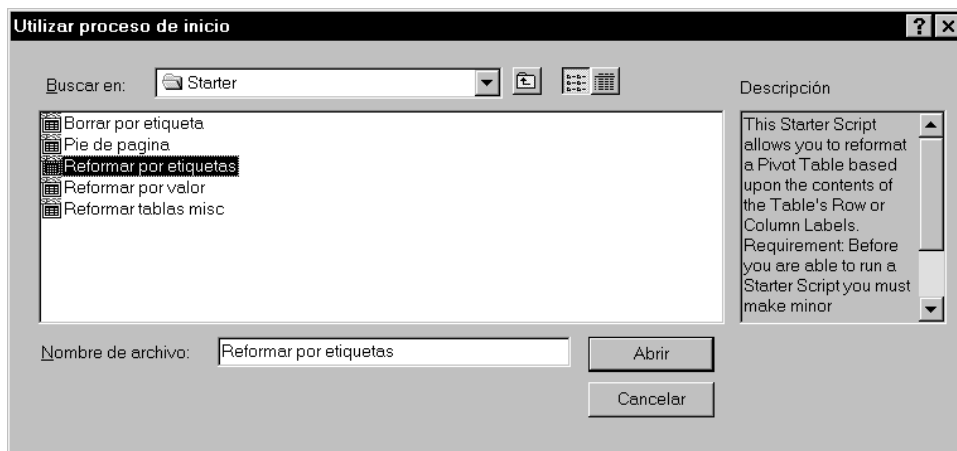


- ▶ Para modificar el color de un tipo de elemento de código, seleccione el elemento y elija un color en la paleta desplegable.

Procesos de iniciación

Si desea crear un proceso nuevo, puede comenzar eligiendo uno de los diversos procesos de iniciación que se ofrecen.

Figura 46-7
Cuadro de diálogo Utilizar proceso de iniciación



Cada proceso de iniciación proporciona el código para uno o varios procedimientos comunes, e incluye comentarios con consejos sobre cómo personalizar el proceso según sus necesidades particulares.

Borrar por etiqueta. Elimina filas o columnas de una tabla pivote en función del contenido de RowLabels o ColumnLabels. Para que funcione este proceso, la opción Ocultar filas y columnas vacías del cuadro de diálogo Propiedades de tabla debe estar seleccionada.

Borrar elementos del navegador. Elimina elementos del Visor basándose en diferentes criterios.

Pie de página. Vuelve a dar formato a una nota al pie de una tabla pivote, modifica el texto de una nota al pie o añade una nota al pie.

Reformar por etiquetas. Vuelve a dar formato a una tabla pivote basándose en las etiquetas de fila, de columna o de capa.

Reformar por valor. Vuelve a dar formato a una tabla pivote basándose en el valor de las casillas de datos o en una combinación de casillas de datos y etiquetas.

Reformar tablas misc. Vuelve a dar formato o modifica el texto de un título de la tabla pivote, el texto de esquina o el texto al pie.

Además, puede utilizar cualquiera de los otros procesos disponibles como procesos de iniciación, aunque quizás no sea tan fácil personalizarlos. Sólo tiene que abrir el proceso y guardarlo con un nombre de archivo diferente.

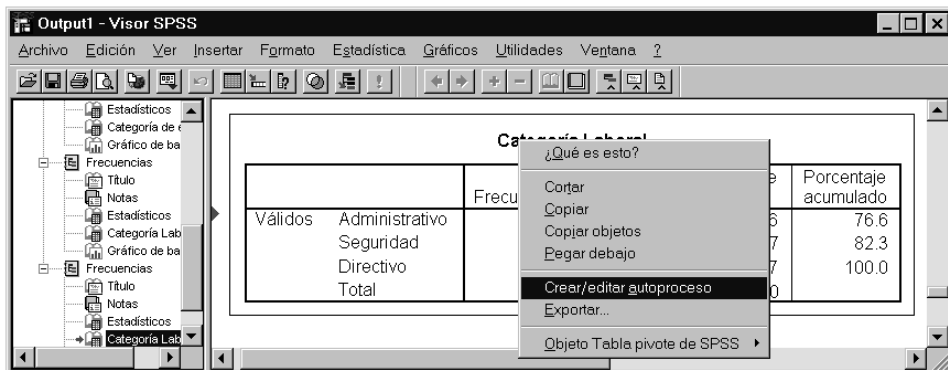
Para crear un proceso

- ▶ Elija en los menús:
Nuevo
Proceso...
- ▶ Seleccione un proceso de iniciación si desea comenzar con uno.
- ▶ Si no desea utilizar un proceso de inicio, pulse en Cancelar.

Creación de autoprosesos

La creación de un autoprosceso empieza por el objeto de resultados que se desea utilizar como desencadenante. Por ejemplo, para crear un autoprosceso que se ejecute siempre que se produzca una tabla de frecuencias, genere una tabla de frecuencias de la manera habitual y pulse en la tabla en el Visor para seleccionarla. A continuación, pulse con el botón derecho del ratón o sírvase del menú Utilidades para crear un nuevo autoprosceso que se active al generar ese tipo de tabla.

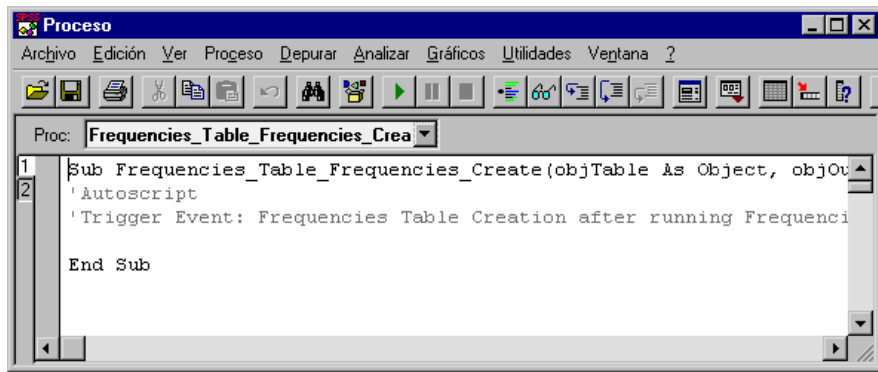
Figura 46-8
Creación de un nuevo autoprosceso



Por defecto, cada autoproceso que se crea se añade al archivo actual de autoprocesos (*autoscript.sbs*) como un nuevo procedimiento. El nombre del procedimiento hace referencia al evento que sirve como desencadenante. Por ejemplo, si crea un autoproceso que se activa cuando Explorar crea una tabla Descriptivos, el nombre de la subrutina de autoproceso sería `Explore_Table_Descriptives_Create`.

Figura 46-9

Nuevo procedimiento de autoproceso en la ventana de proceso



Esto facilita el desarrollo de los autoprocesos porque no es necesario escribir el código para obtener el objeto en el que se quiere operar, pero requiere que los autoprocesos sean específicos para una pieza dada de procedimiento estadístico y de resultados.

Para crear un autoproceso

- ▶ Seleccione el objeto que desee que sirva como desencadenante en el Visor.
- ▶ Elija en los menús:
Utilidades
Crear/editar autoproceso...

Si no existe ningún autoproceso para el objeto seleccionado, se creará uno nuevo. Si existe uno, se mostrará en la pantalla.

- ▶ Escriba el código.
- ▶ En el menú Edición, seleccione Opciones para activar o desactivar el autoproceso.

Eventos que desencadenan autoprosesos

El nombre del procedimiento de autoproseso hace referencia al evento que sirve como desencadenante. Los siguientes eventos pueden desencadenar autoprosesos:

Creación de una tabla pivote. El nombre del procedimiento hace referencia tanto al tipo de tabla como al procedimiento que la creó; por ejemplo, `Correlations_Table_Correlations_Create`.

Figura 46-10

Procedimiento de autoproseso para una tabla de correlaciones

```

Proc: Crosstabs_Table_Crosstabulation_Create
1 Sub Crosstabs_Table_Crosstabulation_Create(objTable As Object, objOutput
2 'Autoscript
'Trigger Event: Crosstabulation Table Creation after running Crosstabs p
'Effects: Goes through the Row and Column Labels and finds "Total" rows/c
' and turns "Total" and associated data cells bold
' Dim bolSelection As Boolean
Call SelectRowLabelsAndData(objTable, cTOTAL, bolSelection)
Call SelectColLabelsAndData(objTable, cTOTAL, bolSelection)
If bolSelection = True Then
objTable.TextStyle = 2 'make text bold
End If
'Set flag so that screen is not updated while we make changes
'NOTE:THIS MUST BE SET BACK TO TRUE AT THE END OF THE AUTOSCRIPT
objTable.UpdateScreen = False
Call ChangeToPercent(objTable) 'Subroutine that changes names of lab
'Row %, Column%, and Total%
'Set flag so that screen is updated so we can view changes
objTable.UpdateScreen = True
End Sub

```

Creación de título. Hace referencia al procedimiento estadístico que lo creó: `Correlations_Title_Create`.

Creación de notas. Hace referencia al procedimiento que lo creó:

`Correlations_Notes_Create.`

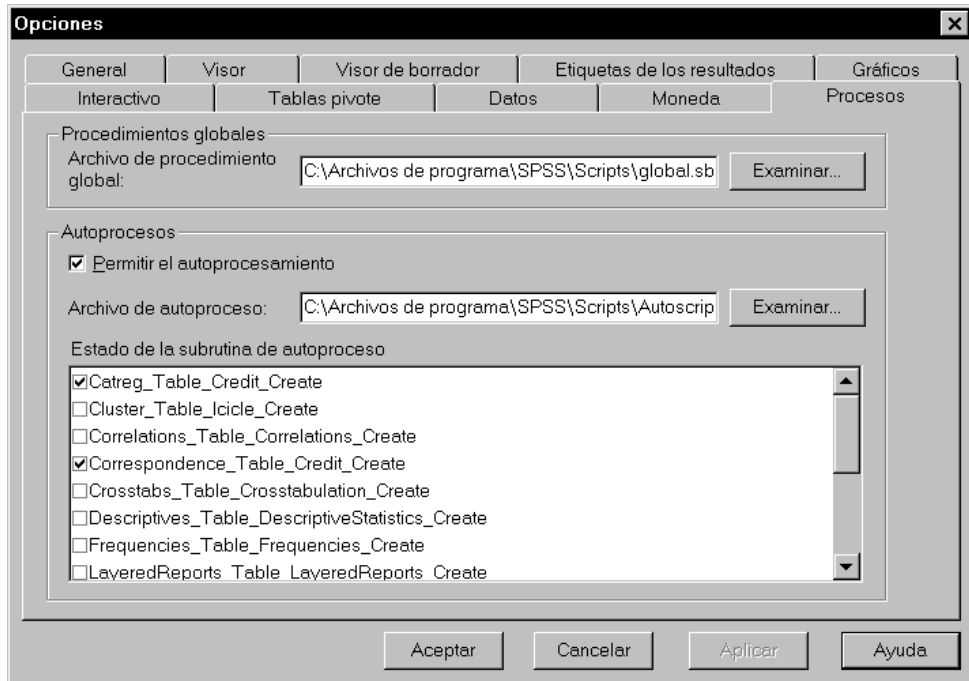
Creación de advertencias. Hace referencia al procedimiento que lo creó.

También puede utilizar un proceso para desencadenar un autoproceso de forma indirecta. Por ejemplo, puede escribir un proceso que invoque al procedimiento Correlaciones, lo que desencadena a su vez el autoproceso registrado en la tabla de correlaciones resultante.

Archivo de autoproceso

Todos los autoprocesos se guardan en un único archivo (a diferencia de otros procesos que se guardan en archivos diferentes). También se añade a este archivo cualquier autoproceso nuevo que se crea. El nombre del archivo de autoproceso actual se muestra en la pestaña Procesos del cuadro de diálogo Opciones (menú Edición).

Figura 46-11
Subrutinas de autoproceso mostradas en el cuadro de diálogo Opciones



El cuadro de diálogo Opciones también muestra todos los autoprocesos del archivo de autoproceso actualmente seleccionado, permitiéndole activar o desactivar procesos individuales.

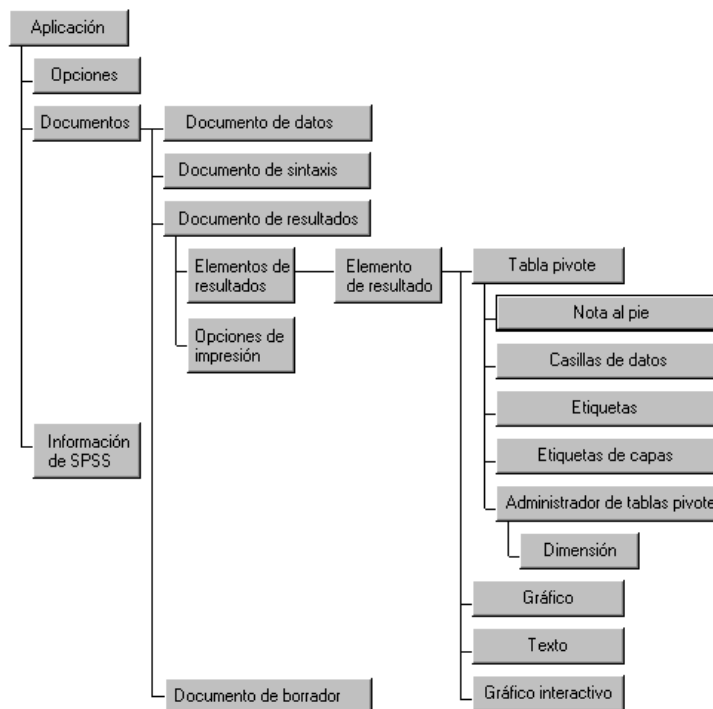
El archivo de autoproceso por defecto es *autscript.sbs*. Puede especificar un archivo de autoproceso diferente, pero sólo puede hacer uno activo.

Funcionamiento de los procesos

Los procesos funcionan gestionando los objetos mediante propiedades y métodos. Por ejemplo, las tablas pivote son una clase de objetos. Con los objetos de esta clase, puede utilizar el método `SelectTable` para seleccionar todos los elementos de la tabla y emplear la propiedad `TextColor` para cambiar el color del texto seleccionado. Además cada una de estas clases posee propiedades y métodos específicos asociados

a ella. El conjunto de todas las clases (o tipos) de objetos se denomina la biblioteca de tipos de SPSS.

Figura 46-12
Presentación en árbol de la jerarquía de objetos



La utilización de objetos es un proceso que consta de dos fases. En primer lugar, se crea una referencia al objeto (que se denomina *obtención* del objeto). Después, se utilizan propiedades y métodos para realizar algo. Se obtienen objetos explorando la jerarquía de objetos y utilizando en cada paso propiedades o métodos de los objetos superiores en la jerarquía para obtener los objetos ubicados por debajo. Por ejemplo, para obtener un objeto de tabla pivote, primero debe obtener el documento de resultados que contiene la tabla pivote y, después, los elementos de ese documento de resultados.

Cada objeto que obtiene se almacena en una variable. (Recuerde que lo que realmente almacena en la variable es una referencia al objeto). A menudo, uno de los primeros pasos en la creación de un proceso consiste en declarar variables para los objetos que necesite.

Sugerencia: Es difícil entender cómo funcionan los procesos si no se conoce el funcionamiento del programa. Antes de escribir un proceso, utilice el ratón para realizar la tarea varias veces de la forma habitual. En cada paso, considere qué objetos está utilizando y qué propiedades de cada objeto va a cambiar.

Declaración de variables

Aunque no siempre es necesario, es una buena idea declarar todas las variables antes de utilizarlas. Esto se realiza principalmente empleando instrucciones de declaración Dim:

```
Dim objOutputDoc As ISpssOutputDoc
Dim objPivotTable As PivotTable
Dim intType As Integer
Dim strLabel As String
```

Cada declaración especifica el nombre y el tipo de variable. Por ejemplo, la primera declaración de arriba crea una variable de objeto que se llama `objOutputDoc` y la asigna a la clase de objeto `ISpssOutputDoc`. La variable aún no tiene un valor porque no se ha establecido para un documento de resultados determinado. Lo que hace la instrucción es declarar que la variable existe. (A este proceso se le ha denominado “renombrar los objetos que desea utilizar”.)

Convenciones de denominación de variables. Por convención, el nombre de cada variable indica su tipo. Los nombres de las variables de objeto empiezan por `obj`, las variables de enteros por `int` y las variables de cadena por `str`. Se trata de simples convenciones (puede dar a sus variables el nombre que desee), pero si las sigue, resultará más fácil entender su código.

Clases de objetos de SPSS. `ISpssOutputDoc` y `PivotTable` son nombres de clases de objetos de SPSS. Cada clase representa un tipo de objeto que SPSS puede crear, tales como un documento de resultados o una tabla pivote. Además cada una de estas

clases posee propiedades y métodos específicos asociados a ella. El conjunto de todas las clases (o tipos) de objetos de SPSS se conoce como biblioteca de tipos de SPSS.

Clases de objetos de SPSS y convenciones de denominación de variables

Los siguientes nombres de variables se utilizan en los procesos de muestra incluidos en el programa y se recomiendan para todos los procesos. Observe que a excepción de las tablas pivote, las clases de objetos poseen nombres que comienzan con `ISpss`.

Objeto	Tipo o Clase	Nombre de variable
Aplicación de SPSS	<code>ISpssApp</code>	<code>objSpssApp</code> , la variable es global y no requiere declaración
Opciones de SPSS	<code>ISpssOptions</code>	<code>objSpssOptions</code>
Información de archivo de SPSS	<code>ISpssInfo</code>	<code>objSpssInfo</code>
Documentos	<code>ISpssDocuments</code>	<code>objDocuments</code>
Documento de datos	<code>ISpssDataDoc</code>	<code>objDataDoc</code>
Documento de sintaxis	<code>ISpssSyntaxDoc</code>	<code>objSyntaxDoc</code>
Documento del Visor	<code>ISpssOutputDoc</code>	<code>objOutputDoc</code>
Opciones de impresión	<code>ISpssPrintOptions</code>	<code>objPrintOptions</code>
Conjunto de elementos de resultados	<code>ISpssItems</code>	<code>objOutputItems</code>
Elemento de resultados	<code>ISpssItem</code>	<code>objOutputItem</code>
Gráficos	<code>ISpssChart</code>	<code>objSPSSChart</code>
Texto	<code>ISpssRtf</code>	<code>objSPSSText</code>
Tabla pivote	<code>PivotTable</code>	<code>objPivotTable</code>
Notas al pie	<code>ISpssFootnotes</code>	<code>objFootnotes</code>
Casillas de datos	<code>ISpssDataCells</code>	<code>objDataCells</code>
Etiquetas de capa	<code>ISpssLayerLabels</code>	<code>objLayerLabels</code>
Etiquetas de columna	<code>ISpssLabels</code>	<code>objColumnLabels</code>
Etiquetas de fila	<code>ISpssLabels</code>	<code>objRowLabels</code>

Objeto	Tipo o Clase	Nombre de variable
Administrador de tabla pivote	ISpssPivotMgr	objPivotMgr
Dimensión	ISpssDimension	objDimension

Obtención de objetos de automatización (Proceso)

Obtener un objeto significa crear una referencia al objeto de forma que se puedan utilizar propiedades y métodos para realizar algo. Cada referencia a un objeto que se obtiene se almacena en una variable. Para obtener un objeto, en primer lugar declare una variable de objeto de la clase apropiada; después defina la variable para el objeto específico. Por ejemplo, para obtener el documento de resultados designado:

```
Dim objOutputDoc As ISpssOutputDoc
Set objOutputDoc = objSpssApp.GetDesignatedOutputDoc
```

Se utilizan propiedades y métodos de objetos superiores en la jerarquía de objetos para llegar hasta los objetos de debajo. La segunda de las instrucciones anteriores obtiene el documento de resultados designado utilizando `GetDesignatedOutputDoc`, un método asociado al objeto de aplicación, que es el objeto del nivel más alto. De modo similar, para obtener un objeto de tabla pivote, en primer lugar se obtiene el documento de resultados que contiene la tabla pivote, después el conjunto de elementos de ese documento y así sucesivamente.

Ejemplo: Obtención de un elemento de resultados

Este proceso obtiene el tercer elemento de resultados del documento de resultados designado y lo activa. Si ese elemento no es un objeto OLE, el proceso genera un error.

```
Sub Main
```

```
Dim objOutputDoc As ISpssOutputDoc'Declara las variables del objeto
Dim objOutputItems As ISpssItems
Dim objOutputItem As ISpssItem
```

```

Set objOutputDoc = objSpssApp.GetDesignatedOutputDoc 'Obtiene referencia para el
documento de resultados designado
Set objOutputItems = objOutputDoc.Items() 'Obtiene la colección de elementos
del documento
Set objOutputItem = objOutputItems.GetItem(2) 'Obtiene el tercer elemento de
resultados '(los números de elemento empiezan en 0, así que 2 corresponde al tercero)

objOutputItem.Activate 'Activa el elemento de resultados

End sub

```

Ejemplo: Obtención de la primera tabla pivote

Este proceso obtiene la primera tabla pivote en el documento de resultados designado y lo activa.

```

Sub Main

Dim objOutputDoc As ISpssOutputDoc 'declare object variables
Dim objOutputItems As ISpssItems
Dim objOutputItem As ISpssItem Dim objPivotTable As PivotTable

Set objOutputDoc = objSpssApp.GetDesignatedOutputDoc 'Obtiene referencia
para el documento de resultados designado
Set objOutputItems = objOutputDoc.Items() 'Obtiene la colección de elementos del documento

Dim intItemCount As Integer 'número de elementos de resultados
Dim intItemType As Integer 'tipo de elemento (definido por la propiedad SpssType)

intItemCount = objOutputItems.Count() 'obtiene el número de elementos de resultados
For index = 0 To intItemCount 'bucle para recorrer los elementos de resultados
Set objOutputItem = objOutputItems.GetItem(index) 'obtiene el elemento actual
intItemType = objOutputItem.SPSSType() 'obtiene el tipo del elemento actual
If intItemType = SPSSPivot Then
Set objPivotTable = objOutputItem.Activate() 'si el elemento es una tabla pivote, actívela
Exit For
End If
Next index

End sub

```

También dispone de ejemplos en la ayuda en pantalla. Puede probarlos usted mismo pegando el código de la Ayuda en la ventana de proceso.

Propiedades y métodos (Proceso)

Los objetos de automatización OLE poseen características y usos, como si fueran objetos reales. En terminología de programación, a las características se las denomina propiedades y a los usos, métodos. Cada clase de objetos posee métodos y propiedades específicos que determinan lo que se puede realizar con ese objeto.

Objeto	Propiedad	Método
lápiz (mundo real)	Dureza Color	Escribir Borrar
Tabla pivote (SPSS)	TextFont DataCellWidths CaptionText	SelectTable ClearSelection HideFootnotes

Ejemplo: Uso de propiedades (Proceso)

Las propiedades definen o devuelven atributos de los objetos, tales como el color o el ancho de la casilla. Cuando una propiedad aparece en el lado izquierdo de un signo igual, está escribiendo en ella. Por ejemplo, para definir el texto al pie de una tabla pivote activada (`objPivotTable`) en "los resultados de Anita":

```
objPivotTable.CaptionText = "los resultados de Anita"
```

Cuando una propiedad aparece en el lado derecho, está leyéndola. Por ejemplo, para obtener el texto al pie de la tabla pivote activada y guardarlo en una variable:

```
strFontName = objPivotTable.CaptionText
```


Ejemplo: Uso de métodos (Proceso)

Los métodos llevan a cabo acciones sobre los objetos, tales como la selección de todos los elementos de una tabla:

```
objPivotTable.SelectTable
```

O la eliminación de una selección:

```
objPivotTable.ClearSelection
```

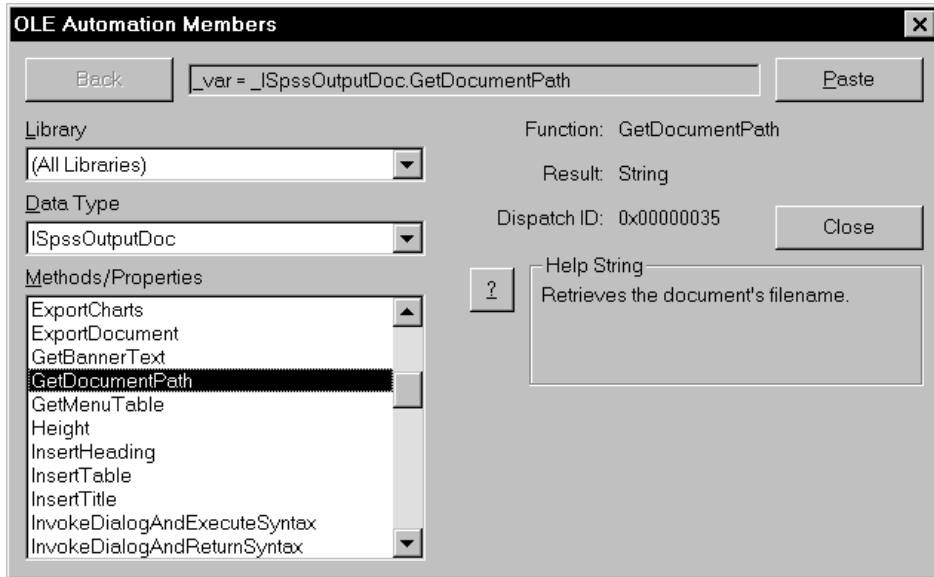
Algunos métodos devuelven otro objeto. Dichos métodos son esenciales para desplazarse por la jerarquía de objetos. Por ejemplo, el método `GetDesignatedOutputDoc` devuelve el documento de resultados designado, permitiendo acceder a los elementos de ese documento:

```
Set objOutputDoc = objSpssApp.GetDesignatedOutputDoc  
Set objItems = objOutputDoc.Items
```

Visor de objetos

El Visor de objetos muestra todas las clases de objetos y los métodos y propiedades asociados a cada una. También puede acceder a la ayuda sobre propiedades y métodos específicos y pegar los que seleccione en su proceso.

Figura 46-13
Visor de objetos



Para utilizar el Visor de objetos

- ▶ En los menús de la ventana de proceso elija:
Depurar
Visor de objetos...
- ▶ Seleccione una clase de objeto de la lista Tipo de datos para mostrar los métodos y propiedades para esa clase.
- ▶ Seleccione propiedades y métodos para obtener ayuda sensible al contexto o para pegarlos en su proceso.

Procedimiento nuevo (Proceso)

Un procedimiento es una secuencia nombrada de instrucciones que se ejecuta como una unidad. La organización del código en los procedimientos hace más fácil administrar y reutilizar las piezas de código. Los procesos deben poseer al menos un

procedimiento (la subrutina `Main`) y a menudo poseen varios. El procedimiento `Main` puede contener pocas instrucciones, aparte de llamadas a subrutinas que realizan la mayor parte del trabajo.

Figura 46-14
Cuadro de diálogo *Procedimiento nuevo*



Los procedimientos pueden ser subrutinas o funciones. Un procedimiento comienza con una instrucción que especifica el tipo de procedimiento y el nombre (por ejemplo, `Sub Main` o `Function DialogMonitor()`) y concluye con la instrucción `End` (`End Sub` o `End Function`) adecuada.

Según se desplaza por la ventana de proceso, el nombre del procedimiento actual se muestra en la parte superior de dicha ventana. Dentro de un proceso, puede invocar cualquier procedimiento tantas veces como desee. También puede invocar cualquier procedimiento del archivo global de procesos, lo cual hace posible compartir procedimientos entre los procesos.

Para añadir un nuevo procedimiento en un proceso

- ▶ Elija en los menús:
 - Proceso
 - Nuevo procedimiento...
- ▶ Escriba un nombre para el procedimiento.
- ▶ Seleccione `Sub` o `Función`.

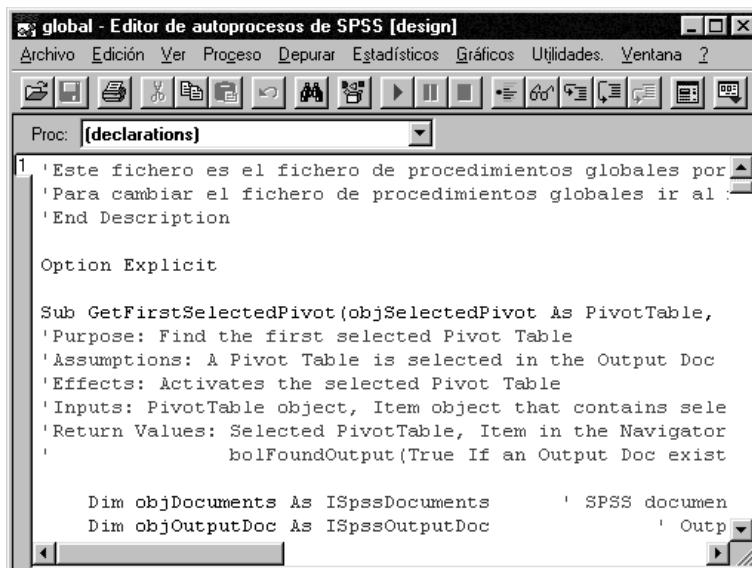
Si lo desea, puede crear un nuevo procedimiento escribiendo las instrucciones que lo definen directamente en el proceso.

Procedimientos globales (Proceso)

Si posee un procedimiento o una función que desea utilizar en varios procesos diferentes, puede añadirlo al archivo global de procesos. Los procedimientos de este archivo pueden invocarse desde todos los procesos restantes.

Figura 46-15

Archivo global de procesos



El archivo global de procesos por defecto es *global.sbs*. Puede añadir los procedimientos que desee a este archivo. También puede especificar otro archivo global en la pestaña Procesos del cuadro de diálogo Opciones (menú Edición), pero sólo puede haber un archivo activo como archivo global en un momento dado. Esto significa que si crea un nuevo archivo global y lo define como el archivo global, los procedimientos y funciones de *global.sbs* ya no estarán disponibles.

Puede ver el archivo global de procesos en cualquier ventana de proceso (pulse en la pestaña 2 en el lateral izquierdo de la ventana justo debajo de la barra de herramientas), pero sólo puede editarlo en una ventana cada vez.

Los procedimientos globales deben ser invocados por otros procedimientos de proceso. No se pueden ejecutar procesos globales directamente desde el menú Utilidades ni desde una ventana de proceso.

Adición de una descripción a un proceso

Puede añadir una descripción para que figure en los cuadros de diálogo Ejecutar proceso y Utilizar proceso de iniciación. Simplemente añada un comentario en la primera línea del proceso que comience con `Begin Description`, seguido del comentario deseado (una o más líneas), seguido de `End Description`. Por ejemplo:

```
'Begin Description  
'Este proceso cambia Sig. a p= en las etiquetas de columna de cualquier tabla pivote.  
'Requisito: Deberá seleccionar la tabla pivote que desea cambiar.  
'End Description
```

La descripción debe poseer formato de comentario, es decir, cada línea debe comenzar con un apóstrofo.

Procesamiento de cuadros de diálogo personalizados

Existen dos pasos para aplicar un cuadro de diálogo personalizado: primero se crea el cuadro de diálogo utilizando el Editor de cuadros de diálogo personalizados (UserDialog Editor) y después se crea una función de control del cuadro de diálogo (DialogFunc) que lo controla y define su funcionamiento.

El cuadro de diálogo propiamente dicho queda definido por un bloque del tipo `Begin Dialog...End Dialog`. No será necesario que escriba este código directamente, ya que el Editor de cuadros de diálogo personalizados proporciona una manera fácil y gráfica de definir el cuadro de diálogo.

Figura 46-16

Creación de un cuadro de diálogo en el Editor de cuadros de diálogo personalizados



El editor muestra al principio un cuadro de diálogo en blanco. Puede añadir controles, tales como botones de radio y casillas de verificación, seleccionando la herramienta apropiada y arrastrando con el ratón. (Sitúe el ratón sobre cada herramienta para obtener su descripción.) También puede arrastrar los lados y las esquinas para cambiar el tamaño del cuadro de diálogo. Después de añadir un control, pulse en él con el botón derecho para definir sus propiedades.

Función de control del cuadro de diálogo. Para crear esta función, pulse con el botón derecho del ratón en el cuadro de diálogo (asegúrese de que no hay ningún control seleccionado en él) y escriba un nombre para la función en el campo correspondiente (Dialog Function). Las instrucciones que definen la función se añadirán a su proceso, aunque tendrá que editar manualmente la función para definir el funcionamiento de cada acción.

Cuando termine, pulse en el icono de guardar y salir (Save and Exit), en el extremo derecho de la barra de herramientas, para añadir el código para el cuadro de diálogo a su proceso.

Para crear un cuadro de diálogo personalizado

- ▶ En la ventana de proceso, pulse en el proceso donde desee insertar el código para el cuadro de diálogo.

- ▶ Elija en los menús:
Proceso
Editor de cuadros de diálogo...
- ▶ Seleccione herramientas de la paleta y pulse en el nuevo cuadro de diálogo para añadir los controles, tales como botones y casillas de verificación (arrástrelos hasta la posición que desee).
- ▶ Modifique el tamaño del cuadro de diálogo arrastrando los tiradores de los lados y las esquinas.
- ▶ Pulse con el botón derecho en el cuadro (sin ningún control seleccionado) y escriba un nombre para la función de control del cuadro de diálogo en el campo correspondiente (Dialog Function).
- ▶ Cuando termine, pulse en el icono de guardar y salir (en el extremo derecho de la barra de herramientas).

Es necesario editar manualmente la función de control del cuadro de diálogo para definir el funcionamiento del cuadro de diálogo.

Funciones de control del cuadro de diálogo (Proceso)

Una función de control de cuadro de diálogo define el funcionamiento de un cuadro de diálogo para cada uno de los casos especificados. La función adopta la siguiente forma (genérica):

```
Function DialogFunc(strDlgItem as String, intAction as Integer, intSuppValue as Integer)
  Select Case intAction
    Case 1 ' inicialización del cuadro de diálogo
      ... 'instrucciones que deben ejecutarse cuando se inicializa
      este cuadro de diálogo
    Case 2 ' cambia el valor o pulsa en el botón
      ... 'instrucciones...
    Case 3 ' ha cambiado el texto de TextBox o ComboBox ...
    Case 4 ' ha cambiado el foco ...
    Case 5 ' inactivo ...
  End Select
End Function
```

Parámetros. La función debe ser capaz de cumplir tres parámetros: Una cadena (`strDlgItem`) y dos enteros (`intAction` e `intSuppValue`). Los parámetros son valores pasados entre la función y el cuadro de diálogo, dependiendo de la acción que se elija.

Por ejemplo, cuando un usuario pulsa en un control del cuadro de diálogo, el nombre del control se pasa a la función como `strDlgItem` (el nombre del campo se especifica en la definición del cuadro de diálogo). El segundo parámetro, `intAction`, es un valor numérico que indica qué acción tuvo lugar en el cuadro de diálogo. El tercer parámetro se utiliza para ofrecer información adicional en algunos casos. Debe incluir los tres parámetros en la definición de la función incluso aunque no vaya a utilizarlos todos.

Select Case intAction. El valor de `intAction` indica la acción que tuvo lugar en el cuadro de diálogo. Por ejemplo, cuando el cuadro de diálogo se abre, `intAction` = 1. Si el usuario pulsa en un botón, `intAction` cambia a 2, etc. Existen cinco acciones posibles y se pueden especificar instrucciones que se ejecuten para cada acción del modo abajo indicado. No es necesario especificar los cinco casos posibles, sólo los aplicables. Por ejemplo, si no desea que se ejecute ninguna instrucción para ejecutarla en la inicialización, omite Case 1.

- **Case intAction = 1.** Especifica las instrucciones que deben ejecutarse cuando se inicializa el cuadro de diálogo. Por ejemplo, podría desactivar uno o varios controles o añadir una señal sonora. La cadena `strDlgItem` es una cadena nula; `intSuppValue` es 0.
- **Case 2.** Se ejecuta al pulsar en un botón o cuando un valor cambia en un control `CheckBox`, `DropListBox`, `ListBox` u `OptionGroup`. Si se pulsa en un botón, `strDlgItem` es el botón, `intSuppValue` no posee significado y debe definir `DialogFunc = True` para evitar que el cuadro de diálogo se cierre. Si un valor cambia, `strDlgItem` es el elemento cuyo valor ha cambiado e `intSuppValue` es el nuevo valor.
- **Case 3.** Se ejecuta cuando un valor cambia en un control `TextBox` o `ComboBox`. La cadena `strDlgItem` es el control cuyo texto cambió y está perdiendo foco; `intSuppValue` es el número de caracteres.

- **Case 4.** Se ejecuta cuando el foco cambia en el cuadro de diálogo. La cadena `strDlgItem` gana foco e `intSuppValue` es el elemento que pierde foco (el primer elemento es 0, el segundo es 1, y así sucesivamente).
- **Case 5.** Procesamiento inactivo. La cadena `strDlgItem` es una cadena nula; `intSuppValue` es 0. Defina `DialogFunc = True` para continuar recibiendo acciones inactivas.

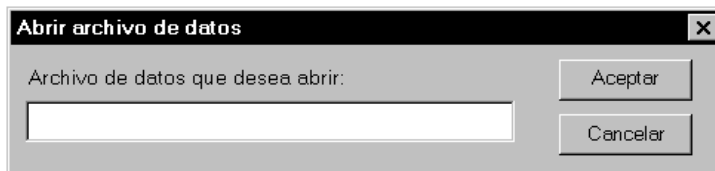
Si desea obtener más información, consulte los ejemplos y el prototipo `DialogFunc` en el archivo de ayuda de referencia sobre el lenguaje Sax BASIC.

Ejemplo: Procesamiento de un cuadro de diálogo sencillo

Este proceso crea un cuadro de diálogo sencillo que abre un archivo de datos. Consulte las secciones relacionadas para obtener explicaciones sobre la subrutina `BuildDialog` y la función de control del cuadro de diálogo.

Figura 46-17

Cuadro de diálogo Abrir archivo de datos creado por el proceso



```
Sub Main
  Call BuildDialog
End Sub
```

```
'define el cuadro de diálogo
Sub BuildDialog
  Begin Dialog UserDialog 580,70,Abrir archivo de datos,.DialogFunc
    Text 40,7,280,21,Archivo de datos para abrir:,.txtDialogTitle
    TextBox 40,28,340,21,.txtFilename
    OKButton 470,7,100,21,.cmdOK
    CancelButton 470,35,100,21,.cmdCancel
  End Dialog
  Dim dlg As UserDialog
  Dialog dlg
End Sub
```

```

'define la función que determina el comportamiento del cuadro de diálogo
Function DialogFunc(strDlgItem As String, intAction As Integer, intSuppValue As Integer)
As Boolean  Select Case intAction
    Case 1' emite una señal sonora cuando se inicializa el cuadro de diálogo
        Beep
    Case 2' cambia el valor o pulsa en el botón
        Select Case strDlgItem
            Case cmdOK'si el usuario pulsa en Aceptar, abre el archivo de datos con el
                nombre de archivos especificado
                strFilename = DlgText(txtFilename)
            Call OpenDataFile(strFilename)
        DialogFunc = False
    Case cmdCancel'Si el usuario pulsa en Cancelar, cierra el cuadro de diálogo
        DialogFunc = False
    End Select
End Select
End Function

Sub OpenDataFile(strFilename As Variant)'Abre el archivo de datos con el nombre
de archivo especificado
Dim objDataDoc As ISpssDataDoc
Set objDataDoc = objSpssApp.OpenDataDoc(strFilename)
End Sub

```

También dispone de ejemplos en la ayuda en pantalla. Puede probarlos usted mismo pegando el código de la Ayuda en la ventana de proceso.

Depuración de procesos

EL menú Depurar permite recorrer el código paso a paso, ejecutando una línea o subrutina por separado y viendo el resultado. También puede insertar un punto de interrupción en el proceso para detener la ejecución en la línea que contiene dicho punto de interrupción.

Para depurar un autoproceso, abra el archivo de autoproceso en una ventana de proceso, inserte los puntos de interrupción en el procedimiento que desee depurar y, seguidamente, ejecute el procedimiento estadístico que desencadena el autoproceso.

Paso hacia dentro. Ejecuta la línea actual. Si la línea actual es una subrutina o una llamada de función, se detiene en la primera línea de esa subrutina o función.

Paso por encima. Ejecuta hasta la línea siguiente. Si la línea actual es una subrutina o llamada de función, ejecuta la subrutina o función por completo.

Paso hacia fuera. Sale de la subrutina o la llamada de función actual.

Paso hacia el cursor. Ejecuta hasta la línea actual.

Alternar salto. Inserta o elimina un punto de interrupción. El proceso se detiene en el punto de interrupción y muestra el panel de depuración.

Vista rápida. Muestra el valor de la expresión actual.

Añadir vista. Añade la expresión actual a la ventana de observación.

Visor de objetos. Muestra el visor de objetos.

Definir sentencia siguiente. Define la siguiente sentencia que se va a ejecutar. Sólo se pueden seleccionar instrucciones de la subrutina o la función actual.

Mostrar sentencia siguiente. Muestra la siguiente sentencia que se va a ejecutar.

Para seguir los pasos de un proceso

- ▶ En el menú Depurar, seleccione cualquiera de las opciones Paso... para ejecutar el código, una línea o una subrutina por vez.

Las pestañas Immediate, Watch, Stack y Loaded aparecerán en la ventana de proceso, junto con la barra de herramientas de depuración.

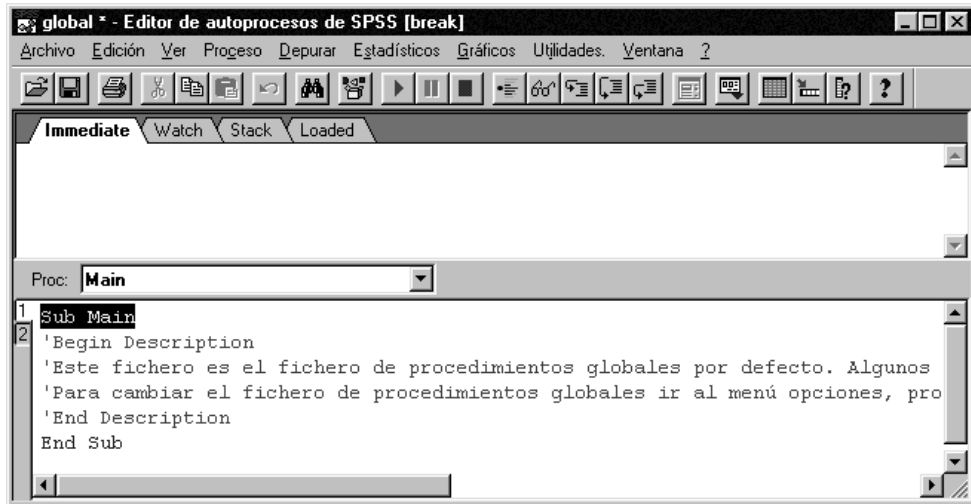
- ▶ Utilice la barra de herramientas (o las teclas de aceleración) para continuar con los pasos del proceso.
- ▶ De forma alternativa, seleccione Alternar salto para insertar un punto de interrupción en la línea actual.

El proceso se detendrá en el punto de interrupción.

Panel de depuración (Proceso)

Cuando se recorre el código por pasos se muestran las pestañas Immediate, Watch, Stack y Loaded.

Figura 46-18
Panel de depuración mostrado en la ventana de proceso



Pestaña Immediate. Pulse en el nombre de cualquier variable y en el icono de gafas para ver el valor actual de la variable. También puede evaluar una expresión, asignar una variable o invocar una subrutina.

- Escriba `?expr` y pulse Intro para mostrar el valor de `expr`.
- Escriba `var = expr` y pulse Intro para cambiar el valor de `var`.
- Escriba `subname args` y pulse Intro para invocar una subrutina o una instrucción pre-incorporada.
- Escriba `Trace` y pulse Intro para activar o desactivar el modo de traza. El modo de traza imprime cada instrucción en la ventana inmediata cuando hay un proceso en ejecución.

Pestaña Watch. Para mostrar una variable, una función o una expresión, pulse en ella y seleccione "Añadir vista" en el menú "Depurar". Los valores mostrados se actualizan cada vez que se detiene la ejecución. Puede editar la expresión a la izquierda de `->`. Pulse en "Entrar" para actualizar todos los valores de manera inmediata. Pulse `Ctrl-Y` para eliminar la línea.

Pestaña Stack. Muestra las líneas que invocaron la instrucción actual. La primera línea es la instrucción actual, la segunda línea es la que invocó a la primera, y así sucesivamente. Pulse en cualquier línea para resaltarla en la ventana de edición.

Pestaña Loaded. Presenta una lista de los procesos activos actualmente. Pulse en una línea para ver ese proceso.

Archivos de procesos y archivos de sintaxis

Los archivos de sintaxis (*.*sps*) no son lo mismo que los archivos de procesos (*.*sbs*). Los primeros poseen comandos escritos en el lenguaje de comandos que permiten ejecutar procedimientos estadísticos y transformaciones de datos. Mientras que los procesos permiten manipular los resultados y automatizar otras tareas que se realizan habitualmente utilizando la interfaz gráfica de los menús y cuadros de dialogo; el lenguaje de comandos proporciona un método alternativo de comunicación directa con el soporte del programa, la parte del sistema que se ocupa de los cálculos estadísticos y las transformaciones de datos.

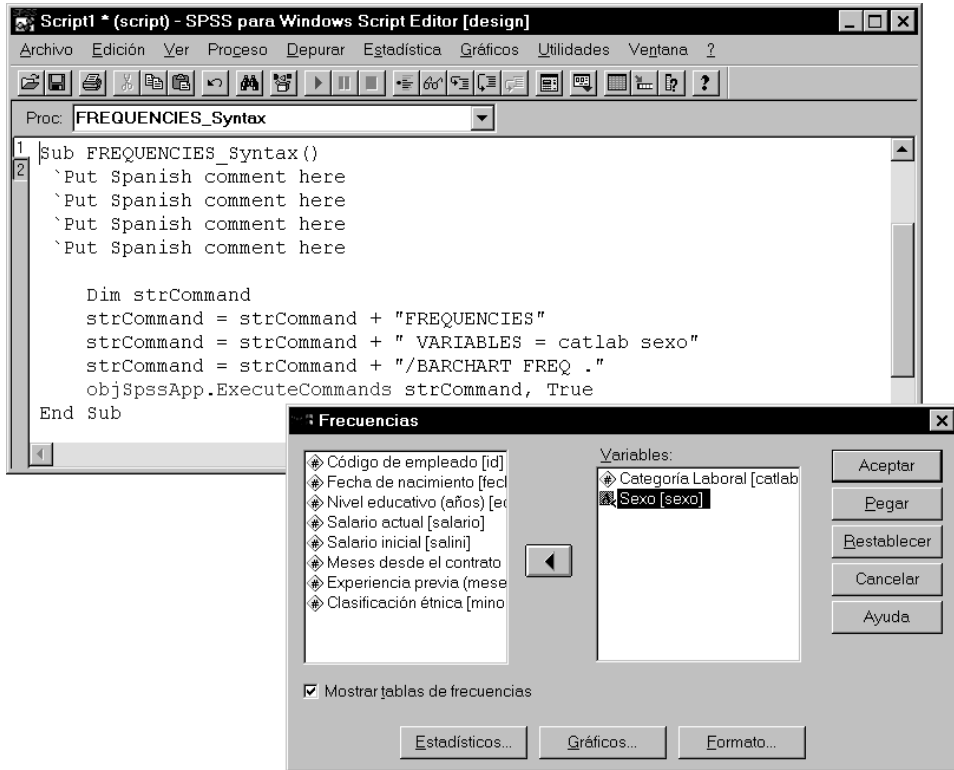
Puede combinar archivos de procesos y de sintaxis para obtener una mayor flexibilidad, ejecutando un proceso desde la sintaxis de comandos o incrustando la sintaxis de comandos dentro de un proceso.

Ejecución de sintaxis de comandos desde un proceso

Puede ejecutar la sintaxis de comandos desde un proceso de automatización utilizando el método `ExecuteCommands`. La sintaxis de comandos permite ejecutar transformaciones de datos y procedimientos estadísticos y generar gráficos. Buena parte de esta funcionalidad no se puede automatizar directamente desde los procesos de comandos.

La manera más fácil de construir un archivo de sintaxis de comandos es efectuar selecciones en los cuadros de diálogo y pegar la sintaxis de las selecciones en la ventana de proceso.

Figura 46-19
Pegado de sintaxis de comandos en un proceso



Al abrir los cuadros de diálogo mediante los menús de la ventana de proceso, el botón Pegar insertará todo el código necesario para ejecutar el comando desde un proceso.

Nota: Debe utilizar los menús de la ventana de proceso para abrir el cuadro de diálogo; de otro modo, los comandos se pegarán en una ventana de sintaxis en lugar de en la ventana de proceso.

Para pegar sintaxis de comandos de SPSS en un proceso

- ▶ En los menús de la ventana de proceso, elija comandos de los menús Estadísticos, Gráficos y Utilidades para abrir cuadros de diálogo.

- ▶ Realice selecciones en el cuadro de diálogo.
- ▶ Pulse en Pegar.

Nota: Debe utilizar los menús de la ventana de proceso para abrir el cuadro de diálogo; de otro modo, los comandos se pegarán en una ventana de sintaxis en lugar de en la ventana de proceso.

Ejecución de un proceso desde la sintaxis de comandos

Puede utilizar el comando `SCRIPT` para ejecutar un proceso desde la sintaxis de comandos. Especifique el nombre del proceso que desea ejecutar, con el nombre de archivo entre comillas, del siguiente modo:

```
SCRIPT 'C:\PROGRAM FILES\SPSS\CLEAN NAVIGATOR.SBS' .
```

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos de SPSS (*SPSS Command Syntax Reference*).

Sistema de gestión de resultados

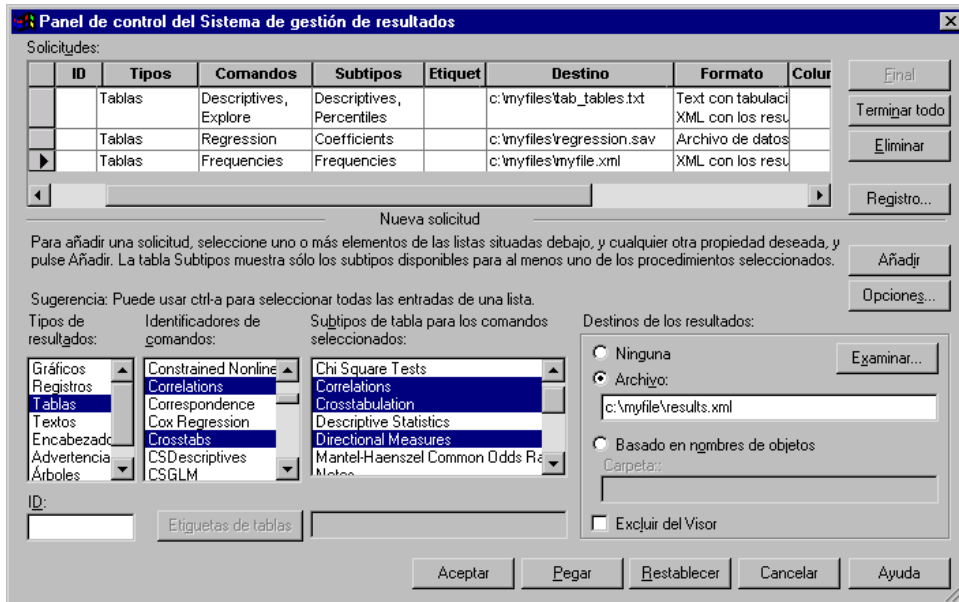
El Sistema de gestión de resultados (SGR) ofrece la posibilidad de escribir de forma automática las categorías de resultados seleccionadas en diferentes archivos de resultados y en distintos formatos. Los formatos incluyen:

- **Formato de archivo de datos de SPSS (.sav).** Los resultados que se muestran en tablas pivote en el Visor se pueden almacenar en forma de un archivo de datos de SPSS, permitiendo de esta manera utilizar los resultados como entrada de posteriores comandos.
- **XML.** Las tablas, los resultados de texto e incluso muchos tipos de gráficos se pueden almacenar en formato XML.
- **HTML.** Las tablas y los resultados de texto se pueden almacenar en formato HTML. Los gráficos normales (no interactivos) y los diagramas de modelo de árbol (opción Árbol de clasificación) se pueden incluir como archivos de imagen.
- **Texto.** Las tablas y los resultados de texto se pueden almacenar como texto delimitado por tabuladores o separado por espacios.

Para utilizar el panel de control del Sistema de gestión de resultados:

- ▶ Elija en los menús:
 - Utilidades
 - Panel de control de SGR...

Figura 47-1
Panel de control del Sistema de gestión de resultados



Puede utilizar el panel de control para iniciar y parar el envío de los resultados a los distintos destinos.

- Cada solicitud de SGR permanece activa hasta que finaliza de forma explícita o hasta el final de la sesión.
- Un archivo de destino especificado en una solicitud de SGR no está disponible para otros procedimientos de SPSS y otras aplicaciones hasta que finaliza la solicitud de SGR.
- Mientras que una solicitud de SGR está activa, los archivos de destino especificados se almacenan en la memoria (RAM); por consiguiente, las solicitudes de SGR activas que escriben una gran cantidad de resultados en archivos externos pueden consumir una cantidad considerable de memoria.
- Las distintas solicitudes de SGR son independientes. Un mismo resultado se puede enviar a distintas ubicaciones en formatos diversos basados en las especificaciones de las distintas solicitudes de SGR.

- El orden de los objetos de resultados en un destino concreto equivale al orden en que se han creado, lo cual se determina mediante el orden y el funcionamiento de los procedimientos que generan los resultados.
- SGR no puede enviar objetos de gráfico o advertencias creados mediante procedimientos de gráficos interactivos (menú Gráficos, submenú Interactivos) ni mapas creados mediante procedimientos de asignación (menú Gráficos, submenú Mapas).

Adición de solicitudes de SGR

Para añadir una solicitud de SGR:

- ▶ Seleccione los tipos de resultados (tablas, gráficos, etc.) que desee incluir. Si desea obtener más información, consulte “Tipos de objetos de resultados” en p. 699.
- ▶ Seleccione los comandos que desee incluir. Si desea incluir todos los resultados, seleccione todos los elementos de la lista. Si desea obtener más información, consulte “Identificadores de comandos y subtipos de tabla” en p. 701.
- ▶ En el caso de los comandos que generan resultados de tablas pivote, seleccione los tipos de tabla concretos que desee incluir. La lista muestra sólo las tablas disponibles en los comandos seleccionados; todo tipo de tabla disponible en uno o más comandos seleccionados se muestra en la lista. Si no se selecciona ningún comando, se muestran todos los tipos de tabla. Si desea obtener más información, consulte “Identificadores de comandos y subtipos de tabla” en p. 701.
- ▶ Para seleccionar tablas basadas en etiquetas de texto en lugar de subtipos, pulse en Etiquetas de tabla. Si desea obtener más información, consulte “Etiquetas de tabla” en p. 702.
- ▶ Pulse en Opciones para especificar el formato del resultado (por ejemplo, archivo de datos de SPSS, XML o HTML). Por defecto, se utiliza el formato XML con los resultados. Si desea obtener más información, consulte “Opciones de SGR” en p. 704.
- ▶ Especifique un archivo de destino de resultados o una ubicación de carpeta para varios archivos de destino basados en nombres de objetos. En caso de haber varios archivos basados en nombres de objetos, se crea un archivo independiente para cada

objeto de resultados con un nombre de archivo basado en los nombres de subtipos de tabla o en las etiquetas de tabla.

Si lo desea, tiene la posibilidad de:

- Excluir los resultados seleccionados del Visor. Si selecciona Excluir del Visor, los tipos de resultados de la solicitud de SGR no se muestran en la ventana del Visor. Si varias solicitudes SGR activas incluyen los mismos tipos de resultados, la presentación de dichos tipos de resultados en el Visor se determina mediante la solicitud de SGR más reciente que contiene los tipos de resultados. Si desea obtener más información, consulte “Exclusión de presentación de resultados del Visor” en p. 710.
- Asigne una cadena de ID a la solicitud. A todas las solicitudes se les asigna automáticamente un valor de ID; puede anular la cadena de ID por defecto del sistema con un ID descriptivo. Esto puede resultar útil si dispone de varias solicitudes activas que desea identificar fácilmente. Los valores de ID asignados no pueden empezar por un signo de dólar (\$).

Unos cuantos consejos sencillos para seleccionar varios elementos de una lista:

- Pulse Ctrl-A en el teclado para seleccionar todos los elementos de la lista.
- Mantenga pulsada la tecla Mayús para seleccionar varios elementos contiguos.
- Mantenga pulsada la tecla Ctrl para seleccionar varios elementos no contiguos.

Finalización y eliminación de solicitudes de SGR

Las solicitudes de SGR activas y nuevas se muestran en la lista Solicitudes con la solicitud más reciente en la parte superior. Puede cambiar los anchos de las columnas de información si pulsa y arrastra los bordes, y puede desplazar la lista horizontalmente para ver más información sobre una solicitud en concreto.

Un asterisco (*) después de la palabra “Activa” en la columna Estado indica que se ha creado una solicitud de SGR con una sintaxis de comandos que incluye funciones no disponibles en el panel de control. Si desea obtener información sobre el comando OMS, consulte la *Referencia de sintaxis de comandos* (disponible en formato PDF en el menú Ayuda).

Para finalizar una solicitud de SGR activa concreta:

- ▶ Pulse en cualquier casilla de la fila para dicha solicitud en la lista Solicitudes.
- ▶ Pulse en Terminar.

Para finalizar todas las solicitudes de SGR activas:

- ▶ Pulse en Terminar todo.

Para eliminar una solicitud nueva (una solicitud añadida que aún no esté activa):

- ▶ Pulse en cualquier casilla de la fila para dicha solicitud en la lista Solicitudes.
- ▶ Pulse en Eliminar.

Nota: las solicitudes de SGR activas no finalizan hasta que pulsa en Aceptar.

Tipos de objetos de resultados

Hay siete tipos de objetos de resultados distintos:

Gráficos. Gráficos (excepto gráficos “interactivos” y mapas). Los objetos de gráfico sólo se incluyen con los formatos de destino XML y HTML.

Logaritmos. Objetos de texto de logaritmo. Los objetos de logaritmo contienen determinados tipos de mensajes error y advertencia. En función de la configuración de las opciones (menú Edición, Opciones, Visor), los objetos de logaritmo pueden contener la sintaxis de comandos ejecutada durante la sesión. Los objetos de logaritmo tienen la etiqueta *Logaritmo* en el panel de titulares del Visor.

Tablas. Objetos de resultados que son tablas pivote en el Visor. Esto incluye las tablas de notas. Las tablas son los únicos objetos de resultados que se pueden dirigir al formato del archivo de datos de SPSS (.sav).

Texto. Objetos de texto que no son logaritmos ni encabezados. Esto incluye los objetos con la etiqueta *Resultados de texto* en el panel de titulares del Visor.

Árboles. Diagramas de modelo de árbol generados por la opción Árbol de clasificación. Los objetos de árbol sólo se incluyen con los formatos de destino XML y HTML.

Encabezados. Objetos de texto con la etiqueta *Título* en el panel de titulares del Visor. Para el formato XML con los resultados, no se incluyen objetos de texto de encabezado.

Advertencias. Objetos de advertencias. Los objetos de advertencias contienen determinados tipos de mensajes error y advertencia.

Figura 47-2

Tipos de objetos de resultados

The screenshot shows the SPSS Results Viewer window titled "Resultados1 - Visor SPSS". The interface includes a menu bar (Archivo, Edición, Ver, Datos, Transformar, Insertar, Formato, Analizar, Gráficos, Utilidades, Ventana, ?) and a toolbar. On the left, a tree view shows the results structure: Resultados > Frecuencias > Título, Notas, Estadístico: Categoría lab, Gráfico de Frecuencias > Título, Notas, Advertenci.

The main content area displays several objects:

- Heading:** "Frecuencias" (Frequency).
- Table:** A table titled "Categoría laboral" (Labor Category) with columns: Frecuencia, Porcentaje, Porcentaje válido, and Porcentaje acumulado. The data is as follows:

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos Administrativo	363	76,6	76,6	76,6
Seguridad	27	5,7	5,7	82,3
Directivo	84	17,7	17,7	100,0
Total	474	100,0	100,0	
- Chart:** A bar chart showing the frequency distribution for the three labor categories: Administrativo (approx. 363), Seguridad (approx. 27), and Directivo (approx. 84).
- Log:** A message indicating an error: ">Error # 1. Command name: comptue >The first word in the line is not recognized as an SPSS command. >This command not executed."
- Warning:** A warning message: "Advertencia Text: newvar An undefined variable name, or a scratch or system variable was specified in a variable list which accepts only standard variables. Check spelling and verify the existence of this variable. This command not executed."

Identificadores de comandos y subtipos de tabla

Identificadores de comandos

Los identificadores de comandos están disponibles para todos los procedimientos estadísticos y de gráficos y para cualquier otro comando que genere bloques de resultados con su propio encabezado identificable en el panel de titulares del Visor. Estos identificadores son por lo general (aunque no siempre) iguales o similares a los nombres de procedimiento de los menús y los títulos de los cuadros de diálogo, los cuales son por lo general (aunque no siempre) similares a los nombres de los comandos de SPSS subyacentes. Por ejemplo, el identificador de comandos para el procedimiento Frecuencias es “Frecuencias” y el nombre del comando subyacente es también el mismo.

No obstante, hay algunos casos en que el nombre del procedimiento y el identificador de comandos o el nombre del comando no son en absoluto similares. Por ejemplo, todos los procedimientos del submenú Pruebas no paramétricas (del menú Analizar) utilizan el mismo comando subyacente y el identificador de comandos es el mismo que el nombre de comando subyacente: Pruebas no paramétricas.

Subtipos de tabla

Los subtipos de tabla son los diferentes tipos de tablas pivote que se pueden generar. Algunos subtipos sólo son generados por un comando; otros subtipos se pueden generar mediante varios comandos (aunque las tablas pueden no presentar un aspecto similar). Aunque los nombres de subtipos de tabla suelen ser bastante descriptivos, puede haber muchos entre los que elegir (sobre todo si ha seleccionado un número considerable de comandos) o bien dos subtipos pueden tener nombres muy similares.

Búsqueda de identificadores de comandos y subtipos de tabla

En caso de duda, puede buscar los nombres de los identificadores de comandos y los subtipos de tabla en la ventana del Visor:

- ▶ Ejecute el procedimiento para generar algunos resultados en el Visor.
- ▶ Pulse con el botón derecho del ratón en el panel de titulares del Visor.

- ▶ En el menú contextual emergente, seleccione Copiar identificador de comandos SGR o Copiar subtipo de tablas SGR.
- ▶ Pegue el nombre del identificador de comandos o del subtipo de tabla copiado en cualquier editor de texto (por ejemplo, una ventana de sintaxis de SPSS).

Etiquetas de tabla

En lugar de los nombres de subtipos de tabla puede seleccionar tablas basadas en el texto mostrado en el panel de titulares del Visor. Las etiquetas resultan útiles para diferenciar varias tablas del mismo tipo en las que el texto del titular refleja algún atributo del objeto de resultados concreto como las etiquetas o los nombres de las variables. Hay, no obstante, ciertos factores que pueden afectar al texto de la etiqueta:

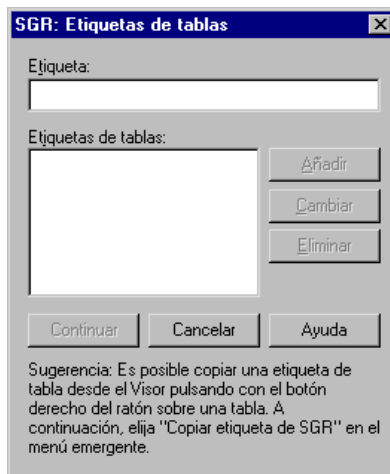
- Si está activado el procesamiento de segmentación del archivo, es posible que se añada a la etiqueta una identificación con el grupo de segmentación del archivo.
- Las etiquetas que incluyen información sobre las variables o los valores se ven afectadas por la configuración actual de las opciones de las etiquetas de resultados (menú Edición, Opciones, pestaña Etiquetas de los resultados).
- Las etiquetas se ven afectadas por el ajuste actual del idioma de los resultados (menú Edición, Opciones, pestaña General).

Para especificar las etiquetas que se van a utilizar para identificar las tablas de resultados:

- ▶ En el panel de control del Sistema de gestión de resultados, seleccione Tablas en la lista Tipos de resultados (también puede seleccionar otros Tipos de resultados, aunque Tablas debe ser uno de los tipos seleccionados) y seleccione uno o más comandos.

Pulse en Etiquetas de tabla.

Figura 47-3
Cuadro de diálogo Etiquetas de tabla



- ▶ Introduzca la etiqueta de tabla *exactamente* como aparece en el panel de titulares de la ventana del Visor. También puede pulsar con el botón derecho del ratón en el elemento del titular, seleccionar Copiar etiqueta de SGR y pegar la etiqueta copiada en el campo de texto Etiqueta.
- ▶ Pulse en Añadir.
- ▶ Repita los pasos para cada etiqueta de tabla que desee incluir.
- ▶ Pulse en Continuar.

Comodines

Puede utilizar un asterisco (*) como último carácter de la cadena de etiqueta y como carácter comodín. Se seleccionarán todas las tablas que empiecen por la cadena especificada (excepto el asterisco). Esto sólo funciona si el asterisco es el último carácter, ya que los asteriscos pueden aparecer como caracteres válidos en una etiqueta de tabla.

Opciones de SGR

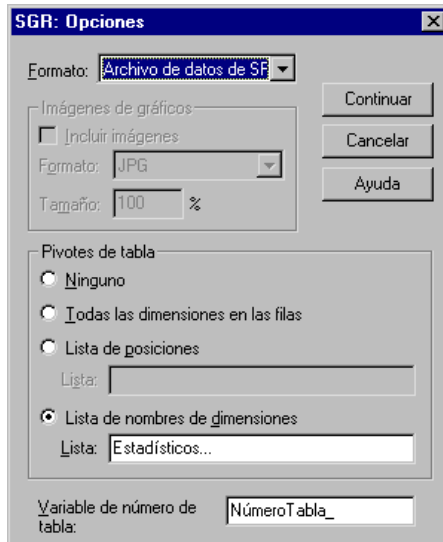
Puede utilizar el cuadro de diálogo Opciones de SGR para:

- Especificar el formato del resultado.
- Incluir o excluir los resultados del diagrama de modelo de árbol y especificar el formato gráfico.
- Especificar qué elementos de la dimensión de tabla deben tener la dimensión de las filas.
- En el caso del formato de archivo de datos de SPSS, incluya una variable que identifique el número de tabla secuencial que sea el origen en cada caso.

Para especificar las opciones de SGR:

- ▶ Pulse en Opciones en el panel de control del Sistema de gestión de resultados.

Figura 47-4
Cuadro de diálogo SGR: Opciones



Formato

XML con los resultados. XML que se adapta al esquema de *resultados de SPSS*. Los gráficos normales se incluyen como XML adaptado al esquema *vizml*. Además, puede exportar todos los gráficos y mapas como archivos independientes en el formato gráfico seleccionado.

HTML. Objetos de resultados que son tablas pivote en el Visor y se convierten en tablas HTML simples. No se admite ningún atributo de aspecto de tabla (características de fuente, estilos de borde, colores, etc.). Los objetos de resultados de texto tienen la etiqueta <PRE> en HTML. Si opta por incluir gráficos, estos se exportan como archivos independientes en el formato gráfico seleccionado y se incrustan mediante referencias () en el documento HTML.

Archivo de datos de SPSS. Se trata de un formato de archivo binario. Todos los tipos de objetos de resultados distintos de las tablas se excluyen. Cada columna de una tabla se convierte en una variable en el archivo de datos. Para utilizar un archivo de datos creado con SGR en la misma sesión, debe finalizar la solicitud de SGR activa para abrir el archivo de datos. Si desea obtener más información, consulte “Envío de resultados a archivos de datos de SPSS” en p. 711.

SVWS XML. XML utilizado por el servidor Web de SmartViewer. Se trata en realidad de un archivo JAR/ZIP que contiene XML, CSV y otros archivos. El servidor Web de SmartViewer es un producto independiente basado en un servidor.

Texto. Texto separado por espacios. Los resultados se escriben como texto con los resultados tabulares alineados con espacios para las fuentes de paso fijo. Se excluyen todos los gráficos y mapas.

Texto con tabulaciones. Texto delimitado por tabulaciones. Para los resultados que son tablas pivote en el Visor, las tabulaciones delimitan los elementos de columnas de tabla. Las líneas de bloque de texto se escriben sin cambios; no se realiza ningún intento de dividirlos con tabulaciones en las posiciones útiles. Se excluyen todos los gráficos y mapas.

Imágenes de gráficos

Para el formato HTML, puede incluir gráficos (excluyendo los gráficos interactivos) y diagramas de modelo de árbol como archivos de imagen. Se crea un archivo de imagen independiente para cada gráfico o árbol, y se incluyen etiquetas estándar en HTML para cada archivo de imagen.

Los archivos de imagen se guardan en un subdirectorio (carpeta) independiente. El nombre del subdirectorio es el nombre del archivo de destino HTML sin ninguna extensión y con “_files” añadido al final. Por ejemplo, si el archivo de destino HTML es *julydata.htm*, el subdirectorio de imágenes se llamará *julydata_files*.

- **Formato de imagen.** Los formatos de imagen disponibles son: PNG, JPG, EMF y BMP.
- **Tamaño.** Puede ajustar la escala del tamaño de imagen del 10% al 200%.

Pivotes de tabla

Para los resultados de las tablas pivote, puede especificar los elementos de dimensión que deben aparecer en las columnas. El resto de los elementos de dimensión aparecen en las filas. Para el formato de archivo de datos de SPSS, las columnas de tabla se convierten en variables y las filas en casos.

- Si especifica varios elementos de dimensión para las columnas, estos se anidan en las columnas en el orden en que aparecen. Para el formato de archivo de datos de SPSS, los nombres de variable se generan mediante elementos de columna

anidados. Si desea obtener más información, consulte “Nombres de variable en los archivos de datos generados por SGR” en p. 719.

- Si la tabla no contiene ninguno de los elementos de dimensión que aparecen, todos los elementos de dimensión para dicha tabla aparecerán en las filas.
- Los pivotes de tabla especificados aquí no surtirán ningún efecto en las tablas que se muestran en el Visor.

Cada dimensión de una tabla (fila, columna, capa) puede contener cero o más elementos. Por ejemplo, una tabla de contingencia sencilla de dos dimensiones contiene un único elemento de dimensión de fila y un único elemento de dimensión de columna, cada uno de los cuales contiene una de las variables utilizadas en la tabla. Puede utilizar argumentos de posición o “nombres” de elementos de dimensión para especificar los elementos de dimensión que desea colocar en la dimensión de columna.

Lista de posiciones. El formato general de un argumento de posición es una letra que indica la posición por defecto del elemento (C para columna, R para fila o L para capa) seguida de un número entero positivo que indica la posición por defecto en la dimensión. Por ejemplo, R1 indica el elemento de dimensión de fila más exterior.

- Para especificar varios elementos de diversas dimensiones, separe cada dimensión con un espacio. Por ejemplo: R1 C2.
- La letra de dimensión seguida de ALL indica todos los elementos de dicha dimensión en el orden por defecto. Por ejemplo, CALL equivale al comportamiento en que se utilizan todos los elementos de columna en el orden por defecto para generar columnas.
- CALL RALL LALL (o RALL CALL LALL, etc.) coloca los elementos de dimensión en las columnas. Para el formato de archivo de datos de SPSS, se genera de este modo una fila o caso por tabla en el archivo de datos.

Figura 47-5
Argumentos de posición de filas y columnas

		R1	R2	R3	C1				
Modelo					Experiencia previa (meses)	Meses desde el contrato	Salario inicial	Fecha de nacimiento	
1	Correlaciones	Experiencia previa (meses)			1,000	,067	-,087	,805	
		Meses desde el contrato			,067	1,000	,012	,085	
		Salario inicial			-,087	,012	1,000	-,075	
		Fecha de nacimiento			,805	,085	-,075	1,000	
	Covarianzas	Experiencia previa (meses)				31,307	12,940	-,022	7,096E-06
		Meses desde el contrato				12,940	1205,248	,019	4,635E-06
		Salario inicial				-,022	,019	,002	-5,236E-09
		Fecha de nacimiento				7,096E-06	4,635E-06	-5,236E-09	2,485E-12

Lista de nombres de dimensiones. En lugar de argumentos de posición, puede utilizar “nombres” de elementos de dimensión que son las etiquetas de texto que aparecen en la tabla. Por ejemplo, una tabla de contingencia sencilla de dos dimensiones contiene un único elemento de dimensión de fila y un único elemento de dimensión de columna, cada uno de los cuales incluye etiquetas basadas en las variables de dichas dimensiones además de un único elemento de dimensión de capa con la etiqueta *Estadísticos* (si el idioma de los resultados es el inglés).

- Los nombres de elementos de dimensión pueden variar en función del idioma de los resultados y la configuración que influye en la presentación de los nombres de variable o las etiquetas de las tablas.
- Cada nombre de elemento de dimensión debe aparecer entre comillas simples o dobles. Para especificar varios nombres de elementos de dimensión, incluya un espacio entre cada nombre entre comillas.

Las etiquetas asociadas con los elementos de dimensión pueden no ser siempre evidentes. Para ver todos los elementos de dimensión y sus etiquetas para una tabla pivote concreta:

- ▶ Active (pulse dos veces en) la tabla en el Visor.

- ▶ Elija en los menús:
Ver
Mostrar todo

y/o
- ▶ Si los paneles de pivotado no se muestran, elija en los menús:
Pivotar
Paneles de pivotado
- ▶ Sitúe el puntero del ratón sobre cada icono de los paneles de pivotado para obtener una pista emergente que muestre la etiqueta.

Figura 47-6

Nombres de elementos de dimensión mostrados en la tabla y los paneles de pivotado

Tabla sin las etiquetas de elementos de dimensión visibles

Sexo		Salario inicial	Salario actual
Mujer	Media	\$13,091.97	\$26,031.92
	N	216	216
Hombre	Media	\$20,301.40	\$41,441.78
	N	258	258
Total	Media	\$17,016.09	\$34,419.57
	N	474	474

Tabla con las etiquetas de elementos de dimensión visibles

Sexo	Estadísticos	Salario inicial	Salario actual
Mujer	Media	\$13,091.97	\$26,031.92
	N	216	216
Hombre	Media	\$20,301.40	\$41,441.78
	N	258	258
Total	Media	\$17,016.09	\$34,419.57
	N	474	474

Paneles de pivotado

Variables

1 elemento: seleccionados (0 ocultos/contraídos) SPS:

Registro

Puede registrar la actividad de SGR en un registro de XML o formato de texto.

- El registro realiza un seguimiento de todas las solicitudes de SGR nuevas para la sesión; no incluye las solicitudes de SGR activas antes de solicitar un registro.
- El archivo de registro actual finaliza si especifica un nuevo archivo de registro o si anula selección (desactiva) Registrar actividad de SGR.

Para especificar el registro de SGR:

- ▶ Pulse en Registro en el panel de control del Sistema de gestión de resultados.

Exclusión de presentación de resultados del Visor

La casilla de verificación Excluir del Visor suprime la presentación de todos los resultados seleccionados en la solicitud de SGR en la ventana del Visor. Esto suele resultar útil para los trabajos de producción que generan una gran cantidad de resultados y en caso de no necesitar los resultados como un documento del Visor (archivo *.spo*). Además, puede utilizarla para suprimir la presentación de determinados objetos de resultados que no desea ver sin enviar ningún otro resultado a un archivo y formato externos.

Para suprimir la presentación de determinados objetos de resultados sin dirigir otros resultados a un archivo externo:

- ▶ Cree una solicitud de SGR que identifique los resultados no deseados.
- ▶ Seleccione Excluir del Visor.
- ▶ Para el destino de los resultados, seleccione Archivo, pero *no* introduzca ninguna especificación de archivo. Deje el campo Archivo en blanco.
- ▶ Pulse en Añadir.

Los resultados seleccionados se excluyen del Visor, mientras que el resto de los resultados se muestran en el Visor del modo normal.

Envío de resultados a archivos de datos de SPSS

Un archivo de datos de SPSS consta de variables en las columnas y de casos en las filas, lo que representa básicamente el modo en que las tablas pivote se convierten en archivos de datos:

- Las columnas de la tabla son variables en el archivo de datos. Los nombres de variable válidos se generan a partir de etiquetas de columna.
- Las etiquetas de fila de la tabla se convierten en variables con nombres de variable genérica (*Var1*, *Var2*, *Var3*, etc.) en el archivo de datos. Los valores de estas variables son etiquetas de fila en la tabla.
- En el archivo de datos se incluyen automáticamente tres variables de identificador de tabla: *Comando_*, *Subtipo_* y *Etiqueta_*. Las tres son variables de cadena. Las dos primeras se corresponden con los identificadores de comandos y subtipos. Si desea obtener más información, consulte “Identificadores de comandos y subtipos de tabla” en p. 701. *Etiqueta_* contiene el texto de título de tabla.
- Las filas de la tabla se convierten en casos en el archivo de datos.

Ejemplo: Tabla sencilla de dos dimensiones

En el caso más simple (una tabla sencilla de dos dimensiones), las columnas de tabla se convierten en variables y las filas se convierten en casos en el archivo de datos.

Figura 47-7
Tabla sencilla de dos dimensiones

Informe

Sexo		Salario actual	Salario inicial
Mujer	Media	\$26.031,92	\$13.091,97
	Mediana	\$24.300,00	\$12.375,00
	Mínimo	\$15.750	\$9.000
	Máximo	\$58.125	\$30.000
Hombre	Media	\$41.441,78	\$20.301,40

Command	Subtype	Label	Var1	Var2	Salario actual	Salario inicial
1	Medias	Informe	Mujer	Media	\$26.031,92	\$13.091,97
2	Medias	Informe	Mujer	Mediana	\$24.300,00	\$12.375,00
3	Medias	Informe	Mujer	Mínimo	\$15.750,00	\$9.000,00
4	Medias	Informe	Mujer	Máximo	\$58.125,00	\$30.000,00
5	Medias	Informe	Hombre	Media	\$41.441,78	\$20.301,40
6	Medias	Informe	Hombre	Mediana	\$32.850,00	\$15.750,00
7	Medias	Informe	Hombre	Mínimo	\$19.650,00	\$9.000,00
8	Medias	Informe	Hombre	Máximo	\$135.000,00	\$79.980,00
9	Medias	Informe	Total	Media	\$34.419,57	\$17.016,09
10	Medias	Informe	Total	Mediana	\$28.875,00	\$15.000,00
11	Medias	Informe	Total	Mínimo	\$15.750,00	\$9.000,00
12	Medias	Informe	Total	Máximo	\$135.000,00	\$79.980,00

- Las primeras tres variables identifican la tabla de origen por comando, subtipo y etiqueta.
- Los dos elementos que definen las filas de la tabla (valores de la variable *Sexo* y medidas estadísticas) reciben los nombres de variable genérica *Var1* y *Var2*. Ambos elementos son variables de cadena.
- Las etiquetas de columna de la tabla se utilizan para crear nombres de variable válidos. En este caso, dichos nombres de variable se basan en las etiquetas de variable de las tres variables de escala resumidas en la tabla. Si las variables no tienen etiquetas de variable definidas o bien ha optado por mostrar los nombres de variable en lugar de las etiquetas de variable como las etiquetas de columna en la tabla, los nombres de variable del nuevo archivo de datos deben ser los mismos que los del archivo de datos de origen.

Ejemplo: Tablas con capas

Además de filas y columnas, una tabla puede contener una tercera dimensión: la dimensión de capas.

Figura 47-8
Tabla con capas

El diagrama muestra dos tablas de capas. La 'Capa 2' (Clasificación de minorías Sí) tiene 40 mujeres y 47 hombres. La 'Capa 1' (Clasificación de minorías No) tiene 166 mujeres y 110 hombres en la categoría Administrativa, 0 mujeres y 14 hombres en Seguridad, y 10 mujeres y 70 hombres en Directivo.

La captura de pantalla del SPSS Editor de datos muestra la siguiente tabla de datos:

	Var1	Var2	Var3	Var4	Mujer	Hombre
1	Categoría laboral	Administrativo	Clasificación de minorías	No	166	110
2	Categoría laboral	Administrativo	Clasificación de minorías	Sí	40	47
3	Categoría laboral	Seguridad	Clasificación de minorías	No	0	14
4	Categoría laboral	Seguridad	Clasificación de minorías	Sí	0	13
5	Categoría laboral	Directivo	Clasificación de minorías	No	10	70
6	Categoría laboral	Directivo	Clasificación de minorías	Sí	0	4

- En la tabla, la variable con la etiqueta *Clasificación étnica* define las capas. En el archivo de datos, esto crea dos variables adicionales: una que identifica el elemento de capa y otra que identifica las categorías del elemento de capa.
- Al igual que en el caso de las variables creadas a partir de elementos de fila, las variables creadas a partir de elementos de capa son variables de cadena con nombres de variable genérica (el prefijo *Var* seguido de un número secuencial).

Archivos de datos creados a partir de varias tablas

Si se envían varias tablas al mismo archivo de datos, cada tabla se añade al archivo de datos de un modo similar a la fusión de archivos de datos mediante la adición de casos de un archivo de datos a otro (menú Datos, Fundir archivos, Añadir casos).

- Cada tabla siguiente añade siempre casos al archivo de datos.
- Si las etiquetas de columna de las tablas difieren, cada tabla puede añadir variables al archivo de datos con valores perdidos para los casos de otras tablas que no tienen una columna con un etiquetado idéntico.

Ejemplo: Varias tablas con las mismas etiquetas de columna

Las tablas que contienen las mismas etiquetas de columna suelen generar los archivos de datos más útiles de inmediato (archivos que no requieren manipulación adicional). Por ejemplo, dos o más tablas de frecuencia del procedimiento Frecuencias tienen etiquetas de columna idénticas.

Figura 47-9
 Dos tablas con etiquetas de columna idénticas

Sexo					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Hombre	258	54,4	54,4	54,4
	Mujer	216	45,6	45,6	100,0

Categoría laboral					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Administrativo	363	76,6	76,6	76,6
	Seguridad	27	5,7	5,7	82,3
	Directivo	84	17,7	17,7	100,0

Subtype_	Label_	Var1	Var2	Frecuencia	Porcentaje	Porcentaje vál	Porcentaje acumulad
1 Frecuencias	Sexo	Válido	Mujer	216	45,6	45,6	45,6
2 Frecuencias	Sexo	Válido	Hombre	258	54,4	54,4	100,0
3 Frecuencias	Sexo	Válido	Total	474	100,0	100,0	.
4 Frecuencias	Categoría laboral	Válido	Adminis	363	76,6	76,6	76,6
5 Frecuencias	Categoría laboral	Válido	Segurid	27	5,7	5,7	82,3
6 Frecuencias	Categoría laboral	Válido	Directiv	84	17,7	17,7	100,0
7 Frecuencias	Categoría laboral	Válido	Total	474	100,0	100,0	.

- La segunda tabla contribuye con casos adicionales (filas) al archivo de datos, pero no aporta ninguna variable nueva dado que las etiquetas de columna son exactamente iguales; por consiguiente, no hay ningún archivo de revisión de datos perdidos de gran tamaño.
- Aunque los valores para *Comando_* y *Subtipo_* son los mismos, el valor *Etiqueta_* identifica la tabla de origen para cada grupo de casos dado que las dos tablas de frecuencia tienen distintos títulos.

Ejemplo: Varias tablas con distintas etiquetas de columna

Se crea una nueva variable en el archivo de datos para cada etiqueta de columna única de las tablas enviadas al archivo de datos, lo que ofrece como resultado una serie de bloques de valores perdidos si las tablas contienen distintas etiquetas de columna.

Figura 47-10
 Dos tablas con distintas etiquetas de columna

Tabla 1			
Sexo		Salario inicial	Salario actual
Mujer	Media	\$13.091,97	\$26.031,92
	Mediana	\$12.375,00	\$24.300,00
Hombre	Media	\$20.301,40	\$41.441,78

Tabla 2			
Sexo		Nivel educativo	Meses desde el contrato
Mujer	Media	12,37	80,38
	Mediana	12,00	81,00
Hombre	Media	14,43	81,72

Subtype_	Label_	Var1	Var2	Salario actual	Salario inicial	Nivel educativo	Meses desde el contrato
1	Informe	Mujer	Media	\$26032	\$13092,0	.	.
2	Informe	Mujer	Mediana	\$24300	\$12375,0	.	.
3	Informe	Hombre	Media	\$41442	\$20301,4	.	.
4	Informe	Hombre	Mediana	\$32850	\$15750,0	.	.
5	Informe	Total	Media	\$34420	\$17016,1	.	.
6	Informe	Total	Mediana	\$28875	\$15000,0	.	.
7	Informe	Mujer	Media	.	.	12,37	80,38
8	Informe	Mujer	Mediana	.	.	12,00	81,00
9	Informe	Hombre	Media	.	.	14,43	81,72
10	Informe	Hombre	Mediana	.	.	15,00	82,00
11	Informe	Total	Media	.	.	13,49	81,11
12	Informe	Total	Mediana	.	.	12,00	81,00

- La primera tabla tiene columnas con las etiquetas *Salario inicial* y *Salario actual*, no presentes en la segunda tabla, lo que resulta en una serie de valores perdidos para dichas variables en los casos de la segunda tabla.
- Por el contrario, la segunda tabla tiene columnas con las etiquetas *Nivel de formación académica* y *Meses desde el contrato*, no presentes en la primera tabla, lo que resulta en una serie de valores perdidos para dichas variables en los casos de la primera tabla.
- Las variables discordantes como las de este ejemplo se pueden generar incluso con tablas del mismo subtipo. De hecho, en este ejemplo ambas tablas son del mismo subtipo.

Ejemplo: Archivos de datos no creados a partir de varias tablas

Si alguna tabla no tiene el mismo número de elementos de fila que otras tablas, no se creará ningún archivo de datos. El número de filas no tiene que ser el mismo; el número de *elementos* de fila que se convierten en variables en el archivo de datos debe ser el mismo. Por ejemplo, una tabla de contingencia de dos variables y una tabla de contingencia de tres variables contienen distintos números de elementos de fila dado que la variable “capa” está anidada en la variable de fila de la presentación de la tabla de contingencia de tres variables por defecto.

Figura 47-11

Tablas con distintos números de elementos de fila

Tabla de contingencia Categoría laboral * Sexo

		Sexo		Total
		Mujer	Hombre	
Categoría laboral	Administrativo	206	157	363
	Seguridad	0	27	27
	Directivo	10	74	84
Total		216	258	474

Tabla de contingencia Categoría laboral * Sexo * Clasificación de minorías

Minority Classification			Sexo		Total
			Mujer	Hombre	
No	Categoría laboral	Administrativo	166	110	276
		Seguridad	0	14	14
		Directivo	10	70	80
	Total	176	194	370	
Sí	Categoría laboral	Administrativo	40	47	87
		Seguridad	0	13	13
		Directivo	0	4	4
	Total	40	64	104	

Control de elementos de columna para las variables de control del archivo de datos

En el cuadro de diálogo Opciones del panel de control de SGR puede especificar qué elementos de dimensión deben estar en las columnas y se utilizan, por tanto, para crear variables en el archivo de datos generado. Esto es equivalente al pivotado de la tabla en el Visor.

Por ejemplo, el procedimiento Frecuencias genera una tabla de estadísticos descriptivos con estadísticos en las filas, mientras que el procedimiento Descriptivos genera una tabla de estadísticos descriptivos con estadísticos en las columnas. Para incluir ambos tipos de tabla en el mismo archivo de datos de forma significativa, debe cambiar la dimensión de columna de una de ellas.

Dado que ambos tipos de tabla utilizan el nombre de elemento “Estadísticos” para la dimensión de estadísticos, podemos colocar los estadísticos de la tabla Frecuencias: Estadísticos en las columnas si especifica “Estadísticos” (entre comillas) en la lista de nombres de dimensiones del grupo Pivotes de tabla del cuadro de diálogo Opciones.

Figura 47-12
Cuadro de diálogo SGR: Opciones

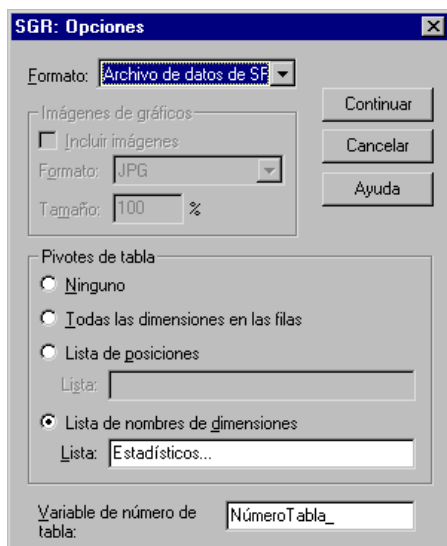


Figura 47-13
Combinación de distintos tipos de tablas en un archivo de datos mediante el pivotado de elementos de dimensión

Estadísticos			
		Salario inicial	Salario actual
N	Válidos	474	474
	Perdidos	0	0
Media		\$17,016.09	\$34,419.57
Media			
Suma			

Estadísticos					
	N		Media	Mediana	Suma
	Válidos	Perdidos			
Salario inicial	474	0	\$17,016.09	474	\$8,065,625
Salario actual	474	0	\$34,419.57	\$28,875.00	\$16,314,875

Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desv. típ.
Meses desde el contrato	474	63	98	81,11	10,061
Experiencia previa (meses)	474	0	476	95,86	104,586

temp.sav - SPSS Editor de datos							
g : Media							
	Var1	Válidos	Perdidos	Media	Mínimo	Máximo	N
1	Salario actual	474	0	17016	9000	79980	.
2	Salario inicial	474	0	34420	15750	135000	.
3	Meses desde el contrato	.	.	81,11	63	98	474
4	Experiencia previa (meses)	.	.	95,86	0	476	474
5	N válido (según lista)	474

Algunas de las variables incluyen valores perdidos dado que las estructuras de tabla aún no son exactamente igual a los estadísticos de las columnas.

Nombres de variable en los archivos de datos generados por SGR

SGR genera nombres de variable exclusivos y válidos a partir de etiquetas de columna:

- A los elementos de fila y capa se les asignan nombres de variable genérica: el prefijo *Var* seguido de un número secuencial.

- Los caracteres no permitidos en los nombres de variable (por ejemplo, espacio o paréntesis) se eliminan. Por ejemplo, “Esta etiqueta (columna)” se convierte en una variable con el nombre *EstaEtiquetaColumna*.
- Si la etiqueta empieza por un carácter permitido en los nombres de variable, pero no permitido como primer carácter (por ejemplo, un número), se inserta “@” como prefijo. Por ejemplo, “2º” se convierte en una variable llamada @2º.
- Los caracteres de subrayado o los puntos al final de las tablas se eliminan de los nombres de variable resultantes. (Los caracteres de subrayado al final de las variables generadas automáticamente *Comando_*, *Subtipo_* y *Etiqueta_* no se eliminan.)
- Si hay más de un elemento en la dimensión de columna, los nombres de variable se generan mediante la combinación de etiquetas de categoría con caracteres de subrayado entre dichas etiquetas. Las etiquetas de grupo no se incluyen. Por ejemplo, si *VarB* se anida bajo *VarA* en las columnas, obtendrá variables como *CatA1_CatB1*, pero no *VarA_CatA1_VarB_CatB1*.

Figura 47-14

Nombres de variable generados a partir de elementos de tabla

Variable de capa		Variable de columna			Variable de capa	
		Entrada	(Salida)	Total		
Máximos	Variable de fila	Sí	1	1	2	
		No	0	1	1	
	Total		1	2	3	
Mínimos	Variable de fila	Sí	1	0	1	
		No	1	1	2	
	Total		2	1	3	

Variable de capa		Máximos			Mínimos		
		Variable de columna		Total	Variable de columna		Total
		Entrada	(Salida)	Total	Entrada	(Salida)	Total
Variable de fila	Sí	1	1	2	1	0	1
	No	0	1	1	1	1	2
	Total	1	2	3	2	1	3

Estructura de tablas OXML

El formato XML de resultados (OXML) equivale a XML adaptado al esquema *resultados de SPSS*. Si desea obtener una descripción detallada del esquema, consulte *SPSSOutputXML_schema.htm* en la carpeta *help\main* de la carpeta de instalación de SPSS.

- Los identificadores de comandos y subtipos de SGR se utilizan como valores de los atributos `command` y `subType` en OXML. Por ejemplo:

```
<command text=Frequencies command=Frequencies...>
  <pivotTable text=Gender label=Gender subType=Frequencies...>
```

- Los valores de los atributos `command` y `subType` de SGR no se ven afectados por el idioma de los resultados o la configuración de presentación para los nombres de variable y etiquetas o para los valores y etiquetas de valor.
- XML distingue entre mayúsculas y minúsculas. El valor del atributo `subType` de “frequencies” *no* es igual al valor del atributo `subType` de “Frequencies”.
- Toda la información mostrada en la tabla se incluye en los valores de atributo de OXML. En el nivel de casillas individuales, OXML consta de elementos “vacíos” que contienen atributos, pero no se incluye ningún “contenido” distinto del que se incluye en los valores de atributo.
- La estructura de tablas en OXML se representa por filas; los elementos que representan las columnas se anidan en las filas y las casillas individuales se anidan en los elementos de columna:

```
<pivotTable...>
  <dimension axis='row'...>
    <dimension axis='column'...>
      <category...>
        <cell text='...' number='...' decimals='...'/>
      </category>
      <category...>
        <cell text='...' number='...' decimals='...'/>
      </category>
    </dimension>
  </dimension> ...
</pivotTable>
```

El ejemplo anterior es una representación simplificada de la estructura que muestra las relaciones descendentes/ascendentes de estos elementos, aunque no necesariamente las relaciones parentales/filiales, ya que suele haber niveles de elementos anidados intercalados.

Las siguientes dos figuras muestran una tabla de frecuencias simple y la representación completa de XML con los resultados de dicha tabla.

Figura 47-15
Tabla de frecuencias simple

		Sexo			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Hombre	258	54,4	54,4	54,4
	Mujer	216	45,6	45,6	100,0
Total		474	100,0	100,0	

```
<?xml version=1.0 encoding=UTF-8 ?>
<outputTreeoutputTree xmlns=http://xml.spss.com/spss/oms
xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance
xsi:schemaLocation=http://xml.spss.com/spss/oms
http://xml.spss.com/spss/oms/spss-output-1.0.xsd>
<command text=Frequencies command=Frequencies
displayTableValues=label displayOutlineValues=label
displayTableVariables=label displayOutlineVariables=label>
<pivotTable text=Gender label=Gender subType=Frequencies
varName=gender variable=true>
<dimension axis=row text=Gender label=Gender
varName=gender variable=true>
<group text=Valid>
<group hide=true text=Dummy>
<category text=Female label=Female string=f
varName=gender>
<dimension axis=column text=Statistics>
<category text=Frequency>
<cell text=216 number=216/>
</category>
<category text=Percent>
<cell text=45.6 number=45.569620253165 decimals=1/>
</category>
<category text=Valid Percent>
<cell text=45.6 number=45.569620253165 decimals=1/>
</category>
```

```
<category text=Cumulative Percent>
  <cell text=45.6 number=45.569620253165 decimals=1/>
</category>
</dimension>
</category>
<category text=Male label=Male string=m varName=gender>
  <dimension axis=column text=Statistics>
    <category text=Frequency>
      <cell text=258 number=258/>
    </category>
    <category text=Percent>
      <cell text=54.4 number=54.430379746835 decimals=1/>
    </category>
    <category text=Valid Percent>
      <cell text=54.4 number=54.430379746835 decimals=1/>
    </category>
    <category text=Cumulative Percent>
      <cell text=100.0 number=100 decimals=1/>
    </category>
  </dimension>
</category>
</group>
<category text=Total>
  <dimension axis=column text=Statistics>
    <category text=Frequency>
      <cell text=474 number=474/>
    </category>
    <category text=Percent>
      <cell text=100.0 number=100 decimals=1/>
    </category>
    <category text=Valid Percent>
      <cell text=100.0 number=100 decimals=1/>
    </category>
  </dimension>
</category>
</group>
</dimension>
</pivotTable>
</command>
</outputTree>
```

Como puede observar, una tabla sencilla de dimensiones reducidas genera una cantidad considerable de XML. Esto se debe en parte a que XML contiene información no evidente en la tabla original, información que puede no estar disponible en la tabla original y una cantidad determinada de redundancia.

- El contenido de la tabla tal y como aparece (o podría aparecer) en una tabla pivote en el Visor se incluye en los atributos de texto. Por ejemplo:

```
<command text=Frequencies command=Frequencies...>
```

- Los atributos de texto se pueden ver afectados por el idioma de los resultados y la configuración que influyen en la presentación de los nombres de variable y etiquetas o valores y etiquetas de valor. En este ejemplo, el valor del atributo de **texto** difiere en función del idioma de los resultados, mientras que el valor del atributo de comando permanece igual independientemente del idioma de los resultados.
- Siempre que las variables o los valores se utilicen en las etiquetas de fila o columna, XML contiene un atributo de **texto** y uno o más valores de atributo adicionales. Por ejemplo:

```
<dimension axis=row text=Gender label=Gender varName=gender>  
...<category text=Female label=Female string=f varName=gender>
```

- Para una variable numérica, hay un atributo de número en lugar de un atributo de cadena. El atributo de etiqueta está presente sólo si la variable o los valores tienen etiquetas definidas.
- Los elementos `<cell>` que contienen valores de casilla para los números contienen el atributo de **texto** y uno o más valores de atributo adicionales. Por ejemplo:

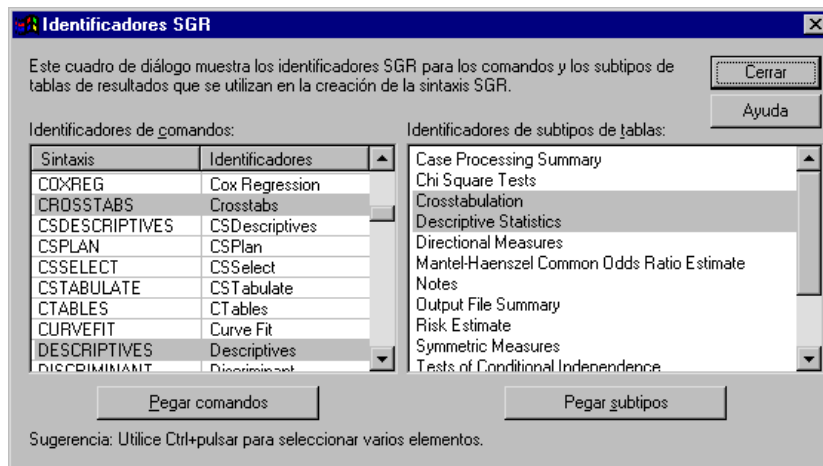
```
<cell text=45.6 number=45.569620253165 decimals=1/>
```

- El atributo de número es el valor numérico real sin redondear y el atributo de decimales indica el número de decimales mostrados en la tabla.
- Dado que las columnas se anidan en las filas, el elemento de categoría que identifica cada columna se repite para cada fila. Por ejemplo, dado que los estadísticos se muestran en las columnas, el elemento <category text=Frequency> aparece tres veces en XML: una vez para la fila de hombre, una vez para la fila de mujer y una vez para la fila total.

Identificadores SGR

El objetivo del cuadro de diálogo Identificadores SGR es ofrecerle asistencia en la escritura de la sintaxis del comando OMS. Se puede utilizar este cuadro de diálogo para pegar los identificadores de subtipos y comandos seleccionados en una ventana de sintaxis de comandos.

Figura 47-16
Cuadro de diálogo Identificadores SGR



Uso del cuadro de diálogo Identificadores SGR

- Elija en los menús:
Utilidades
Identificadores SGR...

- ▶ Seleccione uno o varios identificadores de comandos o de subtipos. Mantenga pulsada la tecla Ctrl si desea seleccionar varios identificadores en cada lista.
- ▶ Pulse en Pegar comandos y/o Pegar subtipos.
 - La lista de subtipos disponibles depende de los comandos seleccionados en ese momento. Si se seleccionan varios comandos, la lista de subtipos disponibles es la unión de todos los subtipos disponibles para cualquiera de los comandos seleccionados. Si no se selecciona ningún comando, en la lista aparecerán todos los subtipos.
 - Los identificadores se pegan en la posición actual del cursor dentro de la ventana de sintaxis de comandos designada. Si no hay abierta ninguna ventana de sintaxis de comandos, se abrirá automáticamente una nueva ventana de sintaxis.
 - Cuando se pega un identificador de subtipos y/o comandos, éste aparece entre comillas ya que la sintaxis del comando OMS exige que así sea.
 - Las listas de identificadores de las palabras clave `COMMANDS` y `SUBTYPES` deben ir entre corchetes, por ejemplo:

```
/IF COMMANDS=['Crosstabs' 'Descriptives']  
SUBTYPES=['Crosstabulation' 'Descriptive Statistics']
```

Copia de identificadores SGR desde los titulares del Visor

copiar y pegar identificadores de subtipos y comandos SGR desde el panel de titulares del Visor.

- ▶ Pulse con el botón derecho del ratón en la entrada del titular del elemento en el panel de titulares.
- ▶ En el menú contextual emergente, seleccione Copiar identificador de comandos SGR o Copiar subtipo de tablas SGR.

Este método presenta una diferencia respecto al cuadro de diálogo Identificadores SGR: el identificador copiado no se pega automáticamente en una ventana de sintaxis de comandos. Sencillamente se copia en el Portapapeles y, a continuación, podrá pegarlo donde desee. Como los valores de los identificadores de subtipos y comandos

son idénticos a los correspondientes valores de los atributos de subtipos y comandos de los resultados con formato XML (OXML), este método de copiar y pegar puede resultar muy útil para escribir transformaciones XSLT.

Copia de etiquetas SGR

En vez de identificadores, puede copiar etiquetas para utilizarlas con la palabra clave LABELS. Las etiquetas se pueden utilizar para diferenciar varios gráficos o varias tablas del mismo tipo en las que el texto del titular refleja algún atributo del objeto de resultados concreto como las etiquetas o los nombres de las variables. Hay, no obstante, ciertos factores que pueden afectar al texto de la etiqueta:

- Si está activado el procesamiento de segmentación del archivo, es posible que se añada a la etiqueta una identificación con el grupo de segmentación del archivo.
- Las etiquetas que incluyen información acerca de variables o valores se ven afectadas por la configuración de la presentación de nombres de variables/etiquetas y valores/etiquetas de valor del panel de titulares (menú Edición, Opciones, pestaña Etiquetas de los resultados).
- Las etiquetas se ven afectadas por el ajuste actual del idioma de los resultados (menú Edición, Opciones, pestaña General).

Para copiar etiquetas SGR:

- ▶ Pulse con el botón derecho del ratón en la entrada del titular del elemento en el panel de titulares.
- ▶ En el menú contextual emergente, seleccione Copiar etiqueta de SGR.

Al igual que ocurría con los identificadores de subtipos y comandos, las etiquetas deben ir entre comillas y toda la lista debe ir entre corchetes, por ejemplo:

```
/IF LABELS=['Employment Category' 'Education Level']
```


Administrador de acceso a bases de datos

El Administrador de acceso a bases de datos es una utilidad diseñada para simplificar las fuentes de datos demasiado extensas o confusas a fin de utilizarlas con el Asistente para bases de datos. Permite a los usuarios y a los administradores personalizar sus fuentes de datos de las formas siguientes:

- Creando alias para las tablas y los campos de la base de datos.
- Creando nombres de variable para los campos.
- Ocultando las tablas y los campos no pertinentes.

El Administrador de acceso a bases de datos no modifica su base de datos. Lo que hace realmente es generar archivos que guardan toda su información y que actúan como *vistas* de la base de datos.

Puede utilizar este Administrador para especificar hasta tres vistas diferentes por base de datos: Nivel de empresa, Nivel de departamento y Nivel personal. Tanto el Administrador como el Asistente para bases de datos reconocen estos archivos por los nombres siguientes:

- Nivel de empresa: *dba01.inf*
- Nivel de departamento: *dba02.inf*
- Nivel personal: *dba03.inf*

Cada archivo contiene información específica del nivel sobre un número cualquiera de fuentes de datos. Por ejemplo, su archivo *dba03.inf* puede contener información de la vista personal para una base de datos de contabilidad corporativa, para la base de datos de registro de horarios de su empresa y para una base de datos donde registre su colección de CD.

Cuando abra el Administrador, éste buscará la ruta de acceso de su sistema para estos archivos y mostrará automáticamente información de cualquier fuente de datos que haya configurado en las tres vistas.

Herencia y prioridades. Siempre que se utiliza el Asistente para bases de datos, éste presenta la vista de nivel más bajo de su fuente de datos que encuentra en la ruta de acceso de su sistema, donde los niveles son, del más alto al más bajo, el de empresa, el de departamento y el personal. El archivo de cada nivel contiene información sobre todas las fuentes de datos para ese nivel. Por ejemplo, su departamento de marketing tendrá un archivo, *dba02.inf*, con información sobre los alias de todas las vistas de base de datos establecidas para el departamento de marketing. Cada persona de este departamento tendrá un archivo, *dba03.inf*, que contenga vistas personalizadas de todas las bases de datos que utilice.

En el Administrador de acceso a bases de datos, los alias, los nombres de variable y el orden de ocultación se heredan desde el nivel superior hacia abajo.

Ejemplo. Si se oculta la tabla Regiones en el Nivel de empresa, no podrá verse en el Nivel de departamento ni en el Nivel personal. Esta tabla no se mostraría en el Asistente para bases de datos.

Ejemplo. El campo *JOBCAT* de la tabla VentasEmpleados no tiene un alias en el Nivel de empresa, pero tiene el alias Categorías de trabajo en el Nivel de departamento. Aparecerá como Categorías de trabajo en el Nivel personal. Además, si a esta tabla se le asignara el alias Información sobre empleados en el Nivel personal, el campo original (*VentasEmpleados.JOBCAT*) aparecería en el Asistente para bases de datos como '*Información sobre empleados.*'Categorías de trabajo'.

Para iniciar el Administrador de acceso a bases de datos, ejecute el archivo *spssdbca.exe*, que está instalado en el directorio de SPSS. Si desea obtener más información sobre el Administrador de acceso a bases de datos, consulte la Ayuda en pantalla.

Personalización de documentos HTML

Puede añadir automáticamente un código HTML personalizado a los documentos exportados al formato HTML, incluyendo:

- Títulos de documentos HTML
- Especificación de tipos de documentos
- Metaetiquetas y código de procesos (por ejemplo, JavaScript)
- Texto mostrado antes y después de los resultados exportados

Para añadir código HTML personalizado a los documentos de resultados exportados

- ▶ Abra el archivo *htmlfram.txt* ubicado en el directorio donde se instaló SPSS, en un editor de texto.
- ▶ Sustituya los comentarios de los “campos” en las líneas entre dos corchetes angulares de apertura (<<), por el texto o el código HTML que desee insertar en sus documentos HTML exportados.
- ▶ Guarde el archivo como un archivo de texto.

Nota: Si cambia el nombre o la ubicación del archivo de texto, deberá modificar el registro del sistema para que utilice el archivo para personalizar los resultados HTML exportados.

Contenido y formato del archivo de texto para HTML personalizado

El código HTML que desee añadir automáticamente a sus documentos HTML debe especificarse en un archivo de sólo texto que contenga seis campos, cada uno delimitado por dos corchetes angulares de apertura en la línea anterior (<<):

<<

Texto o código que desee insertar en la parte superior de los documentos antes de la especificación <HTML> (por ejemplo, comentarios que incluyan las especificaciones de los tipos de documentos)

<<

Texto empleado como título del documento (aparece en la barra de título)

<<

Metaetiquetas o código de proceso (por ejemplo, código JavaScript)

<<

Código HTML que modifica la etiqueta <BODY> (por ejemplo, código que especifique el color de fondo)

<<

Texto y/o código HTML que se inserta después de los resultados exportados (por ejemplo, una nota de copyright)

<<

Texto y/o código HTML que se inserta antes de los resultados exportados (por ejemplo, el nombre de la empresa, el logotipo, etc.)

Para utilizar otro archivo o ubicación para el código HTML personalizado

Si cambia el nombre o la ubicación del archivo *htmlfram.txt*, deberá modificar asimismo el registro del sistema para que utilice el archivo en los resultados HTML personalizados.

- ▶ En el menú Inicio de Windows, seleccione Ejecutar, escriba *regedit* y pulse en Aceptar.

- ▶ En el panel de la izquierda del Editor del registro de configuraciones, seleccione:
 - HHKEY_CURRENT_USER
 - Software o Programa
 - SPSS
 - SPSS para Windows
 - 13.0
 - SPSSWIN
- ▶ En el panel derecho, pulse dos veces en la cadena HTMLFormatFile.
- ▶ En Información del valor, escriba la ruta completa y el nombre del archivo de texto que contenga las especificaciones de HTML personalizadas (por ejemplo, *c:\misarchivos\htmlpersonal.txt*).

Muestra de archivo de texto para HTML personalizado

```
<<
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN>
<<
NVI, Inc.
<<
<META NAME=keywords CONTENT=gizmos, gadgets, gimcracks>
<<
bgcolor=#FFFFFF
<<
<H4 align=center>This page made possible by...
<br><br>
<IMG SRC=spss2.gif align=center></H4>
<<
<h2 align=center>NVI Sales</h2>
<h3 align=center>Regional Data</h3>
```

Muestra de origen HTML para HTML personalizado

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN>
<HTML>
<HEAD>
```

```
<TITLE>
NVI Sales, Inc.
</TITLE>
<META NAME=keywords CONTENT=gizmos, gadgets, gimcracks>
</HEAD>
<BODY bgcolor=#FFFFFF>
<h2 align=center>NVI Sales</h2>
<h3 align=center>Regional Data</h3>
```

[Resultado exportado]

```
<H4 align=center>This page made possible by...
<br><br>
<IMG SRC=spss2.gif align=center></H4>
</BODY>
</HTML>
```

Índice

- acceso a un servidor, 73
- Access (Microsoft), 26
- adición de etiquetas de grupo, 288
- agregación de datos, 200
 - agregación de funciones, 203
 - nombres y etiquetas de variable, 204
- agrupación de filas o columnas, 288
- ajuste, 301
 - control de la anchura de la columna para el texto ajustado, 301
 - etiquetas de variable y de valor, 93
- alfa de Cronbach
 - en Análisis de fiabilidad, 571, 573
- alineación , 95, 244, 311, 616
 - en el Editor de datos, 95
 - en las casillas, 311
 - resultados, 244, 616
- almacenamiento de archivos , 58, 61, 62, 77
 - archivos de datos, 61, 62, 77
 - archivos de datos de SPSS, 58
 - consultas del archivo de base de datos, 42
- almacenamiento de gráficos , 253, 260, 261, 262, 262, 262, 263, 263, 265, 626, 646
 - almacenamiento de gráficos interactivos con datos, 626
 - archivos BMP, 253, 260, 262, 646
 - archivos EPS, 253, 260, 263, 646
 - archivos JPEG, 253, 260, 261, 646
 - archivos PICT, 253, 260, 262, 646
 - archivos PNG, 262
 - archivos PostScript, 263
 - archivos TIFF, 262
 - archivos WMF, 253, 260, 265
 - formato WMF, 646
 - metarchivos, 253, 260, 646
- almacenamiento de resultados , 253, 257, 258, 274, 284, 284, 646, 658
 - almacenamiento de los resultados de borrador como texto, 284
 - formato de texto, 253, 258, 646
 - formato Excel, 253, 257
 - formato HTML, 253, 257, 646
 - Formato PowerPoint, 253, 257
 - formato Word, 253, 257
 - protección de contraseña, 274
 - publicación en Web, 658
 - resultados de borrador, 284
- almacenamiento de vistas de tablas pivote, 294
- análisis alfa, 479
- análisis de componentes principales, 471, 479
- análisis de conglomerados
 - Análisis de conglomerados de K-medias, 505
 - análisis de conglomerados jerárquico, 497
 - eficacia, 510
- Análisis de conglomerados de K-medias, 505
- almacenamiento de información de conglomerados, 512
 - conceptos básicos, 505
 - conglomerado de pertenencia, 512
 - criterios de convergencia, 511
 - distancias entre conglomerados, 512
 - eficacia, 510
 - ejemplos, 505
 - estadísticos, 505, 512
 - iteraciones, 511
 - métodos, 505
 - valores perdidos, 512
- Análisis de conglomerados en dos fases, 487
 - almacenamiento en el archivo de trabajo, 495
 - almacenamiento en un archivo externo, 495
 - estadísticos, 495
 - gráficos, 494
 - opciones, 491
- análisis de conglomerados jerárquico, 497
 - almacenamiento de nuevas variables, 503
 - casos de conglomerado, 497
 - conglomerado de pertenencia, 502, 503
 - dendrogramas, 503
 - diagramas de témpanos, 503
 - ejemplo, 497

- estadísticos, 497, 502
- historial de conglomerado, 502
- matrices de distancias, 502
- medidas de distancia, 501
- medidas de similaridad, 501
- métodos de conglomeración, 501
- orientación de los gráficos, 503
- transformación de medidas, 501
- transformación de valores, 501
- variables de conglomerado, 497
- análisis de datos , 13
 - pasos básicos, 13
- Análisis de fiabilidad , 571
 - coeficiente de correlación intraclase , 573
 - correlaciones y covarianzas inter-elementos, 573
 - descriptivos, 573
 - ejemplo, 571
 - estadísticos, 571, 573
 - funciones adicionales del comando, 575
 - Kuder-Richardson 20, 573
 - Prueba de aditividad de Tukey, 573
 - tabla de ANOVA, 573
 - T-cuadrado de Hotelling, 573
- análisis de la varianza
 - en ANOVA de un factor, 393
 - en Estimación curvilínea, 451
 - en Medias, 370
 - en Regresión lineal, 446
- análisis de respuestas múltiples
 - Frecuencias de respuestas múltiples, 548
 - tabla de contingencia, 550
 - Tablas de contingencia de respuestas múltiples, 550
 - tablas de frecuencias, 548
- análisis de segmentación de archivo, 205
- análisis de series temporales
 - predicción, 456
 - predicción de casos, 456
- Análisis discriminante, 459
 - almacenamiento de variables de clasificación, 468
 - análisis de casos de reserva, 660
 - coeficientes de la función, 464
 - criterios, 465
 - definición de rangos, 463
 - Distancia de Mahalanobis, 465
 - ejemplo, 459
 - estadísticos, 459, 464
 - estadísticos descriptivos, 464
 - exportación de información del modelo, 468
 - gráficos, 467
 - lambda de Wilks, 465
 - matrices, 464
 - matriz de covarianzas, 467
 - métodos de inclusión por pasos, 459
 - métodos discriminantes, 465
 - opciones de presentación, 465, 467
 - probabilidades previas, 467
 - selección de casos, 463
 - valores perdidos, 467
 - variables de agrupación, 459
 - variables independientes, 459
 - V de Rao, 465
- análisis factorial, 471
 - análisis de casos de reserva, 660
 - conceptos básicos, 471
 - convergencia, 479, 481
 - descriptivos, 478
 - ejemplo, 471
 - estadísticos, 471, 478
 - formato de presentación de los coeficientes, 484
 - gráficos de saturaciones, 481
 - métodos de extracción, 479
 - métodos de rotación, 481
 - puntuaciones factoriales, 483
 - selección de casos, 477
 - valores perdidos, 484
- análisis imagen, 479
- ancho de columna , 95, 301, 308, 627
 - control de la anchura máxima, 301
 - control de la anchura para el texto ajustado, 301
 - control de la anchura por defecto, 627
 - en el Editor de datos, 95
 - tablas pivote, 308
- años , 630
 - valores de dos dígitos, 630
- ANOVA
 - en ANOVA de un factor, 393
 - en Medias, 370
 - en MLG Univariante, 403
 - modelo, 407

- ANOVA de un factor, 393
 comparaciones múltiples, 397
 contrastes, 396
 contrastes polinómicos, 396
 contrastes post hoc, 397
 estadísticos, 400
 opciones, 400
 valores perdidos, 400
 variables del factor, 393
- apertura de archivos , 21, 21, 21, 21, 21, 21, 21, 22,
 23, 24, 25, 25, 26, 26, 44
 archivos de datos, 21, 22
 archivos de datos de texto, 44
 archivos de dBASE, 21, 25
 archivos de Excel, 21, 23
 archivos de hoja de cálculo, 21, 24
 archivos delimitados por tabuladores, 21
 archivos de Lotus 1-2-3, 21
 archivos SYSTAT, 21
- archivo activo, 66, 68, 68
 almacenamiento en la caché, 68
 archivo activo virtual, 66
 creación de un archivo activo temporal, 68
- archivo activo temporal, 68
- archivo de diario, 614
- archivos , 248
 adición de un archivo de texto al Visor, 248
 apertura, 21
- archivos BMP , 253, 260, 262, 646
 exportación de gráficos, 253, 260, 262, 646
- archivos de bases de datos, 25, 25, 26, 26, 26, 26,
 26, 26, 28, 29, 29, 31, 32, 34, 34, 37, 40, 42
 acceso a una base de datos, 28
 almacenamiento de consultas, 42
 cláusula Where, 34
 consultas de parámetros, 34, 37
 conversión de cadenas en variables numéricas, 40
 creación de relaciones, 31
 definición de variables, 40
 especificación de criterios, 34
 expresiones condicionales, 34
 lectura, 25, 26, 26, 26, 29
 Microsoft Access, 26
 muestreo aleatorio, 34
 pedir el valor al usuario, 37
 propiedades de la relación, 32
 seguridad de la base de datos, 28
 selección de campos de datos, 29
 selección de una fuente de datos, 26
 sintaxis de SQL, 42
 uniones entre tablas, 31, 32
 verificación de los resultados, 42
- archivos de datos , 21, 22, 44, 57, 57, 58, 61, 62,
 64, 68, 76, 77, 78, 214
 adición de comentarios, 608
 almacenamiento, 58, 61, 62
 almacenamiento de resultados como archivos de
 datos con formato SPSS, 695
 almacenamiento de subconjuntos de variables,
 64
 apertura, 21, 22
 información sobre el archivo, 57, 57
 información sobre el diccionario, 57, 57
 mejora del rendimiento para archivos grandes, 68
 protección, 65
 reestructuración, 214
 servidores remotos, 76, 77, 78
 texto, 44
 transposición, 193
 volteado, 193
- archivos de dBASE , 21, 25, 61, 62
 almacenamiento, 61, 62
 apertura, 21, 25
- archivos de Excel , 21, 23, 23, 61, 62, 635
 adición de elementos de menú para enviar datos
 a Excel, 635
 almacenamiento, 61, 62
 apertura, 21, 23
 lectura de nombres de variables, 23
 rangos de lectura, 23
- archivos de hoja de cálculo, 21, 23, 24, 65
 apertura, 24
 escritura de nombres de variable, 65
 lectura de nombres de variables, 23
 rangos de lectura, 23
- archivos delimitados por tabuladores , 21, 23, 44,
 61, 62, 65
 almacenamiento, 61, 62
 apertura, 21
 escritura de nombres de variable, 65
 lectura de nombres de variables, 23

- archivos de Lotus 1-2-3 , 21, 61, 62, 635
 - adición de elementos de menú para enviar datos a Lotus, 635
 - almacenamiento, 61, 62
 - apertura, 21
- archivos de SAS
 - almacenamiento, 61
 - apertura, 21
- archivos EPS , 253, 260, 263, 646
 - exportación de gráficos, 253, 260, 263, 646
- archivos JPEG , 253, 260, 261, 646
 - exportación de gráficos, 253, 260, 261, 646
- archivos PICT , 253, 260, 262, 646
 - exportación de gráficos, 253, 260, 262, 646
- archivos PNG , 253, 262
 - exportación de gráficos, 253, 262
- archivos portátiles
 - por defecto, 62
- archivos PostScript (encapsulado) , 253, 263, 646
 - exportación de gráficos, 253, 263, 646
- archivos SYSTAT , 21
 - apertura, 21
- archivos TIFF , 262
 - exportación de gráficos, 253, 260, 262, 646
- archivos WMF , 253, 260, 265, 646
 - exportación de gráficos, 253, 260, 265, 646
- Asesor estadístico, 13
- asignación de memoria
 - en Análisis de conglomerados en dos fases, 491
- asignación de rangos a los casos, 156
 - percentiles, 157
 - puntuaciones de Savage, 157
 - rangos fraccionales, 157
 - valores empatados, 159
- asimetría
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en el Informe de estadísticos en columnas, 566
 - en el Informe de estadísticos en filas, 559
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
- asociación lineal por lineal
 - en Tablas de contingencia, 354
- Aspectos de tabla, 298, 298, 299
 - aplicación, 298
 - creación, 299
- atributos de variable , 95, 96
 - copia y pegado, 95, 96
- automatización OLE, 659, 672, 675, 676, 678, 679
 - conceptos básicos, 672, 675, 676, 679
 - convenciones de denominación de variables, 675
 - métodos, 678
 - procesamiento con, 659
 - propiedades, 678
 - uso de objetos, 672, 676, 679
- autoprocesos, 633, 668, 670
 - archivo de autoproceso, 671
 - creación, 668
 - eventos de activación, 670
- autovalores
 - en Análisis factorial, 478, 479
 - en Regresión lineal, 446
- ayuda contextual, 290
 - búsqueda de definiciones de etiqueta en las tablas pivote, 290
- ayuda del botón derecho del ratón
 - de los cuadros de diálogo, 12
- ayuda en pantalla , 15
 - Asesor estadístico, 13
- barra de estado, 7
 - ocultación, 8
 - presentación, 8
- barras de herramientas , 637, 638, 639, 640, 641
 - creación, 637, 639
 - creación de nuevas herramientas, 640
 - edición de iconos de mapa de bits, 641
 - personalización, 637, 639
 - presentación en distintas ventanas, 638
 - presentación y ocultación, 637
- Bonferroni
 - en ANOVA de un factor, 397
 - en MLG, 413
- bordes , 277, 305, 306
 - presentación de bordes ocultos, 306
 - Visor de borrador, 277

- bordes de casillas , 277
 - Visor de borrador, 277
- botón Aceptar, 10
- botón Ayuda, 10
- botón Cancelar, 10
- botones , 641
 - edición de iconos de mapa de bits de las barras de herramientas, 641
- botón Pegar, 10
- botón Restablecer, 10

- caché, 68
 - archivo activo, 68
- cálculo de variables, 141
 - cálculo de variables de cadena nuevas, 144
- cambio del orden de filas y columnas, 287
- cambio del tamaño de los gráficos exportados, 260
- capas , 266, 290, 290, 292, 292, 301, 306
 - cambio, 292
 - creación, 290
 - en las tablas pivote, 290
 - en Tablas de contingencia, 354
 - impresión, 266, 301, 306
 - presentación, 290, 292
- casillas de las tablas pivote, 297, 303, 307, 308, 310, 311, 313, 314, 316
 - alineación, 311
 - anchuras, 308
 - formatos, 303
 - formatos de valores, 310
 - fuentes, 307
 - márgenes, 313
 - modificación de texto, 316
 - ocultación, 295
 - presentación, 295
 - sombreado, 314
 - titulares, 314
- casos , 101, 103, 104, 214
 - búsqueda de duplicados, 125
 - búsqueda en el Editor de datos, 103, 104
 - inserción de nuevos casos, 101
 - ordenación, 192
 - ponderación, 211
 - reestructuración en variables, 214
 - selección de subconjuntos, 206, 208, 210
- casos de reserva, 660
 - análisis, 660
- casos duplicados (registros)
 - búsqueda y filtrado, 125
- casos filtrados, 104
 - en el Editor de datos, 104
- categoría de referencia
 - en MLG, 410, 411
- CCI
 - Véase coeficiente de correlación intraclase, 573
- C de Dunnnett
 - en ANOVA de un factor, 397
 - en MLG, 413
- centrado de resultados, 244, 616
- chi-cuadrado de la razón de verosimilitud
 - en Tablas de contingencia, 354
- chi-cuadrado de Pearson
 - en Tablas de contingencia, 354
- clasificación
 - en Curva COR..., 603
- coeficiente alfa
 - en Análisis de fiabilidad, 571, 573
- coeficiente de contingencia
 - en Tablas de contingencia, 354
- coeficiente de correlación de los rangos
 - en Correlaciones bivariadas, 421
- coeficiente de correlación de Spearman
 - en Correlaciones bivariadas, 421
 - en Tablas de contingencia, 354
- coeficiente de correlación intraclase (CCI)
 - en Análisis de fiabilidad, 573
- coeficiente de correlación r
 - en Correlaciones bivariadas, 421
 - en Tablas de contingencia, 354
- coeficiente de dispersión (CDD)
 - en Estadísticos de la razón, 587
- coeficiente de incertidumbre
 - en Tablas de contingencia, 354
- coeficiente de variación (CDV)
 - en Estadísticos de la razón, 587
- coeficiente eta
 - en Medias, 370
 - en Tablas de contingencia, 354

- coeficientes beta
 - en Regresión lineal, 446
- coeficientes de regresión
 - en Regresión lineal, 446
- colapsar categorías de datos, 128
- colores en las tablas pivote, 305, 307, 314
 - bordes, 305
 - fondo de la casilla, 314
 - fuelle, 307
 - máscara de la casilla, 314
- columnas, 308, 316
 - cambio de la anchura en tablas pivote, 308
 - selección en tablas pivote, 316
- columna total
 - en informes, 567
- comando EXECUTE
 - pegado desde cuadros de diálogo, 328
- comparación de grupos
 - en Cubos OLAP, 379
- comparación de variables
 - en Cubos OLAP, 379
- comparaciones múltiples
 - en ANOVA de un factor, 397
- comparaciones múltiples post hoc, 397
- conceptos básicos, 729
- conglomeración
 - selección de procedimientos, 485
- conjuntos de variables, 609, 610
 - definición, 609
 - utilización, 610
- consejos sobre los procesos, 659, 664, 666, 672, 674, 676, 678, 679, 680, 683, 683, 688
 - adición de una descripción, 683
 - cuadros de diálogo personalizados, 683
 - declaraciones de variables, 674
 - depuración, 688
 - funcionamiento de los procesos, 672
 - obtención de objetos de automatización, 676
 - procedimientos, 680
 - procesos de iniciación, 666
 - propiedades y métodos, 678
 - ventana de proceso, 664
 - visor de objetos, 679
- construcción de términos, 408
- contrastes
 - en ANOVA de un factor, 396
 - en MLG, 410, 411
 - contrastes de desviación
 - en MLG, 410, 411
 - contrastes de diferencia
 - en MLG, 410, 411
 - Contrastes de Helmert
 - en MLG, 410, 411
 - contrastes de linealidad
 - en Medias, 370
 - contrastes polinómicos
 - en ANOVA de un factor, 396
 - en MLG, 410, 411
 - contrastes repetidos
 - en MLG, 410, 411
 - contrastes simples
 - en MLG, 410, 411
- control de página
 - en el informe de estadísticos en columnas, 568
 - en informes de estadísticos en filas, 561
- convergencia
 - en Análisis de conglomerados de K-medias, 511
 - en Análisis factorial, 479, 481
- copia , 243, 249
 - gráficos, 249
 - resultados, 243
 - tablas pivote, 249
- corrección por continuidad de Yates
 - en Tablas de contingencia, 354
- Correlación de Pearson
 - en Correlaciones bivariadas, 421
 - en Tablas de contingencia, 354
- correlaciones
 - bivariadas, 421
 - de orden cero, 428
 - en Correlaciones bivariadas, 421
 - en Correlaciones parciales, 425
 - en Tablas de contingencia, 354
- Correlaciones bivariadas
 - coeficientes de correlación, 421
 - estadísticos, 423
 - funciones adicionales del comando, 424
 - niveles de significación, 421
 - opciones, 423
 - valores perdidos, 423

- correlaciones de orden cero
 - en Correlaciones parciales, 428
- Correlaciones parciales, 425
 - correlaciones de orden cero, 428
 - en Regresión lineal, 446
 - estadísticos, 428
 - opciones, 428
 - valores perdidos, 428
- cuadros de diálogo , 11, 609, 610, 614, 614, 683, 685
 - controles, 10
 - definición de conjuntos de variables, 609
 - especificaciones opcionales, 10
 - información sobre la variable, 11
 - obtención de ayuda, 12
 - orden de presentación de variables, 614
 - presentación de etiquetas de variable, 9, 614
 - presentación de nombres de variable, 9, 614
 - procesamiento, 683, 685
 - reordenación de listas de destino, 611
 - selección de variables, 11
 - subcuadros de diálogo, 10
 - utilización de conjuntos de variables, 610
 - variables, 8
- cuartiles
 - en Frecuencias, 334
- Cubos OLAP, 373
 - estadísticos, 376
 - títulos, 380
- curtosis
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en el Informe de estadísticos en columnas, 566
 - en el Informe de estadísticos en filas, 559
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
- Curva COR, 603
 - estadísticos y gráficos, 606
- d
 - en Tablas de contingencia, 354
- Data List, 66
 - frente al comando Get Data, 66
- datos categóricos, 113
 - conversión de datos de intervalo en categorías discretas, 128
- datos con formato CSV, 44
- datos de cadena , 98
 - introducción de datos, 98
- datos delimitados por comas, 44
- datos delimitados por espacios, 44
- datos de serie temporal
 - creación de nuevas variables de serie temporal, 183
 - definición de variables de datos, 181
 - funciones de transformación, 185
 - sustitución de valores perdidos, 186
 - transformaciones de los datos, 181
- datos ponderados , 237
 - y de datos reestructurados, 237
- d de Somers
 - en Tablas de contingencia, 354
- declaraciones de variables, 674, 675
 - convenciones de denominación, 675
 - en procesos, 674, 675
- definición de variables, 86, 89, 92, 92, 93, 95, 95, 96, 96, 108
 - aplicación de un diccionario de datos, 116
 - copia y pegado de atributos, 95, 96
 - etiquetas de valor, 92, 108
 - etiquetas de variable, 92
 - plantillas, 95, 96
 - tipos de datos, 89
 - valores perdidos, 93
- definir conjuntos de respuestas múltiples, 546
 - categorías, 546
 - dicotomías, 546
 - etiquetas del conjunto, 546
 - nombres del conjunto, 546
- dendrogramas
 - en Análisis de conglomerados jerárquico, 503
- depuración de procesos, 688, 689
 - avance por los procesos, 688
 - panel de depuración, 689
 - puntos de interrupción, 688
- descomposición jerárquica , 408

- Descriptivos, 339
 - almacenamiento de puntuaciones z, 339
 - estadísticos, 341
 - orden de presentación, 341
- desplazamiento de filas y columnas, 287
- desviación absoluta promedio (DAP)
 - en Estadísticos de la razón, 587
- desviación típica
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en el Informe de estadísticos en columnas, 566
 - en el Informe de estadísticos en filas, 559
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en MLG Univariante, 418
 - en Resumir, 364
- DfAjuste
 - en Regresión lineal, 443
- DfBeta
 - en Regresión lineal, 443
- diagnósticos de colinealidad
 - en Regresión lineal, 446
- diagnósticos por caso
 - en Regresión lineal, 446
- diagramas de caja
 - comparación de niveles del factor, 348
 - comparación de variables, 348
 - en Explorar, 348
- diagramas de dispersión
 - en Regresión lineal, 441
- diagramas de dispersión por nivel
 - en Explorar, 348
 - en MLG Univariante, 418
- diagramas de témpanos
 - en Análisis de conglomerados jerárquico, 503
- diario de la sesión, 614
- diccionario, 57, 57
- diccionario de datos
 - aplicar desde otro archivo, 116
- diferencia de configuración
 - en Distancias, 431
- diferencia de tamaño
 - en Distancias, 431
- diferencia honestamente significativa de Tukey
 - en ANOVA de un factor, 397
 - en MLG, 413
- diferencial relativo al precio (DRP)
 - en Estadísticos de la razón, 587
- diferencia menos significativa
 - en ANOVA de un factor, 397
 - en MLG, 413
- diferencias entre grupos
 - en Cubos OLAP, 379
- diferencias entre variables
 - en Cubos OLAP, 379
- directorio temporal , 614, 614
 - definir ubicación en modo local, 614
 - variable de entorno SPSSTMPDIR, 614
- distancia de bloques
 - en Distancias, 431
- distancia de Chebychev
 - en Distancias, 431
- Distancia de Cook
 - en MLG, 416
 - en Regresión lineal, 443
- Distancia de Mahalanobis
 - en Análisis discriminante, 465
 - en Regresión lineal, 443
- distancia de Minkowski
 - en Distancias, 431
- Distancia euclídea
 - en Distancias, 431
- distancia euclídea al cuadrado
 - en Distancias, 431
- Distancias, 429
 - cálculo de distancias entre casos, 429
 - cálculo de distancias entre variables, 429
 - ejemplo, 429
 - estadísticos, 429
 - medidas de disimilaridad, 431
 - medidas de similaridad, 432
 - transformación de medidas, 431, 432
 - transformación de valores, 431, 432
- división
 - división entre columnas del informe, 567
- divisor de paneles
 - Editor de datos, 104
- divisor de ventanas
 - Editor de datos, 104

- DMS de Fisher
 - en MLG, 413
- edición de datos, 100, 100, 101
- Editor de datos, 83, 86, 95, 95, 97, 98, 98, 99, 100, 100, 101, 101, 102, 103, 103, 103, 104, 104, 104, 106, 635
 - alineación, 95
 - ancho de columna, 95
 - búsqueda de casos, 103, 104
 - cambiar tipo de datos, 103
 - casos filtrados, 104
 - definición de variables, 86
 - desplazamiento de variables, 103
 - edición de datos, 100, 100, 101
 - envío de datos a otras aplicaciones, 635
 - impresión, 106
 - inserción de nuevas variables, 102
 - inserción de nuevos casos, 101
 - introducción de datos, 97
 - introducción de datos no numéricos, 98
 - introducción de datos numéricos, 98
 - opciones de presentación, 104
 - restricciones de los valores de datos, 99
 - varias vistas/paneles, 104
 - Vista de datos, 84
- eliminación de etiquetas de grupo, 288
- eliminación de resultados, 243
- eliminación de varios comandos EXECUTE en
 - archivos de sintaxis, 328
- eliminación hacia atrás
 - en Regresión lineal, 439
- encabezados, 269, 270
- entrada de datos, 97
- error típico
 - en Curva COR..., 606
 - en Descriptivos, 341
 - en Explorar, 347
 - en Frecuencias, 334
 - en MLG, 416, 418
 - error típico de la asimetría
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
 - error típico de la curtosis
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
 - error típico de la media
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- escala , 87
 - en Análisis de fiabilidad, 571
 - en Escalamiento multidimensional, 577
 - nivel de medida, 87, 113
- escalamiento de tablas pivote, 301, 306
- Escalamiento multidimensional , 577
 - condicionalidad, 581
 - creación de matrices de distancias, 580
 - critérios, 582
 - definición de la forma de los datos, 579
 - dimensiones, 581
 - ejemplo, 577
 - estadísticos, 577
 - funciones adicionales del comando, 583
 - medidas de distancia, 580
 - modelos de escalamiento, 581
 - niveles de medida, 581
 - opciones de presentación, 582
 - transformación de valores, 580
- espacio en disco, 66, 68
 - temporal, 66, 68
- espacio temporal en disco, 66, 68
- estadístico de Brown-Forsythe
 - en ANOVA de un factor, 400
- estadístico de Cochran
 - en Tablas de contingencia, 354
- estadístico de la mediana
 - en Cubos OLAP, 376
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364

- estadístico del rango
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en Estadísticos de la razón, 587
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
- estadístico de Mantel-Haenszel
 - en Tablas de contingencia, 354
- estadístico de Welch
 - en ANOVA de un factor, 400
- estadístico Durbin-Watson
 - en Regresión lineal, 446
- estadístico R
 - en Medias, 370
 - en Regresión lineal, 446
- Estadísticos de la razón, 585
 - estadísticos, 587
- estadísticos descriptivos
 - en Análisis de conglomerados en dos fases, 495
 - en Descriptivos, 339
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
 - en MLG Univariante, 418
 - en Resumir, 364
- Estimación curvilínea, 451
 - almacenamiento de intervalos de pronóstico, 456
 - almacenamiento de residuos, 456
 - almacenamiento de valores pronosticados, 456
 - análisis de la varianza, 451
 - inclusión de constante, 451
 - modelos, 455
 - predicción, 456
- estimaciones de Blom, 157
- estimaciones de los parámetros
 - en MLG Univariante, 418
- estimaciones de potencia
 - en MLG Univariante, 418
- estimaciones de proporción
 - en la asignación de rangos a los casos, 157
- estimaciones de Rankit, 157
- estimaciones de tamaño de efecto
 - en MLG Univariante, 418
- estimaciones de Tukey, 157
- estimaciones de Van der Waerden, 157
- estimador bponderado de Tukey
 - en Explorar, 347
- estimador en onda de Andrews
 - en Explorar, 347
- Estimadores robustos centrales
 - en Explorar, 347
- estimador-M de Huber
 - en Explorar, 347
- Estimador-M redescendente de Hampel
 - en Explorar, 347
- estudio de control de casos
 - Prueba T para muestras relacionadas, 385
- estudio de pares relacionados
 - en Prueba T para muestras relacionadas, 385
- eta-cuadrado
 - en Medias, 370
 - en MLG Univariante, 418
- etiquetas, 288, 288
 - eliminación, 288
 - frente a nombres de subtipos en SGR, 702
 - inserción de etiquetas de grupo, 288
- etiquetas de grupo, 288
- etiquetas de valor, 92, 99, 104, 108, 619
 - aplicación a varias variables, 114
 - copia, 114
 - en archivos de datos fusionados, 197
 - en el Editor de datos, 104
 - en el panel de titulares, 619
 - en las tablas pivote, 619
 - inserción de saltos de línea, 93
 - uso para entrada de datos, 99
- etiquetas de variable , 92, 614, 619
 - de los cuadros de diálogo, 9, 614
 - en archivos de datos fusionados, 197
 - en el panel de titulares, 619
 - en las tablas pivote, 619
 - inserción de saltos de línea, 93
- eventos de activación, 670
 - autoproscesos, 670
- exclusión de resultados del Visor con SGR, 710
- Explorar, 343
 - estadísticos, 347
 - gráficos, 348
 - opciones, 350
 - transformaciones de potencia, 349
 - valores perdidos, 350

- exportación de datos , 635
 - adición de elementos de menú para exportar datos, 635
- exportación de gráficos, 253, 260, 261, 262, 262, 262, 263, 263, 265, 643, 646
 - producción automatizada, 643
 - tamaño del gráfico, 260
- exportación de resultados, 253, 257, 258, 646, 658
 - formato Excel, 253, 257
 - formato HTML, 257
 - Formato PowerPoint, 253
 - formato Word, 253, 257
 - publicación en Web, 658
 - SGR, 695
- extremos
 - en Explorar, 347

- factor de inflación de la varianza
 - en Regresión lineal, 446
- factorización de ejes principales, 479
- fiabilidad de dos mitades
 - en Análisis de fiabilidad, 571, 573
- fiabilidad de Spearman-Brown
 - en Análisis de fiabilidad, 573
- filas, 316
 - selección en tablas pivote, 316
- F múltiple de Ryan-Einot-Gabriel-Welsch
 - en ANOVA de un factor, 397
 - en MLG, 413
- formato , 277
 - columnas en informes, 558
 - resultados de borrador, 277
- formato con comas, 89, 91
- Formato de archivo de datos de SPSS
 - envío de resultados a un archivo de datos, 704, 711
- formato de archivo SAV
 - envío de resultados a un archivo de datos de SPSS, 704, 711
- formato de cadena, 89
- formato de columnas, 311
 - cambio en las tablas pivote, 311
- formato de dólar, 89, 91
- formato de puntos, 89, 91
- formato Excel
 - exportación de resultados, 253, 257
- formato fijo, 44
- formato HTML , 253, 257, 646, 731
 - adición de código personalizado, 731
 - exportación de resultados, 253, 257, 646
- formato libre, 44
- formato numérico, 89, 91
- Formato PowerPoint
 - exportación de resultados, 253
- formatos de entrada, 91
- formatos de fecha, 89, 91, 630
 - años de dos dígitos, 630
- formatos de moneda, 631
- formatos de moneda personalizados, 89, 631
- formatos de presentación, 91
- formato Word
 - exportación de resultados, 253, 257
- Frecuencias, 331
 - estadísticos, 334
 - formatos, 337
 - gráficos, 336
 - orden de presentación, 337
 - supresión de tablas, 337
- frecuencias de los conglomerados
 - en Análisis de conglomerados en dos fases, 495
- Frecuencias de respuestas múltiples, 548
 - valores perdidos, 548
- fuentes , 104, 247, 282, 307, 307
 - colores, 307
 - en el Editor de datos, 104
 - en el panel de resultados, 247
 - en las casillas, 307
 - en Visor de borrador, 282
- función de adelanto, 185
- función de diferencia, 185
- función de diferencia estacional, 185
- función de media móvil anterior, 185
- función de media móvil centrada, 185
- función de medianas móviles, 185
- función de retardo, 185
- función de suavizado, 185
- función de suma acumulada, 185
- funciones, 144
 - tratamiento de los valores perdidos, 145

- fusión de archivos de datos
 - archivos con casos distintos, 194
 - archivos con distintas variables, 197
 - cambio de nombre de las variables, 196
 - información sobre el diccionario, 197
- gamma
 - en Tablas de contingencia, 354
- gamma de Goodman y Kruskal
 - en Tablas de contingencia, 354
- Get Data, 66
 - frente al comando Data List, 66
 - frente al comando Get Capture, 66
- gráficos , 241, 249, 249, 251, 253, 621, 646
 - conceptos básicos, 589
 - copia, 249
 - copia en otras aplicaciones, 249
 - creación, 589
 - en Curva COR..., 603
 - etiquetas de caso, 451, 599
 - exportación, 253, 646
 - modificación, 591
 - notas al pie, 595
 - ocultación, 241
 - pegado en otras aplicaciones, 251
 - plantillas, 601, 621
 - relación de aspecto, 621
 - subtítulos, 595
 - títulos, 595
 - valores perdidos, 597
- gráficos de barras
 - en Frecuencias, 336
- gráficos de importancia
 - en Análisis de conglomerados en dos fases, 494
- gráficos de la importancia de las variables
 - en Análisis de conglomerados en dos fases, 494
- gráficos de los residuos
 - en MLG Univariante, 418
- gráficos de perfil
 - en MLG, 412
- gráficos de probabilidad normal
 - en Explorar, 348
 - en Regresión lineal, 441
- gráficos de saturaciones
 - en Análisis factorial, 481
- gráficos de sectores
 - en Frecuencias, 336
 - títulos, 596
- gráficos de tallo y hojas
 - en Explorar, 348
- gráficos interactivos , 249, 250, 626, 626
 - almacenamiento de datos con gráficos, 626
 - copia en otras aplicaciones, 249
 - incrustación como objetos ActiveX, 250
 - opciones, 626
- gráficos normales sin tendencia
 - en Explorar, 348
- gráficos parciales
 - en Regresión lineal, 441
- GT2 de Hochberg
 - en ANOVA de un factor, 397
 - en MLG, 413
- H de Kruskal-Wallis
 - en Pruebas para dos muestras independientes, 537
- histogramas
 - en Explorar, 348
 - en Frecuencias, 336
 - en Regresión lineal, 441
- iconos , 641
 - edición de iconos de mapa de bits de las barras de herramientas, 641
- identificadores de comandos, 701
- idioma
 - cambio del idioma de los resultados, 614
- impresión , 106, 266, 267, 269, 270, 272, 282, 283, 301, 306, 318
 - capas, 266, 301, 306
 - control de las rupturas de tabla, 318
 - datos, 106
 - encabezados y pies, 269, 270
 - escalamiento de tablas, 301, 306

- espacio entre los elementos de resultados, 272
- gráficos, 266
- números de páginas, 272
- preparar página, 269
- presentación preliminar, 267
- resultados de borrador, 282, 283
- resultados de texto, 266
- tablas pivote, 266
- tamaño del gráfico, 272
- incrustación de gráficos interactivos, 250
- incrustación de tablas pivote, 250

- índice de concentración
 - en Estadísticos de la razón, 587
- información sobre el archivo, 57, 57
- información sobre la variable, 607
- informe de estadísticos en columnas, 564
- Informe de estadísticos en columnas, 564
 - columnas totales, 567
 - control de página, 568
 - diseño de página, 562
 - formato de columnas, 558
 - funciones adicionales del comando, 570
 - numeración de páginas, 569
 - subtotales, 568
 - total final, 569
 - valores perdidos, 569
- Informe de estadísticos en filas, 555
 - columnas de datos, 555
 - columnas de ruptura, 555
 - control de página, 560
 - diseño de página, 562
 - espaciado de ruptura, 560
 - formato de columnas, 558
 - funciones adicionales del comando, 570
 - numeración de páginas, 561
 - ordenación de secuencias, 555
 - pies, 563
 - títulos, 563
 - valores perdidos, 561
 - variables en los títulos, 563
- informes
 - columnas totales, 567
 - comparación de columnas, 567
 - división de valores de las columnas, 567
 - informe de estadísticos en columnas, 564
 - informes de estadísticos en filas, 555
 - multiplicación de valores de las columnas, 567
 - totales compuestos, 567
- inserción de etiquetas de grupo, 288
- intervalos de confianza
 - almacenamiento en Regresión lineal, 443
 - en ANOVA de un factor, 400
 - en Curva COR..., 606
 - en Explorar, 347
 - en MLG, 410, 418
 - en Pruebas t para muestras independientes, 385
 - en Prueba T para muestras relacionadas, 388
 - en Prueba t para una muestra, 391
 - en Regresión lineal, 446
- intervalos de pronóstico
 - almacenamiento en Estimación curvilínea, 456
 - almacenamiento en Regresión lineal, 443
- introducción de datos, 97, 98, 98, 99
 - no numéricos, 98
 - numéricos, 98
 - uso de etiquetas de valor, 99
- iteraciones
 - en Análisis de conglomerados de K-medias, 511
 - en Análisis factorial, 479, 481

- justificación , 244, 616
 - resultados, 244, 616

- kappa
 - en Tablas de contingencia, 354
- kappa de Cohen
 - en Tablas de contingencia, 354
- KR20
 - en Análisis de fiabilidad, 573
- Kuder-Richardson 20
 - en Análisis de fiabilidad, 573

- lambda
 - en Tablas de contingencia, 354
- lambda de Goodman y Kruskal
 - en Tablas de contingencia, 354
- lambda de Wilks
 - en Análisis discriminante, 465
- lenguaje de comandos, 321
- líneas de cuadrícula, 306
 - tablas pivote, 306
- listado de casos, 361
- listas de destino, 611
- listas de variables, 611
 - reordenación de listas de destino, 611

- márgenes, 269, 313
 - en las casillas, 313
- márgenes de la página, 269
- matriz de configuración
 - en Análisis factorial, 471
- matriz de correlaciones
 - en Análisis discriminante, 464
 - en Análisis factorial, 471, 478
- matriz de covarianzas
 - en Análisis discriminante, 464, 467
 - en MLG, 416
 - en Regresión lineal, 446
- matriz de transformación
 - en Análisis factorial, 471
- máxima verosimilitud
 - en Análisis factorial, 479
- máximo
 - comparación de columnas del informe, 567
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
- media
 - de varias columnas del informe, 567
 - en ANOVA de un factor, 400
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en el Informe de estadísticos en columnas, 566
 - en el Informe de estadísticos en filas, 559
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
 - subgrupo, 373
- media armónica
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- media geométrica
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- mediana agrupada
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- media ponderada
 - en Estadísticos de la razón, 587
- media recortada
 - en Explorar, 347
- medias
 - subgrupo, 367
- Medias, 367, 660
 - estadísticos, 370
 - opciones, 370
 - resultados en ASCII, 660
- medias de grupo, 367, 373
- medias de subgrupo, 367, 373
- medias marginales estimadas
 - en MLG Univariante, 418
- medias observadas
 - en MLG Univariante, 418
- medida de disimilaridad de Lance y Williams, 431
 - en Distancias, 431
- medida de distancia de chi-cuadrado
 - en Distancias, 431
- medida de distancia de phi cuadrado
 - en Distancias, 431
- medidas de dispersión
 - en Descriptivos, 341
 - en Estadísticos de la razón, 587

- en Explorar, 347
 - en Frecuencias, 334
- medidas de distancia
 - en Análisis de conglomerados jerárquico, 501
 - en Distancias, 431
- medidas de distribución
 - en Descriptivos, 341
 - en Frecuencias, 334
- medidas de similaridad
 - en Análisis de conglomerados jerárquico, 501
 - en Distancias, 432
- medidas de tendencia central
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
- memoria, 614
- menús, 7, 635
 - personalización, 635
- metarchivos , 253, 260, 646
 - exportación de gráficos, 253, 260, 646
- métodos, 678
 - objetos de automatización OLE, 678
- métodos de selección, 316
 - selección de filas y columnas en tablas pivote, 316
- Microsoft Access, 26
- mínimo
 - comparación de columnas del informe, 567
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en Estadísticos de la razón, 587
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
- mínimos cuadrados generalizados
 - en Análisis factorial, 479
- mínimos cuadrados no ponderados
 - en Análisis factorial, 479
- mínimos cuadrados ponderados
 - en Regresión lineal, 435
- MLG
 - almacenamiento de matrices, 416
 - almacenamiento de variables, 416
 - contrastes post hoc, 413
 - gráficos de perfil, 412
 - modelo, 407
 - suma de cuadrados, 407
- MLG Univariante, 403, 419
 - contrastes, 410, 411
 - diagnósticos, 418
 - medias marginales estimadas, 418
 - opciones, 418
 - presentación, 418
- moda
 - en Frecuencias, 334
- modelo compuesto
 - en Estimación curvilínea, 455
- modelo cuadrático
 - en Estimación curvilínea, 455
- modelo cúbico
 - en Estimación curvilínea, 455
- modelo de crecimiento
 - en Estimación curvilínea, 455
- modelo de curva S
 - en Estimación curvilínea, 455
- modelo de Guttman
 - en Análisis de fiabilidad, 571, 573
- modelo de potencia
 - en Estimación curvilínea, 455
- modelo estrictamente paralelo
 - en Análisis de fiabilidad, 571, 573
- modelo exponencial
 - en Estimación curvilínea, 455
- modelo inverso
 - en Estimación curvilínea, 455
- modelo lineal
 - en Estimación curvilínea, 455
- modelo logarítmico
 - en Estimación curvilínea, 455
- modelo logístico
 - en Estimación curvilínea, 455
- modelo paralelo
 - en Análisis de fiabilidad, 571, 573
- modelos factoriales completos
 - en MLG, 407
- modelos personalizados
 - en MLG, 407
- modificadores de la línea de comando , 655
 - Unidad de producción, 655

- modo distribuido, 71, 71, 73, 74, 76, 77, 78, 80, 81, 651
 - acceso a un archivo de datos, 76, 78
 - almacenamiento de archivos de datos, 77
 - procedimientos disponibles, 80
 - rutas de acceso UNC, 81
 - Unidad de producción, 651
- muestra aleatoria , 34
 - archivos de bases de datos, 34
 - selección, 209
 - semilla de aleatorización, 146
- muestras relacionadas, 533, 541
- muestreo
 - muestra aleatoria, 209
- multiplicación
 - multiplicación entre columnas del informe, 567
- Newman-Keuls
 - en MLG, 413
- nivel de medida , 87, 87, 113, 113
 - definición, 87, 87
- nombres , 74
 - servidores, 74
- nominal , 87
 - nivel de medida, 87, 113
- notación científica, 89, 311, 614
 - en las tablas pivote, 311
 - supresión en resultados, 614
- notas al pie, 302, 315, 315, 317
 - adición a una tabla, 317
 - en los gráficos, 595
 - marcadores, 302, 315
 - renumerar, 315
- nuevas funciones
 - SPSS 13,0, 2
- numeración de páginas, 269, 272
 - en el informe de estadísticos en columnas, 569
 - en informes de estadísticos en filas, 561
- número de casos
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- número máximo de ramas
 - en Análisis de conglomerados en dos fases, 491
- números de puerto, 74
- objetos, 672, 675, 676, 679
 - conceptos básicos, 672, 675
 - convenciones de denominación de variables, 675
 - uso en procesos, 672, 676, 679
- objetos ActiveX, 250
- objetos de automatización, 672, 675, 676, 678, 679
 - convenciones de denominación de variables, 675
 - métodos, 678
 - propiedades, 678
 - tipos, 675
 - uso en procesos, 672, 676, 679
 - visor de objetos, 679
- ocultación , 241, 241, 242, 296, 297, 297, 297, 637
 - barras de herramientas, 637
 - etiquetas de dimensión, 297
 - filas y columnas, 296
 - notas al pie, 297
 - resultados, 241, 241
 - resultados de un procedimiento, 242
 - textos al pie, 297
 - títulos, 297
- ocultación (exclusión) de resultados del Visor con SGR, 710
- opciones , 614, 614, 616, 617, 619, 621, 626, 627, 630, 630, 631, 633
 - años de dos dígitos, 630
 - aspecto de tablas pivote, 627
 - datos, 630
 - directorio temporal, 614
 - etiquetas de los resultados, 619
 - generales, 614
 - gráficos, 621
 - gráficos interactivos, 626
 - moneda, 631
 - procesos, 633
 - Visor, 616
 - Visor de borrador, 617
- opciones del gráfico, 621
- ordenación de casos, 192

- orden de presentación, 287, 287
- ordinal , 87
 - nivel de medida, 87, 113
- OXML, 726

- pares de variables , 214
 - creación, 214
- pegado , 251, 251, 251, 252
 - gráficos, 251
 - objetos especiales, 252
 - tablas pivote, 251, 251
 - tablas pivote como tablas, 251
- percentiles
 - en Explorar, 347
 - en Frecuencias, 334
- permisos, 80
- personalizadas, 318
 - control de las rupturas de tabla, 318
- pestaña Immediate, 689
 - ventana de proceso, 689
- pestaña Loaded, 689
 - ventana de proceso, 689
- pestaña Stack, 689
 - ventana de proceso, 689
- pestaña Watch, 689
 - ventana de proceso, 689
- phi
 - en Tablas de contingencia, 354
- pies, 269, 270
- pivotado
 - control con SGR para resultados exportados, 717
- plantillas , 95, 96, 621
 - definición de variables, 95, 96
 - en los gráficos, 601, 621
 - uso de un archivo de datos externo como plantilla, 116
- ponderación de casos, 211
 - ponderaciones fraccionarias en Tablas de contingencia, 211
- porcentajes
 - en Tablas de contingencia, 358
- porcentajes de fila
 - en Tablas de contingencia, 358
- porcentajes de la columna
 - en Tablas de contingencia, 358
- porcentajes totales
 - en Tablas de contingencia, 358
- por defecto , 86, 614
 - ajuste de los nombres largos de variable en los resultados, 86
 - archivos portátiles, 62
 - de los cuadros de diálogo, 9, 614
 - generados por SGR, 719
 - nombres de variable de casos mixtos, 86
 - reglas, 86
 - truncado de nombres de variable largos en versiones anteriores, 62
- PowerPoint, 257
 - exportación de resultados como PowerPoint, 257
- predicción
 - en Estimación curvilínea, 456
- preparar página, 269, 270, 272
 - encabezados y pies, 270
 - tamaño del gráfico, 272
- presentación , 241, 296, 297, 297, 297, 637
 - barras de herramientas, 637
 - etiquetas de dimensión, 297
 - filas o columnas, 296
 - notas al pie, 297
 - resultados, 241
 - textos al pie, 297
 - títulos, 297
- primera
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- procedimientos, 680
 - procesos, 680
- procedimientos de función, 680
- procedimientos de subrutinas, 680
- procedimientos globales, 633, 682
- procesos , 633, 635, 640, 659, 659, 660, 661, 663, 666, 668, 668, 683, 683, 688
 - adición a menús, 635
 - adición de una descripción, 683
 - archivo de autoprosceso, 633, 671
 - archivo de procedimientos globales, 633, 682
 - autoproscesos, 661, 668, 671
 - conceptos básicos, 659

- creación, 663, 668
- cuadros de diálogo, 683, 685
- declaración de variables, 674, 675
- depuración, 688, 689
- ejecución, 659
- ejecución mediante los botones de la barra de herramientas, 640
- mediante sintaxis de comandos, 691, 691, 693
- procesos de iniciación, 666
- uso de objetos de automatización, 672, 675, 676, 679
- ventana de proceso, 664, 666
- procesos de iniciación, 666
- procesos globales, 682
- producción automatizada, 643
- profundidad del árbol
 - en Análisis de conglomerados en dos fases, 491
- programación con lenguaje de comandos, 321
- propiedades, 300, 301, 678
- objetos de automatización OLE, 678
- tabla, 301
- tablas pivote, 300
- protección de contraseña, 274
- Proximidades
 - en Análisis de conglomerados jerárquico, 497
- prueba *b* de Tukey
 - en MLG, 413
- Prueba binomial, 520
 - dicotomías, 520
 - estadísticos, 522
 - funciones adicionales del comando, 522
 - opciones, 522
 - valores perdidos, 522
- Prueba de aditividad de Tukey
 - en Análisis de fiabilidad, 571, 573
- Prueba de chi-cuadrado, 516
 - estadísticos, 519
 - funciones adicionales del comando, 520
 - opciones, 519
 - rango esperado, 518
 - valores esperados, 518
 - valores perdidos, 519
- prueba de esfericidad de Bartlett
 - en Análisis factorial, 478
- prueba de Friedman
 - en pruebas para varias muestras relacionadas, 543
- prueba de Gabriel
 - en ANOVA de un factor, 397
 - en MLG, 413
- prueba de Games-Howell
 - en ANOVA de un factor, 397
 - en MLG, 413
- Prueba de Kolmogorov-Smirnov para una muestra, 526
 - distribución de contraste, 526
 - estadísticos, 528
 - funciones adicionales del comando, 529
 - opciones, 528
 - valores perdidos, 528
- prueba de la mediana
 - en Pruebas para dos muestras independientes, 537
- prueba de Levene
 - en ANOVA de un factor, 400
 - en Explorar, 348
 - en MLG Univariante, 418
- prueba de Lilliefors
 - en Explorar, 348
- prueba de los signos
 - en Pruebas para dos muestras relacionadas, 533
- prueba de McNemar
 - en Pruebas para dos muestras relacionadas, 533
 - en Tablas de contingencia, 354
- Prueba de rachas, 523
 - estadísticos, 525
 - funciones adicionales del comando, 526
 - opciones, 525
 - punto de corte, 525
 - puntos de corte, 523
 - valores perdidos, 525
- prueba de rangos múltiples de Duncan
 - en ANOVA de un factor, 397
 - en MLG, 413
- prueba de reacciones extremas de Moses
 - en Pruebas para dos muestras independientes, 531
- prueba de Scheffé
 - en MLG, 413

- Prueba de Scheffé
 - en ANOVA de un factor, 397
- prueba de Shapiro-Wilk
 - en Explorar, 348
- prueba de Wilcoxon de los rangos con signo
 - en Pruebas para dos muestras relacionadas, 533
- prueba exacta de Fisher
 - en Tablas de contingencia, 354
- Prueba M de Box
 - en Análisis discriminante, 464
- pruebas de chi-cuadrado
 - asociación lineal por lineal, 354
 - corrección por continuidad de Yates, 354
 - en Tablas de contingencia, 354
 - para la independencia, 354
 - Pearson, 354
 - prueba exacta de Fisher, 354
 - prueba para una muestra, 516
 - razón de verosimilitud, 354
- pruebas de homogeneidad de las varianzas
 - en ANOVA de un factor, 400
 - en MLG Univariante, 418
- pruebas de normalidad
 - en Explorar, 348
- pruebas no paramétricas
 - Prueba de chi-cuadrado, 516
 - Prueba de Kolmogorov-Smirnov para una muestra, 526
 - Prueba de rachas, 523
 - Pruebas para dos muestras independientes, 529
 - Pruebas para dos muestras relacionadas, 533
 - Pruebas para varias muestras independientes, 537
 - Pruebas para varias muestras relacionadas, 541
- Pruebas para dos muestras independientes, 529
 - definición de grupos, 532
 - estadísticos, 533
 - funciones adicionales del comando, 533
 - opciones, 533
 - tipos de pruebas, 531
 - valores perdidos, 533
 - variables de agrupación, 532
- Pruebas para dos muestras relacionadas, 533
 - estadísticos, 537
 - funciones adicionales del comando, 537
 - opciones, 537
 - tipos de pruebas, 536
 - valores perdidos, 537
- pruebas para la independencia
 - chi-cuadrado, 354
- Pruebas para varias muestras independientes, 537
 - definición de rango, 540
 - estadísticos, 540
 - funciones adicionales del comando, 541
 - opciones, 540
 - tipos de pruebas, 539
 - valores perdidos, 540
 - variables de agrupación, 540
- Pruebas para varias muestras relacionadas, 541
 - estadísticos, 543
 - funciones adicionales del comando, 544
 - tipos de pruebas, 543
- pruebas t
 - en MLG Univariante, 418
 - en Pruebas t para muestras independientes, 381
 - en Prueba T para muestras relacionadas, 385
 - en Prueba t para una muestra, 389
- prueba t de Dunnett
 - en ANOVA de un factor, 397
 - en MLG, 413
- prueba t dependiente
 - en Prueba T para muestras relacionadas, 385
- prueba t de Sidak
 - en ANOVA de un factor, 397
 - en MLG, 413
- prueba t de Student, 381
- prueba t de Waller-Duncan
 - en ANOVA de un factor, 397
 - en MLG, 413
- prueba t para dos muestras
 - en Pruebas t para muestras independientes, 381
- Prueba T para muestras independientes, 381
 - definición de grupos, 384
 - intervalos de confianza, 385
 - opciones, 385
 - valores perdidos, 385
 - variables de agrupación, 384
 - variables de cadena, 384
- Prueba T para muestras relacionadas, 385
 - opciones, 388
 - selección de variables relacionadas, 385
 - valores perdidos, 388

- Prueba T para una muestra, 389
 intervalos de confianza, 391
 opciones, 391
 valores perdidos, 391
- prueba Tukey-b
 en ANOVA de un factor, 397
- publicación de resultados , 658
 mediante la Unidad de producción, 657
- puntos de interrupción, 688
 en procesos, 688
- puntuación, 189
- puntuaciones de Savage, 157
- puntuaciones factoriales, 483
- puntuaciones factoriales de Anderson-Rubin, 483
- puntuaciones factoriales de Bartlett, 483
- puntuaciones normales
 en la asignación de rangos a los casos, 157
- puntuaciones z
 almacenamiento como variables, 339
 en Descriptivos, 339
 en la asignación de rangos a los casos, 157
- Q de Cochran
 en pruebas para varias muestras relacionadas,
 543
- Rachas de Wald-Wolfowitz
 en Pruebas para dos muestras independientes,
 531
- rango múltiple de Ryan-Einot-Gabriel-Welsch
 en ANOVA de un factor, 397
 en MLG, 413
- razón entre covarianzas
 en Regresión lineal, 443
- R-cuadrado
 cambio en R cuadrado, 446
 en Medias, 370
 en Regresión lineal, 446
- R-cuadrado corregida
 en Regresión lineal, 446
- recodificación de variables, 128, 149, 150, 152,
 153, 160
- recuento de apariciones, 147
- recuento esperado
 en Tablas de contingencia, 358
- recuento observado
 en Tablas de contingencia, 358
- reducción de gráficos exportados, 260
- reestructuración de los datos , 213, 213, 214, 217,
 218, 218, 220, 223, 225, 225, 226, 228, 229,
 231, 233, 235, 237
 conceptos básicos, 213
 creación de una única variable de índice para
 variables a casos, 226
 creación de variables de índice para variables a
 casos, 223
 creación de varias variables de índice para
 variables a casos, 228
 ejemplo de casos a variables, 218
 ejemplo de dos índices para variables a casos,
 225
 ejemplo de un índice para variables a casos, 225
 ejemplo de variables a casos, 217
 grupos de variables para variables a casos, 218
 opciones para casos a variables, 235
 opciones para variables a casos, 229
 ordenación de los datos para reestructurar casos
 a variables, 233
 selección de datos para reestructurar casos a
 variables, 231
 selección de datos para reestructurar variables a
 casos, 220
 tipos de reestructuración, 214
 y datos ponderados, 237
- regresión
 gráficos, 441
 Regresión lineal, 435
 regresión múltiple, 435
- Regresión lineal, 435
 almacenamiento de nuevas variables, 443
 bloques, 435
 estadísticos, 446
 exportación de información del modelo, 443
 gráficos, 441
 métodos de selección de variables, 439, 448
 ponderaciones, 435

- residuos, 443
- valores perdidos, 448
- variable de selección, 441
- regresión múltiple
 - en Regresión lineal, 435
- R-E-G-W F
 - en ANOVA de un factor, 397
 - en MLG, 413
- R-E-G-W Q
 - en ANOVA de un factor, 397
 - en MLG, 413
- relación de aspecto, 621
- rendimiento, 68
 - caché de datos, 68
- residuos
 - almacenamiento en Estimación curvilínea, 456
 - almacenamiento en Regresión lineal, 443
 - en Tablas de contingencia, 358
- residuos eliminados
 - en MLG, 416
 - en Regresión lineal, 443
- residuos estudentizados
 - en Regresión lineal, 443
- residuos no tipificados
 - en MLG, 416
- residuos tipificados
 - en MLG, 416
 - en Regresión lineal, 443
- Respuesta múltiple
 - funciones adicionales del comando, 554
- resultados , 239, 241, 242, 242, 243, 243, 244, 248, 249, 252, 253, 273, 275, 316, 616, 646
 - alineación, 244, 616
 - almacenamiento, 273
 - borrador, 275
 - cambio del idioma de los resultados, 614
 - centrado, 244, 616
 - copia, 242, 243
 - copia en otras aplicaciones, 249
 - copia y pegado de múltiples elementos, 252
 - desplazamiento, 242, 242
 - eliminación, 242, 243
 - exportación, 253, 646
 - modificación, 316
 - ocultación, 241
 - pegado en otras aplicaciones, 248
 - presentación, 241
 - Visor, 239
- Resumir, 361
 - estadísticos, 364
 - opciones, 363
- rho
 - en Correlaciones bivariadas, 421
 - en Tablas de contingencia, 354
- riesgo
 - en Tablas de contingencia, 354
- R múltiple
 - en Regresión lineal, 446
- rotación de etiquetas, 289
- rotación equamax
 - en Análisis factorial, 481
- rotación oblimin directa
 - en Análisis factorial, 481
- rotación quartimax
 - en Análisis factorial, 481
- rotación varimax
 - en Análisis factorial, 481
- rupturas de tabla, 318
- saltos de línea
 - etiquetas de variable y de valor, 93
- selección de casos, 206
 - muestra aleatoria, 209
 - rango de casos, 210
 - rango de fechas, 210
 - rango de horas, 210
 - según criterios de selección, 208
- selección hacia delante
 - en Regresión lineal, 439
- selección por pasos
 - en Regresión lineal, 439
- semilla de aleatorización, 146
- señalización de vistas de tablas pivote, 294
- señalizadores, 294
- separación de tablas, 318
 - control de las rupturas de tabla, 318
- servidores , 73, 74, 74
 - acceso, 73
 - adición, 74

- edición, 74
- nombres, 74
- números de puerto, 74
- servidores remotos, 71, 71, 73, 74, 76, 77, 78, 80, 81, 651
- acceso, 73
- acceso a un archivo de datos, 76, 78
- adición, 74
- almacenamiento de archivos de datos, 77
- edición, 74
- procedimientos disponibles, 80
- rutas de acceso UNC, 81
- Unidad de producción, 651
- SGR, 695, 725
- control de pivotes de tabla, 704, 717
- exclusión de resultados del Visor, 710
- Formato de archivo de datos de SPSS, 704, 711
- formato de archivo SAV, 704, 711
- formato de texto, 704
- identificadores de comandos, 701
- nombres de variable en los archivos SAV, 719
- subtipos de tabla, 701
- tipos de objetos de resultados, 699
- Uso de XSLT con OXML, 726
- XML, 704, 721
- signo de porcentaje, 311
- en las tablas pivote, 311
- signo dólar, 311
- en las tablas pivote, 311
- sintaxis, 321, 328, 614, 616, 617, 639, 643, 691, 693
 - anotación, 616, 617
 - anotaciones de los resultados, 324
 - archivo de diario, 326, 328, 614
 - ejecución, 328
 - ejecución de sintaxis de comandos mediante botones de las barras de herramientas, 639
 - instalación de referencia de sintaxis de comandos, 15
 - mediante procesos, 691, 691, 693
 - pegado, 323
 - pegado en los procesos, 692
 - reglas de la Unidad de producción, 643
 - reglas de sintaxis, 322
 - sintaxis de comandos, 321, 328, 614, 616, 617, 635, 640, 643
 - adición a menús, 635
 - anotación, 616, 617
 - anotaciones de los resultados, 324
 - archivo de diario, 326, 328, 614
 - ejecución, 328
 - ejecución mediante los botones de la barra de herramientas, 640
 - formato de la Unidad de producción, 654
 - instalación de referencia de sintaxis de comandos, 15
 - pegado, 323
 - reglas de la Unidad de producción, 643
 - reglas de sintaxis, 322
- Sistema de gestión de resultados, 695, 725
- sistemas de medida, 614
- sombreado, 314
- en las casillas, 314
- SPSS
 - pasos básicos, 13
- S-stress
 - en Escalamiento multidimensional, 577
- stress
 - en Escalamiento multidimensional, 577
- Student-Newman-Keuls
 - en ANOVA de un factor, 397
 - en MLG, 413
- suavizado T4253H, 185
- subconjuntos de casos
 - muestra aleatoria, 209
 - selección, 206, 208, 210
- subtipos, 701
- frente a etiquetas, 702
- subtipos de tabla, 701
- frente a etiquetas, 702
- subtítulos
 - en los gráficos, 595
- subtotales
 - en el informe de estadísticos en columnas, 568
- suma
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364

- suma de cuadrados, 408
 - en MLG, 407
- sustitución de valores perdidos
 - interpolación lineal, 188
 - media de la serie, 188
 - media de los puntos adyacentes, 188
 - mediana de los puntos adyacentes, 188
 - tendencia lineal, 188
- T2 de Tamhane
 - en ANOVA de un factor, 397
 - en MLG, 413
- T3 de Dunnett
 - en ANOVA de un factor, 397
 - en MLG, 413
- tabla de contingencia
 - en Tablas de contingencia, 351
 - respuesta múltiple, 550
- tablas de contingencia, 351
- Tablas de contingencia , 351
 - capas, 354
 - estadísticos, 354
 - formatos, 360
 - gráficos de barras agrupadas, 354
 - ponderaciones fraccionarias, 211
 - presentación de casillas, 358
 - supresión de tablas, 351
 - variables de control, 354
- Tablas de contingencia de respuestas múltiples, 550
 - definición de rangos de valores, 552
 - emparejamiento de las variables entre los conjuntos de respuestas, 553
 - porcentajes basados en casos, 553
 - porcentajes basados en respuestas, 553
 - porcentajes de casilla, 553
 - valores perdidos, 553
- tablas de frecuencias
 - en Explorar, 347
 - en Frecuencias, 331
- tablas pivote , 241, 248, 248, 249, 249, 250, 251, 251, 251, 252, 253, 266, 285, 285, 285, 285, 286, 287, 287, 287, 287, 288, 288, 289, 290, 290, 295, 298, 298, 300, 301, 301, 302, 303, 305, 306, 306, 308, 316, 317, 318, 627, 646, 652
 - adición de texto al pie, 317
 - agrupación de filas o columnas, 288
 - ajustes en la anchura de columna por defecto, 627
 - anchuras de casillas, 308
 - aspecto por defecto para las tablas nuevas, 627
 - bordes, 305
 - búsqueda de definiciones de etiqueta, 290
 - cambio del aspecto, 298, 298
 - cambio del orden de presentación, 287, 287
 - capas, 290
 - control de las rupturas de tabla, 318
 - control del formato en los trabajos de producción, 652
 - copia, 249
 - copia en otras aplicaciones, 249
 - copia y pegado de múltiples tablas, 252
 - desagrupación de filas o columnas, 288
 - desplazamiento de filas y columnas, 287
 - edición, 285, 285
 - edición de dos o más, 285
 - eliminación de etiquetas de grupo, 288
 - escalamiento para ajustarse a la página, 301, 306
 - exportación como HTML, 253, 646
 - formatos de casilla, 303
 - gestión, 285
 - identificación de dimensiones, 287
 - impresión de capas, 266
 - impresión de tablas grandes, 318
 - incrustación como objetos ActiveX, 250
 - inserción de etiquetas de grupo, 288
 - líneas de cuadrícula, 306
 - ocultación, 241
 - pegado como metarchivos, 251
 - pegado como tablas, 248, 251
 - pegado como texto, 251
 - pegado en otras aplicaciones, 248
 - pivotado, 285, 286
 - presentación de bordes ocultos, 306
 - presentación y ocultación de casillas, 295
 - propiedades, 300
 - propiedades de las notas al pie, 302
 - propiedades generales, 301
 - restablecimiento de los valores por defecto, 290

- rotación de etiquetas, 289
- selección de filas y columnas, 316
- texto de continuación, 306
- transposición de filas y columnas, 287
- utilizando iconos, 286
- tamaños , 247
 - en titulares, 247
- tasa de riesgo relativo
 - en Tablas de contingencia, 354
- tau-b
 - en Tablas de contingencia, 354
- Tau-b de Kendall
 - en Correlaciones bivariadas, 421
 - en Tablas de contingencia, 354
- tau-c
 - en Tablas de contingencia, 354
- Tau-c de Kendall, 354
 - en Tablas de contingencia, 354
- tau de Goodman y Kruskal
 - en Tablas de contingencia, 354
- tau de Kruskal
 - en Tablas de contingencia, 354
- T-cuadrado de Hotelling
 - en Análisis de fiabilidad, 571, 573
- términos de interacción, 408
- texto , 44, 247, 248, 253, 258, 275, 284, 316, 646
 - adición al Visor, 247
 - adición de un archivo de texto al Visor, 248
 - archivos de datos, 44
 - creación de resultados de texto, 275
 - en las casillas, 316
 - exportación de resultados como texto, 253, 258, 646
 - exportación de resultados de borrador como texto, 284
- texto de continuación, 306
 - para tablas pivote, 306
- texto de etiqueta vertical, 289
- textos al pie, 317
 - adición a una tabla, 317
- tipificación
 - en Análisis de conglomerados en dos fases, 491
- tipos de datos , 89, 89, 91, 103, 631
 - cambio, 103
 - definición, 89
 - formatos de entrada, 91
 - formatos de presentación, 91
 - moneda personalizada, 89, 631
- tipos de objetos de resultados
 - en SGR, 699
- titulares , 244, 246, 246
 - cambio de los niveles, 246
 - contracción, 246
 - en el Visor, 244
 - expansión, 246
- títulos , 247
 - adición al Visor, 247
 - en Cubos OLAP, 380
 - en los gráficos, 595
- tolerancia
 - en Regresión lineal, 446
- totales, 660
 - aplicación automática de la negrita en los resultados, 660
- totales finales
 - en el informe de estadísticos en columnas, 569
- transformaciones condicionales, 143
- transformaciones de archivos , 213, 214
 - agregación de datos, 200
 - análisis de segmentación de archivo, 205
 - fusión de archivos de datos, 194, 197
 - ordenación de casos, 192
 - ponderación de casos, 211
 - reestructuración de los datos, 213, 214
 - transposición de variables y casos, 193
- transformaciones de los datos, 630
 - asignación de rangos a los casos, 156
 - cálculo de variables, 141
 - funciones, 144
 - recodificación de los valores, 149, 150, 152, 153, 160
 - retraso de la ejecución, 630
 - transformaciones condicionales, 143
 - variables de cadena, 144
 - variables de series temporales, 181, 183
- transposición de filas y columnas, 287
- transposición de variables y casos, 193
- tratamiento del ruido
 - en Análisis de conglomerados en dos fases, 491

- U de Mann-Whitney
 - en Pruebas para dos muestras independientes, 531
- última
 - en Cubos OLAP, 376
 - en Medias, 370
 - en Resumir, 364
- umbral inicial
 - en Análisis de conglomerados en dos fases, 491
- Unidad de producción , 614, 616, 617, 643, 643, 643, 646, 648, 650, 651, 651, 652, 655, 657
 - archivos de resultados, 643
 - control del formato en las tablas pivote, 652
 - control del formato mediante sintaxis de comandos, 654
 - ejecución de varios trabajos de producción, 655
 - entrada de macro, 650
 - especificación de un servidor remoto, 651
 - exportación de gráficos, 643, 646
 - exportación de resultados, 646
 - modificadores de la línea de comando, 655
 - opciones, 651
 - programación de trabajos de producción, 655
 - publicación de resultados, 657
 - publicación en Web, 657
 - reglas de sintaxis, 643
 - sustitución de valores en archivos de sintaxis, 648
 - uso de la sintaxis de comandos de archivos de diario, 614
 - uso de la sintaxis de comandos de la anotación, 616, 617
- unidades compartidas, 80
- V
 - en Tablas de contingencia, 354
- valores, 311
 - formato de presentación de las tablas pivote, 311
- valores atípicos
 - en Análisis de conglomerados en dos fases, 491
 - en Explorar, 347
 - en Regresión lineal, 441
- valores de datos de categorización, 128
- valores definidos por el usuario como perdidos, 93
- valores de influencia
 - en MLG, 416
 - en Regresión lineal, 443
- valores perdidos , 93, 93
 - definición, 93
 - en Análisis factorial, 484
 - en ANOVA de un factor, 400
 - en Correlaciones bivariadas, 423
 - en Correlaciones parciales, 428
 - en Curva COR..., 606
 - en el informe de estadísticos en columnas, 569
 - en el Informe de estadísticos en filas, 561
 - en Explorar, 350
 - en la prueba de chi-cuadrado, 519
 - en las frecuencias de respuestas múltiples, 548
 - en las funciones, 145
 - en las tablas de contingencia de respuestas múltiples, 553
 - en los gráficos, 597, 599
 - en Prueba binomial, 522
 - en Prueba de Kolmogorov-Smirnov para una muestra, 528
 - en Prueba de rachas, 525
 - en Pruebas para dos muestras independientes, 533
 - en Pruebas para dos muestras relacionadas, 537
 - en Pruebas para varias muestras independientes, 540
 - en Pruebas t para muestras independientes, 385
 - en Prueba T para muestras relacionadas, 388
 - en Prueba t para una muestra, 391
 - en Regresión lineal, 448
 - reemplazo en datos de serie temporal, 186
 - variables de cadena, 93
- valores pronosticados
 - almacenamiento en Estimación curvilínea, 456
 - almacenamiento en Regresión lineal, 443
- valores pronosticados ponderados
 - en MLG, 416

- valores tipificados
 - en Descriptivos, 339
- variable de entorno SPSSTMPDIR, 614
- variable de selección
 - en Regresión lineal, 441
- variables , 11, 86, 102, 103, 214, 607, 609, 614
 - cambio de nombre de los archivos de datos fusionados, 196
 - definición, 86
 - definición de conjuntos de variables, 609
 - de los cuadros de diálogo, 8
 - desplazamiento, 103
 - información sobre la definición, 607
 - información sobre variables en los cuadros de diálogo, 11
 - inserción de nuevas variables, 102
 - orden de presentación de los cuadros de diálogo, 614
 - recodificación, 149, 150, 152, 153, 160
 - reestructuración en casos, 214
 - selección en los cuadros de diálogo, 11
- variables de agrupación , 214
 - creación, 214
- variables de cadena , 93
 - cálculo de variables de cadena nuevas, 144
 - de los cuadros de diálogo, 8
 - fragmentación de cadenas largas en versiones anteriores, 62
 - recodificación a valores enteros consecutivos, 160
 - valores perdidos, 93
- variables de control
 - en Tablas de contingencia, 354
- variables de entorno , 614
 - SPSSTMPDIR, 614
- variables de escala
 - categorización para la generación de variables categóricas, 128
- variables de fecha
 - adición o sustracción de variables de fecha/hora, 163
 - creación de variable de fecha/hora a partir de conjunto de variables, 163
 - creación de variable de fecha/hora a partir de una cadena, 163
 - definición para datos de serie temporal, 181
 - extracción de parte de variable de fecha/hora, 163
- variables de segmentación
 - en Agregar datos, 200
- varianza
 - en Cubos OLAP, 376
 - en Descriptivos, 341
 - en el Informe de estadísticos en columnas, 566
 - en el Informe de estadísticos en filas, 559
 - en Explorar, 347
 - en Frecuencias, 334
 - en Medias, 370
 - en Resumir, 364
- varias vistas/paneles
 - Editor de datos, 104
- V de Cramér
 - en Tablas de contingencia, 354
- V de Rao
 - en Análisis discriminante, 465
- velocidad, 68
 - caché de datos, 68
- ventana activa, 6
- ventana de proceso, 664, 666, 679
 - menú Depurar, 688
 - pestaña Immediate, 689
 - pestaña Loaded, 689
 - pestaña Stack, 689
 - pestaña Watch, 689
 - propiedades, 666
 - visor de objetos, 679
- ventana designada, 6
- ventanas, 4
 - ventana activa, 6
 - ventana designada, 6
- ventanas de ayuda, 15
- Visor , 239, 241, 242, 243, 243, 244, 246, 246, 247, 247, 252, 272, 273, 616, 619
 - almacenamiento de documentos, 273
 - cambio de las fuentes de los titulares, 247
 - cambio de los niveles de titulares, 246
 - cambio de los tamaños de los titulares, 247
 - contracción de titulares, 246
 - copia de resultados, 243
 - desplazamiento de los resultados, 242
 - eliminación de resultados, 243

- espacio entre los elementos de resultados, 272
- exclusión de tipos de resultados con SGR, 710
- expansión de titulares, 246
- muestra de los valores de datos, 619
- ocultación de resultados, 241
- opciones de presentación, 616
- paneles de resultados, 239
- paneles de titulares, 239
- pegado de objetos especiales, 252
- presentación de etiquetas de valor, 619
- presentación de etiquetas de variable, 619
- presentación de nombres de variable, 619
- titulares, 244
- Visor de borrador, 275, 276, 277, 277, 282, 282, 283, 284, 614, 617
 - almacenamiento de resultados, 284
 - bordes de casillas, 277
 - bordes de columnas, 277
 - bordes de filas, 277
 - cambio de fuentes, 282
 - caracteres de caja, 277
 - control de la presentación de los resultados por defecto, 614
 - establecimiento del tipo de Visor por defecto, 276
 - formato de resultados, 277
 - impresión, 282, 283
 - opciones de presentación, 617
- visor de objetos, 679
- Vista de datos, 84
- W de Kendall
 - en pruebas para varias muestras relacionadas, 543
- Web , 658
 - publicación de resultados en, 658
- XML
 - almacenamiento de resultados como XML, 695
 - envío de resultados a XML, 704
 - estructura de tablas en OXML, 721
 - Resultados OXML del SGR, 726
- XSLT
 - uso con OXML, 726
- Z de Kolmogorov-Smirnov
 - en Prueba de Kolmogorov-Smirnov para una muestra, 526
 - en Pruebas para dos muestras independientes, 531