

Guía breve de IBM SPSS Statistics 22

IBM

Nota

Antes de utilizar esta información y el producto al que da soporte, lea la información que se incluye en el apartado "Avisos" en la página 91.

Información sobre el producto

Esta edición se aplica a la versión 22, release 0, modificación 0 de IBM SPSS Statistics y a todas las versiones y modificaciones posteriores hasta que se indique lo contrario en nuevas ediciones.

Contenido

Capítulo 1. Introducción 1

Archivos muestrales	1
Apertura de un archivo de datos	1
Ejecución de un análisis	3
Creación de gráficos	4

Capítulo 2. Lectura de datos 7

Estructura básica de archivos de datos de IBM SPSS Statistics	7
Lectura de archivos de datos de IBM SPSS Statistics	7
Lectura de datos de hojas de cálculo	8
Lectura de datos de una base de datos	9
Lectura de datos desde un archivo de texto	12

Capítulo 3. Uso del Editor de datos 15

Introducción de datos numéricos	15
Introducción de datos de cadena	16
Definición de datos	17
Adición de etiquetas de variable	17
Cambio de tipo de variable y formato	18
Añadir etiquetas de valores	19
Tratamiento de los datos perdidos	19
Valores perdidos de una variable numérica	20
Valores perdidos de una variable de cadena	20

Capítulo 4. Examen de estadísticos de resumen para variables individuales 23

Nivel de medición	23
Medidas de resumen para datos categóricos	23
Gráficos para datos categóricos	24
Medidas de resumen para variables de escala	25
Histogramas para variables de escala	26

Capítulo 5. Creación y edición de gráficos 29

Procedimiento básico de creación de gráficos	29
Uso de la galería del generador de gráficos	30
Definición de variables y estadísticos	30
Adición de texto	32
Creación del gráfico	33

Capítulo 6. Cómo trabajar con resultados 35

Uso del Visor	35
Uso del Editor de tablas pivote	36
Acceso a las definiciones de resultados	36
Tablas de pivotado	37
Creación y visualización de capas	38
Edición de tablas	39
Ocultación de filas y columnas	40
Cambio de formato de presentación de datos	40
TableLook	42
Utilización de formatos predefinidos	42
Personalización de estilos de aspectos de tabla	42

Cambio de los formatos de tablas predeterminadas	45
Personalización de las opciones de presentación inicial	46
Mostrar las etiquetas de variable y de valor	47
Uso de los resultados en otras aplicaciones	48
Pegado de resultados como tablas de Word	49
Pegado de resultados como texto	49
Exportación de resultados a archivos de Microsoft Word, PowerPoint y Excel	50
Exportación de resultados a PDF	55
Exportación de resultados a HTML	57

Capítulo 7. Trabajo con sintaxis 59

Pegado de la sintaxis	59
Edición de la sintaxis	60
Apertura y ejecución de un archivo de sintaxis	61
Utilización de puntos de separación	61

Capítulo 8. Modificación de los valores de los datos 63

Creación de una variable categórica a partir de una variable de escala	63
Cálculo de nuevas variables	65
Uso de funciones en las expresiones	66
Uso de Expresiones condicionales	67
Trabajo con fechas y horas	68
Cálculo de la distancia temporal entre dos fechas	69
Adición de una duración a una fecha	70

Capítulo 9. Ordenación y selección de datos 73

Ordenación de datos	73
Procesamiento de archivos segmentados	74
Ordenación de casos para el procesamiento de archivos segmentados	75
Activación y desactivación del procesamiento de archivos segmentados	76
Selección de subconjuntos de casos	76
Para seleccionar subconjuntos de casos basados en una expresión condicional	77
Para seleccionar una muestra aleatoria de casos	78
Selección de un rango de tiempo o un rango de casos	79
Tratamiento de los casos no seleccionados	79
Estado de selección de casos	80

Capítulo 10. Archivos muestrales 81

Avisos 91

Marcas comerciales	93
------------------------------	----

Índice 95

Capítulo 1. Introducción

Esta guía le mostrará cómo utilizar muchas de las características disponibles. Está diseñado para proporcionar una guía práctica paso a paso. Todos los ejemplos que aparecen en los ejemplos vienen instalados con la aplicación para que pueda seguirlos a la hora de realizar el mismo análisis y obtener los mismos resultados que aparecen aquí.

Si quiere ver ejemplos detallados de diversas técnicas de análisis estadísticos, acceda a los Estudios de casos, disponibles desde el menú Ayuda.

Archivos muestrales

La mayoría de los ejemplos que se ofrecen utilizan el archivo de datos *demo.sav*. Este archivo de datos es un estudio ficticio de varios miles de personas que contiene información básica demográfica y de consumo.

Si utiliza la versión para estudiantes, su versión de *demo.sav* será una muestra representativa del archivo de datos original, reducido para ajustarse al límite de 1.500 casos. Los resultados que obtenga con ese archivo de datos serán diferentes de los resultados que aparecen aquí.

Los archivos muestrales instalados con el producto se encuentran en el subdirectorio *Samples* del directorio de instalación. Existe otra carpeta con los subdirectorios de muestra de cada uno de los idiomas siguientes: inglés, francés, alemán, italiano, japonés, coreano, polaco, ruso, chino simplificado, español y chino tradicional.

No todos los archivos muestrales están disponibles en todos los idiomas. Si un archivo muestral no está disponible en un idioma, esa carpeta de idioma contendrá una versión en inglés del archivo muestral.

Apertura de un archivo de datos

Para abrir un archivo de datos:

1. Elija en los menús:

Archivo > Abrir > Datos...

Aparece un cuadro de diálogo para abrir archivos.

De forma predeterminada, aparecerán los archivos de datos de IBM® SPSS Statistics (extensión *.sav*).

En este ejemplo se utiliza el archivo *demo.sav*.

	age	marital	address	income	inccat	car
1	55	Marital status	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figura 1. Archivo demo.sav en el Editor de datos

El archivo de datos aparece en el Editor de datos. En el Visor de datos, si sitúa el cursor sobre un nombre de variable (las cabeceras de columnas), aparecerá una etiqueta de variable más descriptiva (si se ha definido para esa variable).

De forma predeterminada, aparecen los valores de los datos reales. Para visualizar etiquetas:

2. Elija en los menús:

Ver > Etiquetas de valores



Figura 2. Botón Etiquetas de valor

Si lo desea, puede utilizar el botón Etiquetas de valor de la barra de herramientas.

	age	marital	address	income	inccat	car
1	55	Married	12	72.00	\$50 - \$74	36.
2	56	Unmarried	29	153.00	\$75+	76.
3	28	Married	9	28.00	\$25 - \$49	13.
4	24	Married	4	26.00	\$25 - \$49	12.
5	25	Unmarried	2	23.00	Under \$25	11.
6	45	Married	9	76.00	\$75+	37.
7	42	Unmarried	19	40.00	\$25 - \$49	19.
8	35	Unmarried	15	57.00	\$50 - \$74	28.
9	46	Unmarried	26	24.00	Under \$25	12.
10	34	Married	0	89.00	\$75+	46.
11	55	Married	17	72.00	\$50 - \$74	35.

Figura 3. Etiquetas de valor en el Editor de datos

Ahora aparecerán las etiquetas de valor descriptivas para facilitar la interpretación de las respuestas.

Ejecución de un análisis

Si tiene alguna opción adicional, el menú Analizar contiene una lista de categorías de informes generales y de análisis estadísticos.

Para comenzar, crearemos una tabla de frecuencias simple (tabla de recuentos). Este ejemplo requiere la opción Statistics Base.

1. Elija en los menús:

Analizar > Estadísticos descriptivos > Frecuencias...

Aparecerá el cuadro de diálogo Frecuencias.

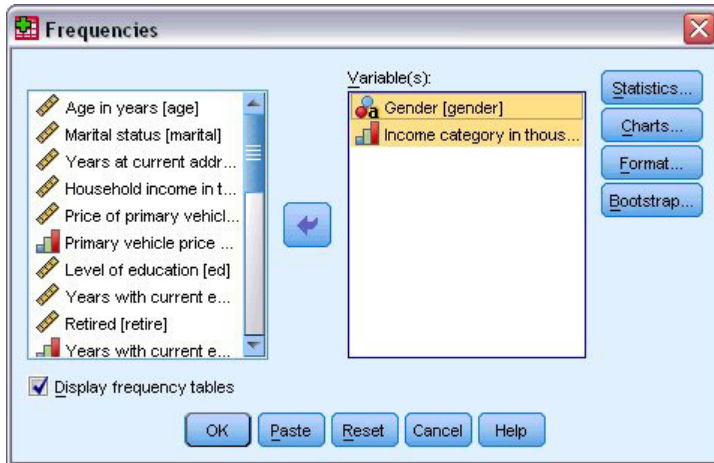


Figura 4. Cuadro de diálogo Frecuencias

Un icono junto a cada variable proporciona información acerca del tipo de datos y el nivel de medición.

	Numérico	Cadena	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Si el nombre o la etiqueta de la variable aparecen truncadas en la lista, puede ver el texto completo del nombre o la etiqueta situando el cursor sobre ella. El nombre de la variable *cating* aparecerá entre corchetes después de la etiqueta de variable descriptiva. *Categoría de ingresos (en miles)* es la etiqueta de la variable. Si no hubiese etiqueta de variable, sólo aparecería el nombre de la variable en el cuadro de lista.

Puede ajustar el tamaño de los cuadros de diálogo como si fueran ventanas, pulsando y arrastrando los bordes o esquinas exteriores. Por ejemplo, si aumenta el ancho del cuadro de diálogo, las listas de variables también serán más anchas.

En el cuadro de diálogo, seleccione las variables que desea analizar de la lista de origen de la izquierda y arrástrelas y suéltelas en la lista Variable(s) de la derecha. El botón **Aceptar**, que ejecuta el análisis, está desactivado hasta que se sitúe al menos una variable en la lista Variable(s).

En muchos cuadros de diálogo puede obtener más información si pulsa con el botón derecho del ratón en cualquier nombre de variable de la lista y selecciona **Información de variable** en el menú emergente.

2. Pulse en *Sexo [sexo]* en la lista de variables de origen y arrastre la variable hasta la lista de destino Variable(s).
3. Pulse en *Categoría de ingresos en miles [cating]* en la lista de origen y arrástrela a la lista de destino.

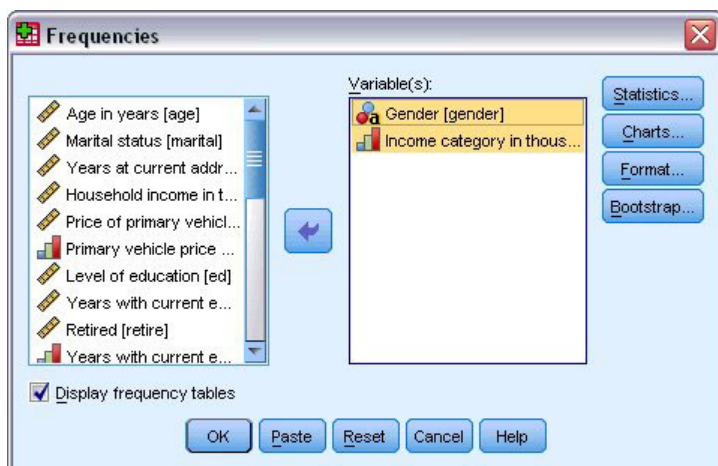


Figura 5. Variables seleccionadas para el análisis

4. Pulse en **Aceptar** para iniciar el procedimiento.

Los resultados aparecen en la ventana Visor.

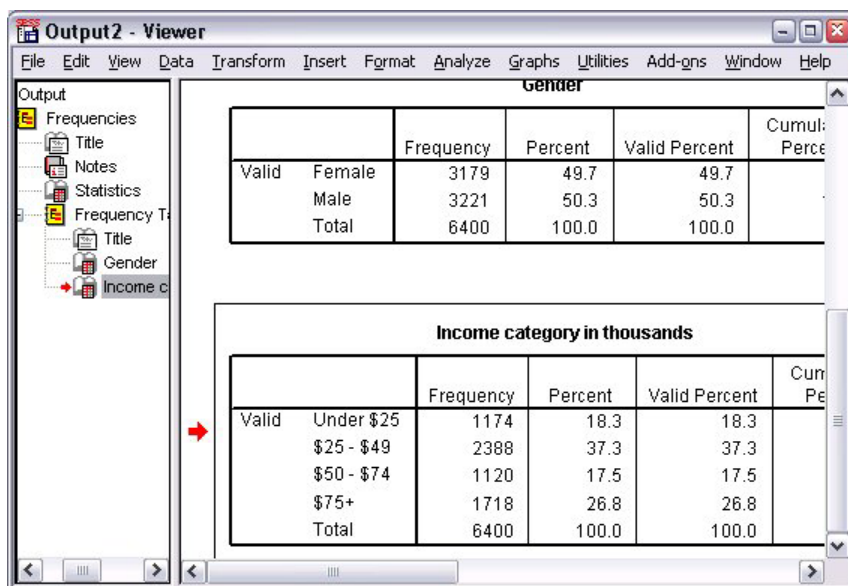


Figura 6. Tabla de frecuencias de las categorías de ingresos

Creación de gráficos

Aunque algunos procedimientos estadísticos pueden crear gráficos, también puede utilizar el menú Gráficos para crear gráficos.

Por ejemplo, puede crear un gráfico que muestre la relación entre el servicio de telefonía inalámbrica y el hecho de poseer una agenda digital (PDA).

1. Elija en los menús:

Gráficos > Generador de gráficos...

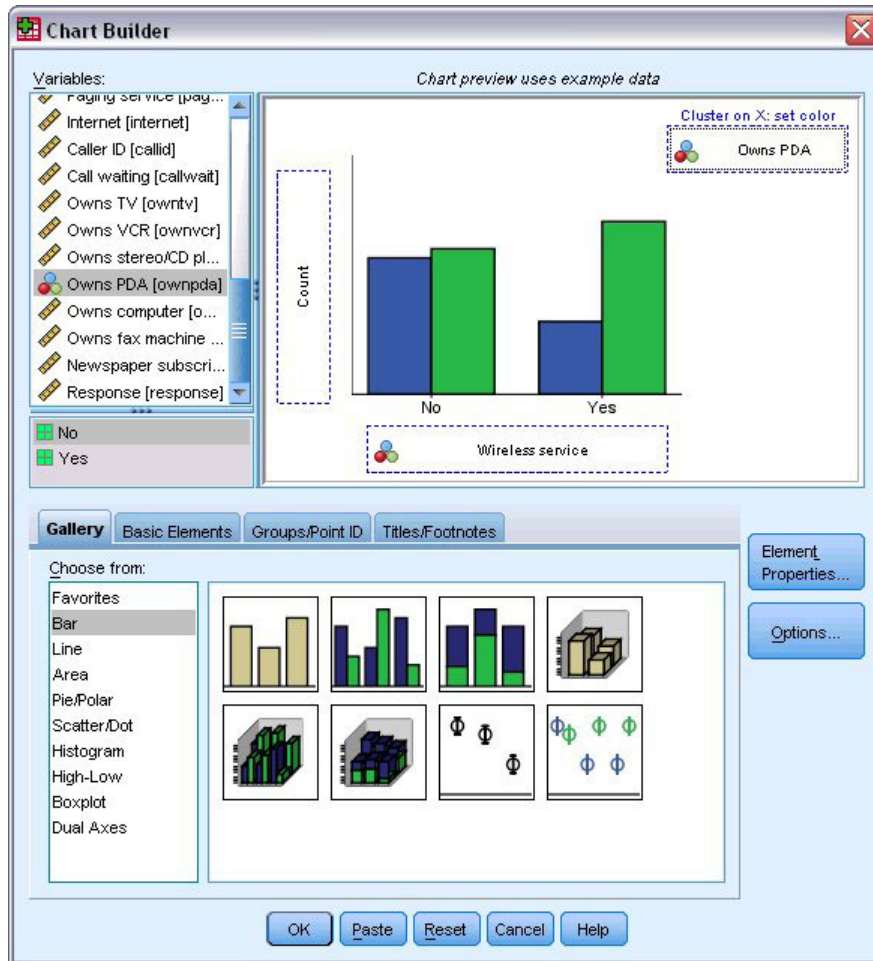


Figura 7. Cuadro de diálogo Generador de gráficos con las zonas de colocación rellenas

2. Pulse en la pestaña **Galería** (si no está seleccionada).
3. Pulse en **Barra** (si no está seleccionada).
4. Arrastre el icono de barras agrupadas al lienzo, que es la zona grande que hay encima de la galería.
5. Desplace hacia abajo la lista de variables, pulse con el botón derecho del ratón en *Servicio inalámbrico [inalam]* y elija **Nominal** como nivel de medición.
6. Arrastre la variable *Servicio inalámbrico [inalam]* al eje *x*.
7. Pulse con el botón derecho del ratón en *Tiene PDA [pda]* y elija **Nominal** como nivel de medición.
8. Arrastre la variable *Tiene PDA [pda]* a la zona de colocación de agrupación situada en la esquina superior derecha del lienzo.
9. Pulse en **Aceptar** para crear el gráfico.

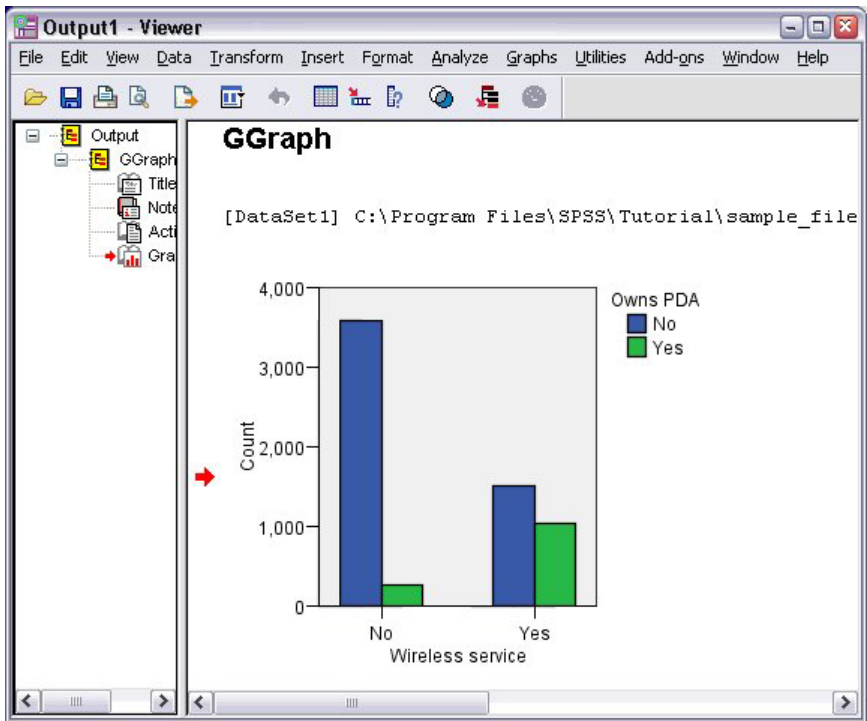


Figura 8. Gráfico de barras en la ventana Visor

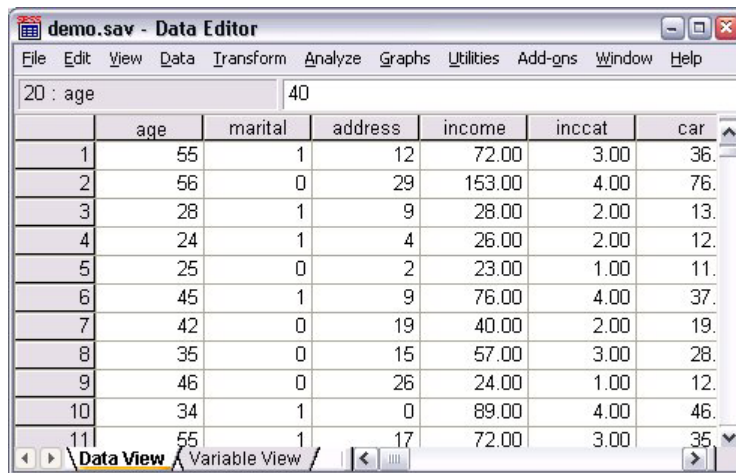
El gráfico de barras aparecerá en el Visor. El gráfico muestra que es mucho más probable que las personas que disponen de servicios de telefonía inalámbrica tengan también agendas digitales que los que no disponen de tales servicios.

Puede editar los gráficos y las tablas pulsando dos veces en ellos en el panel de contenidos de la ventana Visor y, a continuación, puede copiar y pegar los resultados en otras aplicaciones. Abordaremos esos temas más adelante.

Capítulo 2. Lectura de datos

Los datos se pueden introducir directamente, o importarse desde diferentes orígenes. En este capítulo se tratarán los procesos para leer datos almacenados en archivos de datos IBM SPSS Statistics, aplicaciones de hojas de cálculo como Microsoft Excel, aplicaciones de bases de datos como Microsoft Access y archivos de texto.

Estructura básica de archivos de datos de IBM SPSS Statistics



	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figura 9. Editor de datos

Los archivos de datos IBM SPSS Statistics están organizados por casos (filas) y variables (columnas). En este archivo de datos, los casos representan a los encuestados individuales que responden a una encuesta. Las variables representan respuestas a cada una de las preguntas realizadas en la encuesta.

Lectura de archivos de datos de IBM SPSS Statistics

Los archivos de datos IBM SPSS Statistics, que tienen una extensión de archivo *.sav*, contiene los datos guardados.

1. Elija en los menús:
Archivo > Abrir > Datos...
2. Acceda a *demo.sav* y ábralo. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Los datos aparecerán en el Editor de datos.

	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figura 10. Archivo de datos abierto

Lectura de datos de hojas de cálculo

En lugar de escribir todos los datos directamente en el Editor de datos, puede leer datos desde aplicaciones como Microsoft Excel. También puede leer los encabezados de las columnas como nombres de variables.

1. Elija en los menús:

Archivo > Abrir > Datos...

2. Seleccione **Excel (*.xls)** para el tipo de archivo que desea ver.
3. Abra demo.xls. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Aparecerá el cuadro de diálogo Apertura de origen de datos de Excel, que permite especificar si se incluyen los nombres de las variables en la hoja de cálculo, así como las casillas que se desea importar. En Excel 95 o posterior, también se pueden especificar las hojas de cálculo que se desea importar.

4. Asegúrese de que está seleccionada la opción **Leer nombre de variables de la primera fila de datos**. Esta opción lee los encabezados de columnas como nombres de variables.

Si los encabezados de las columnas no cumplen las normas de denominación de variables de IBM SPSS Statistics, se convertirán en nombres de variables válidos y los encabezados originales de las columnas se guardarán como etiquetas de variable. Si desea importar sólo una parte de la hoja de cálculo, especifique el rango de casillas que se va a importar en el campo de texto Rango.

5. Pulse en **Aceptar** para leer el archivo de Excel.

Los datos aparecerán en el Editor de datos, con los encabezados de las columnas utilizados como nombres de variables. Al no poder contener espacios los nombres de variables, se han quitado los espacios de los encabezados de las columnas originales. Por ejemplo, el *Estado civil* del archivo de Excel se ha convertido en la variable *Estadocivil*. El encabezado de la columna original se conservará como etiqueta de la variable.

	Age	Maritalstatus	Address	Income	IncomeCategory
1	55	1	12	72.00	3.00
2	56	0	29	153.00	4.00
3	28	1	9	28.00	2.00
4	24	1	4	26.00	2.00
5	25	1	2	23.00	1.00
6	45	0	9	76.00	4.00
7	44	1	17	144.00	4.00
8	46	1	20	75.00	4.00
9	41	0	10	26.00	2.00
10	29	0	4	19.00	1.00

Figura 11. Datos de Excel importados

Lectura de datos de una base de datos

Los datos procedentes de bases de datos se importan con gran facilidad mediante el Asistente para bases de datos. Se puede leer directamente cualquier base de datos que utilice controladores de ODBC (Open Database Connectivity), una vez instalados los controladores. En el CD de instalación, se suministran controladores ODBC para diferentes formatos de bases de datos. Se pueden obtener controladores adicionales de otros fabricantes. En este ejemplo se trata una de las aplicaciones de bases de datos más conocidas, Microsoft Access.

Nota: este ejemplo es específico para Microsoft Windows y requiere un controlador ODBC para Access. Los pasos son similares en otras plataformas, pero es posible que requieran un controlador ODBC de otro fabricante para Access.

1. Elija en los menús:
Archivo > Abrir base de datos > Nueva consulta...

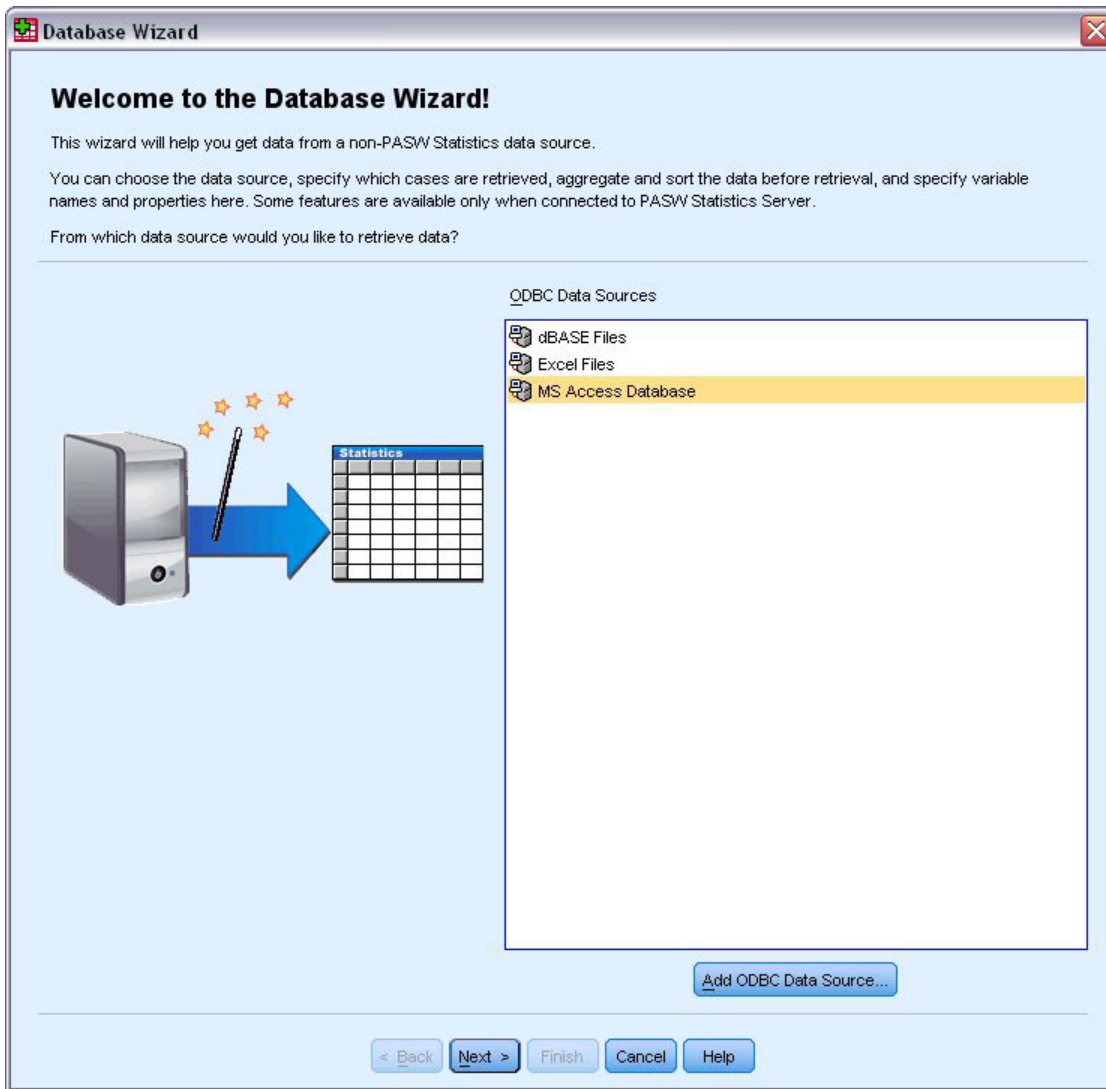


Figura 12. Cuadro de diálogo Bienvenido al asistente para bases de datos

2. Seleccione **MS Access Database** de la lista de orígenes de datos y pulse en **Siguiente**.

Nota: dependiendo de la instalación, es posible que también aparezca una lista de los orígenes de datos OLEDB en la parte izquierda del asistente (sólo sistemas operativos Windows), pero este ejemplo utiliza la lista de orígenes de datos ODBC que aparece a la derecha.

3. Pulse en **Examinar** para acceder al archivo de base de datos de Access que desea abrir.
4. Abrir *demo.mdb*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.
5. Pulse en **Aceptar** en el cuadro de diálogo de acceso.

En el siguiente paso se pueden especificar las tablas y variables que desea importar.

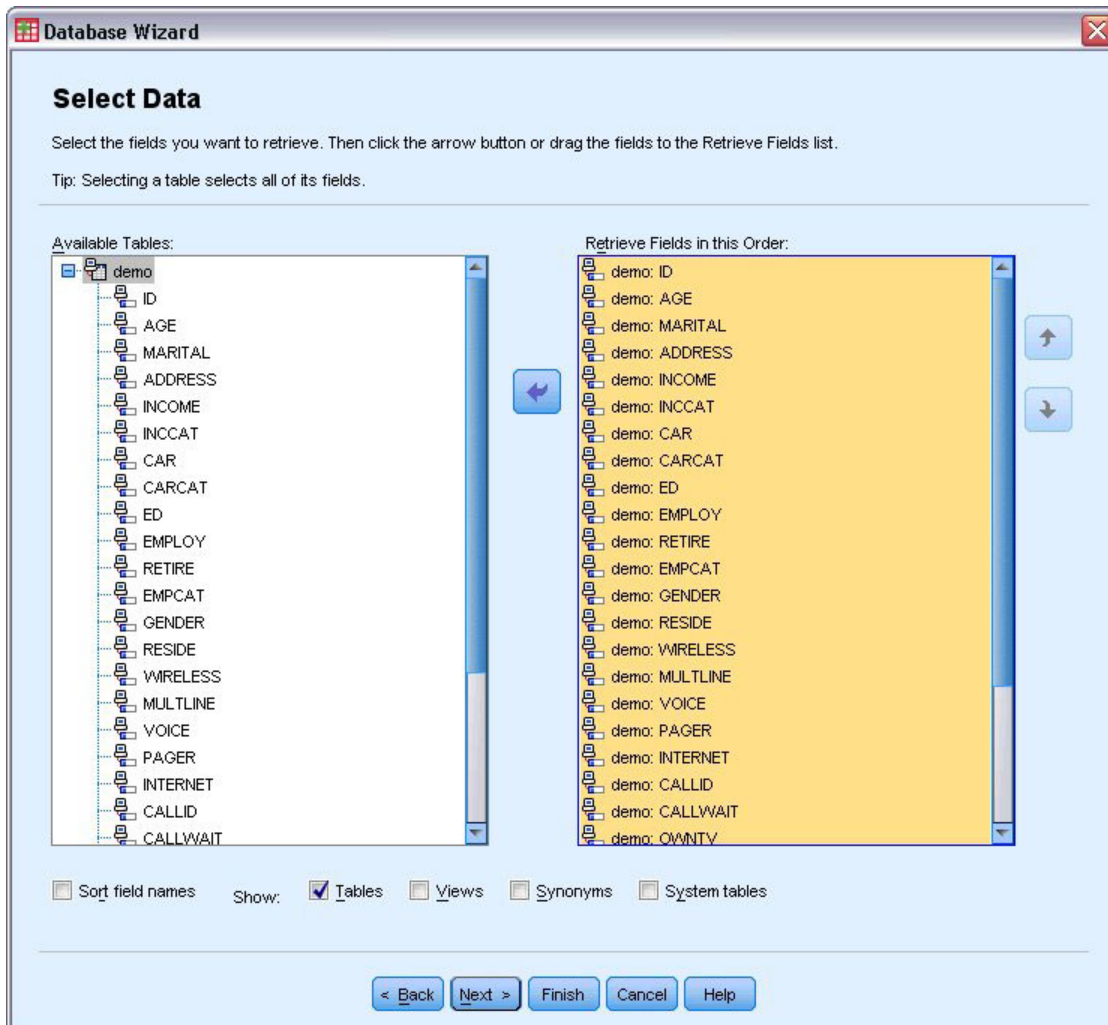


Figura 13. Paso Seleccionar datos

6. Arrastre toda la tabla **demo** hasta la lista Recuperar los campos en este orden.
7. Pulse en **Siguiente**.

En el siguiente paso, se pueden seleccionar los registros (casos) que desea importar.

Si no desea importar todos los casos, puede importar un subconjunto de casos (por ejemplo, los varones mayores de 30) o una muestra aleatoria de casos del origen de datos. Para grandes orígenes de datos, es posible que desee limitar el número de casos a una pequeña y representativa muestra para reducir el tiempo de procesamiento.

8. Pulse en **Siguiente** para continuar.

Los nombres de campo se utilizan para crear nombres de variable. Si es preciso, los nombres se convierten en nombres de variable válidos. Los nombres de campos originales se reservan como etiquetas de variable. También puede cambiar los nombres de variables antes de importar la base de datos.

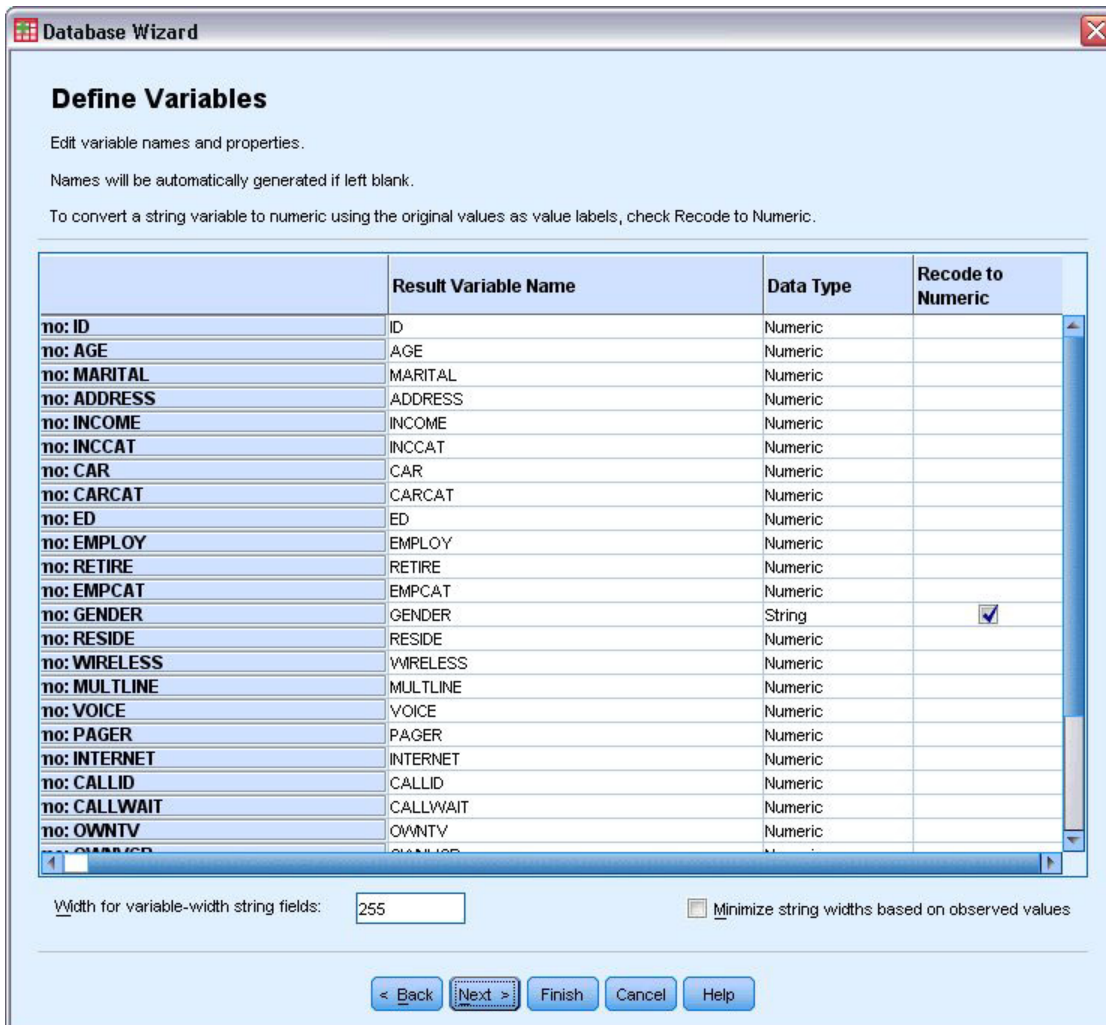


Figura 14. Paso Definir las variables

9. Pulse en la casilla **Recodificar como numérica** en el campo Género. Esta opción convierte las variables de cadena en variables enteras y conserva los valores originales como etiquetas de valor para la nueva variable.
10. Pulse en **Siguiente** para continuar.
En el cuadro de diálogo Resultados aparecerá la sentencia SQL creada a partir de las selecciones realizadas en el Asistente para bases de datos. Esta sentencia se puede ejecutar ahora o se puede guardar en un archivo para su uso posterior.
11. Pulse en **Finalizar** para importar los datos.

En el editor de datos, ya están disponibles todos los datos de la base de datos de Access seleccionados para su importación.

Lectura de datos desde un archivo de texto

Los archivos de texto son otro origen habitual de datos. Muchos programas de hojas de cálculo y de bases de datos pueden guardar su contenido en uno de los varios formatos de archivo de texto. Los archivos delimitados por comas o por tabuladores hacen referencia a filas de datos que utilizan comas o tabuladores para indicar cada variable. En este ejemplo, los datos están delimitados por tabulaciones.

1. Elija en los menús:

Archivo > Leer datos de texto...

2. Seleccione **Texto (*.txt)** para el tipo de archivo que desea ver.
3. Abra *demo.txt*. Consulte el tema Capítulo 10, "Archivos muestrales", en la página 81 para obtener más información.

El Asistente para la importación de texto le guiará a través del proceso para definir cómo se debe interpretar el archivo de texto especificado.

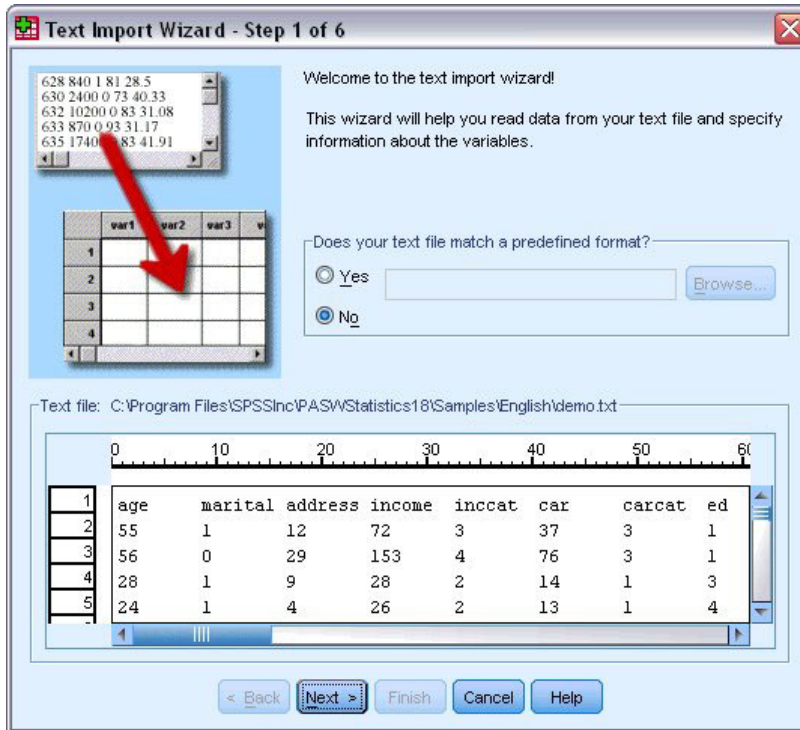


Figura 15. Asistente para la importación de texto: paso 1 de 6

4. En el Paso 1, se puede seleccionar un formato predefinido o crear uno nuevo en el asistente. Seleccione **No** para indicar que se debe crear un nuevo formato.
5. Pulse en **Siguiente** para continuar.
Como ya se ha comentado, este archivo utiliza un formato delimitado por tabulaciones. Asimismo, los nombres de variables se definen en la línea superior de este archivo.
6. En el paso 2 del asistente, seleccione **Delimitadas** para indicar que los datos utilizan una estructura de formato delimitado.
7. Seleccione **Sí** para indicar que se deben leer los nombres de las variables de la parte superior del archivo.
8. Pulse en **Siguiente** para continuar.
9. En paso 3, introduzca 2 para el número de línea donde comienza el primer caso de datos (ya que los nombres de variables están en la primera línea).
10. Mantenga los valores predeterminados del resto de este paso y pulse en **Siguiente** para continuar.
La vista previa de datos del paso 4 proporciona una forma rápida de asegurarse de los datos se han leído correctamente.
11. Seleccione **Tabulador** y desactive las demás opciones para delimitadores.
12. Pulse en **Siguiente** para continuar.
Como es posible que se hayan modificado los nombres de las variables conforme a las normas de denominación, el paso 5 ofrece la oportunidad de editar cualquier nombre no deseado.

También se pueden definir aquí los tipos de datos. Por ejemplo, se puede asumir que la variable de ingresos está destinada a contener una cierta cantidad de dólares.

Para cambiar un tipo de datos:

13. En Vista previa de datos, seleccione la variable que desea modificar, que, en este caso, es *Ingresos*.
14. Seleccione **Dólar** de la lista desplegable Formato de datos.

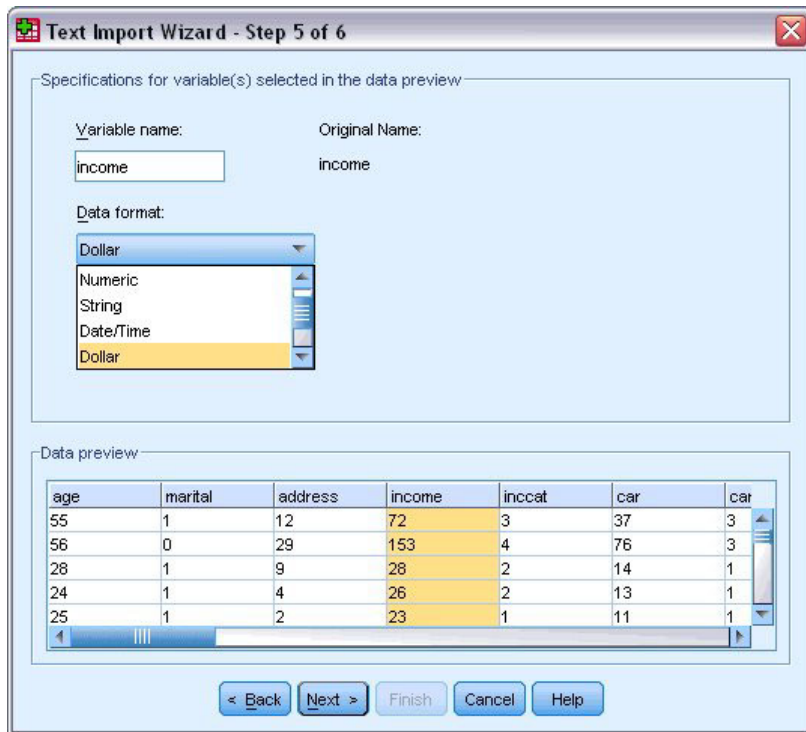


Figura 16. Cambio del tipo de datos

15. Pulse en **Siguiente** para continuar.
16. Mantenga las selecciones predeterminadas del último paso y pulse en **Finalizar** para importar los datos.

Capítulo 3. Uso del Editor de datos

En el Editor de datos se muestra el contenido del archivo de datos activo. La información incluida en el Editor de datos consta de variables y casos.

- En la Vista de datos, las columnas representan las variables y las filas representan los casos (observaciones).
- En la Vista de variables, cada fila es una variable y cada columna es un atributo asociado a dicha variable.

Las variables se utilizan para representar los diferentes tipos de datos que haya recopilado. Una analogía común es la de una encuesta. La respuesta a cada pregunta de una encuesta equivale a una variable. Las variables son de distintos tipos, incluyendo números, cadenas, moneda y fechas.

Introducción de datos numéricos

Los datos se pueden introducir en el Editor de datos, lo que puede resultar útil para archivos de datos pequeños o para realizar tareas de edición menores en archivos de datos más grandes.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.

Debe definir las variables que va a utilizar. En este caso, sólo se necesitan tres variables: *edad*, *estado civil* e *ingresos*.

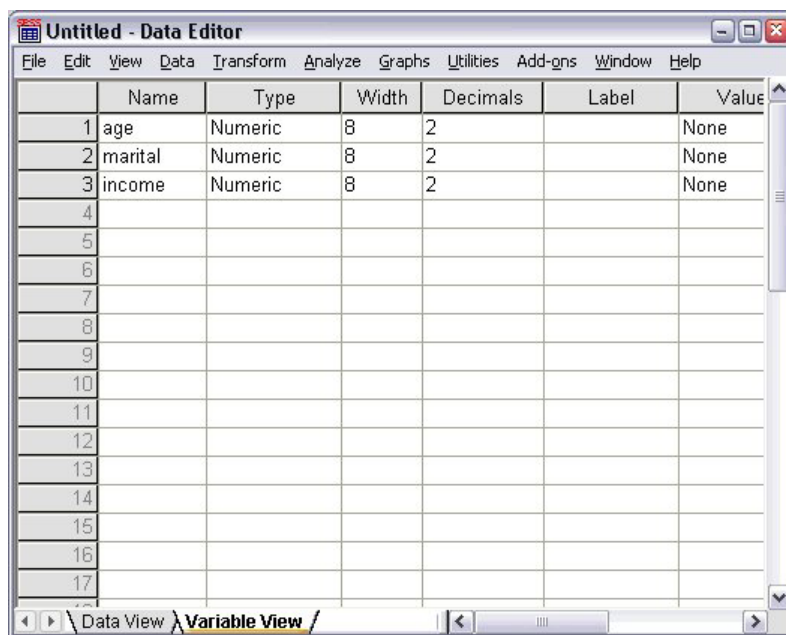


Figura 17. Nombres de variables en la Vista de variables

2. Escriba edad en la primera fila de la primera columna.
3. En la segunda fila, escriba ecivil.
4. En la tercera fila, escriba ingresos.

Se asignará automáticamente el tipo de datos Numérico a las nuevas variables.

Si no introduce nombres de variable, se crearán nombres exclusivos automáticamente. No obstante, estos nombres no son descriptivos ni recomendables para archivos de datos más grandes.

5. Pulse en la pestaña **Vista de datos** para seguir introduciendo los datos.

Los nombres introducidos en la pestaña Vista de variables serán, a partir de ahora, los encabezados de las tres primeras columnas de Vista de datos.

Empiece a introducir datos en la primera fila de la primera columna.

	age	marital	income	var	var	var
1	55.00	1.00	72000.00			
2	53.00	.00	153000.0			
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						

Figura 18. Valores introducidos en Vista de datos

6. En la columna *edad*, escriba 55.

7. En la columna *civil*, escriba 1.

8. En la columna *ingresos*, escriba 72000.

9. Mueva el cursor hasta la segunda fila de la primera columna para añadir los datos del siguiente sujeto.

10. En la columna *edad*, escriba 53.

11. En la columna *civil*, escriba 0.

12. En la columna *ingresos*, escriba 153000.

En la actualidad, en las columnas de la *edad* y el *estado civil* aparecen separaciones decimales, aunque sus valores sean números enteros. Para ocultar las separaciones decimales en estas variables:

13. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.

14. En la columna *Decimales* de la fila *edad*, escriba 0 para ocultar el decimal.

15. En la columna *Decimales* de la fila *civil*, escriba 0 para ocultar el decimal.

Introducción de datos de cadena

También se pueden introducir datos no numéricos como, por ejemplo, cadenas de texto, en el Editor de datos.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.

2. Escriba *sexo* como nombre de variable en la primera casilla de la primera fila vacía.

3. Pulse en la casilla *Tipo* que hay junto a su entrada.

4. Pulse en el botón que hay a la derecha de la casilla *Tipo* para abrir el cuadro de diálogo Tipo de variable.
5. Seleccione **Cadena** para especificar el tipo de variable.
6. Pulse en **Aceptar** para guardar la selección y volver al Editor de datos.

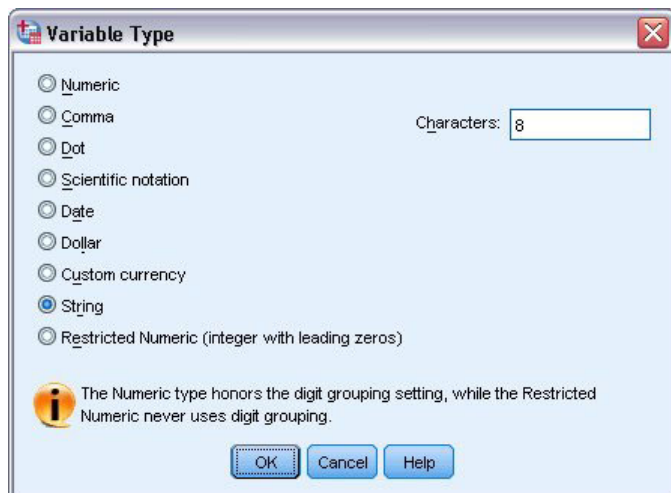


Figura 19. Cuadro de diálogo Tipo de variable

Definición de datos

Además de definir los tipos de datos, también puede definir etiquetas descriptivas de variable y etiquetas de valor para los nombres de variables y los valores de datos. Estas etiquetas descriptivas se emplean en los gráficos e informes estadísticos.

Adición de etiquetas de variable

Las etiquetas están destinadas a proporcionar descripciones de las variables. Estas descripciones suelen ser versiones extendidas de los nombres de variable. Las etiquetas pueden ocupar hasta 255 bytes. Estas etiquetas se utilizan en los resultados para identificar a las diferentes variables.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. En la columna *Etiqueta* de la fila *edad*, escriba Edad del encuestado.
3. En la columna *Etiqueta* de la fila *ecivil*, escriba Estado civil.
4. En la columna *Etiqueta* de la fila *ingresos*, escriba Ingresos familiares.
5. En la columna *Etiqueta* de la fila *sexo*, escriba Sexo.

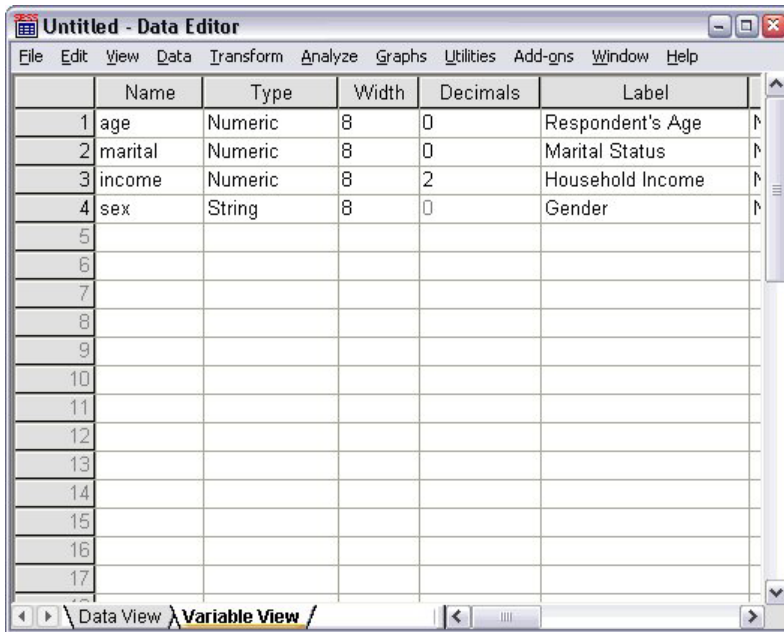


Figura 20. Etiquetas de variable introducidas en Vista de variables

Cambio de tipo de variable y formato

En la columna *Tipo* aparecen los tipos de datos actuales de cada variable. Los tipos de datos más comunes son los numéricos y las cadenas, pero se admiten otros muchos formatos. En el archivo de datos actual, la variable de *ingresos* se define como un tipo numérico.

1. Pulse en la casilla *Tipo* para la fila *ingresos* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Tipo de variable.
2. Seleccione **Dólar**.

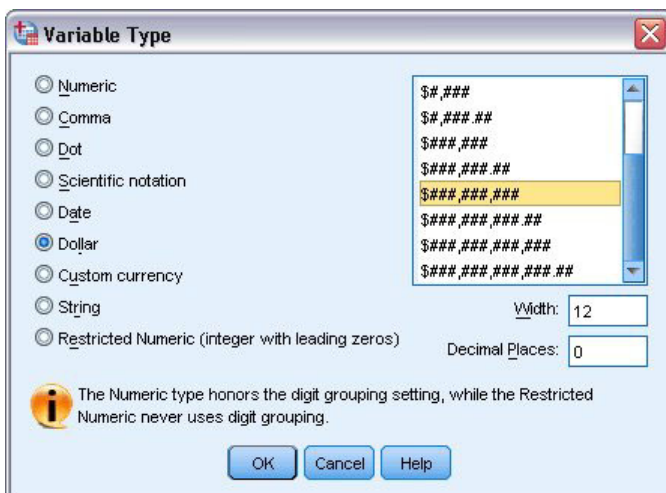


Figura 21. Cuadro de diálogo Tipo de variable

Aparecerán las opciones de formato del tipo de datos seleccionado en ese momento.

3. Para el formato de la moneda de este ejemplo, seleccione **\$###,###,###**.
4. Pulse en **Aceptar** para guardar los cambios.

Añadir etiquetas de valores

Las etiquetas de valor proporcionan un método para correlacionar los valores de variable con una etiqueta de texto. En este ejemplo, hay dos valores aceptables para la variable *ecivil*. Un valor 0 significa que el sujeto está soltero/a, y un valor 1, que está casado/a.

1. Pulse en la casilla *Valores* de la fila *ecivil* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Etiquetas de valor.

El **valor** corresponde al valor numérico real.

La **etiqueta de valor** es la etiqueta de cadena que se aplica al valor numérico especificado.

2. Escriba 0 en el campo Valor.
3. Escriba Sin casar en el campo Etiqueta.
4. Pulse en **Añadir** para que se añada esta etiqueta a la lista.

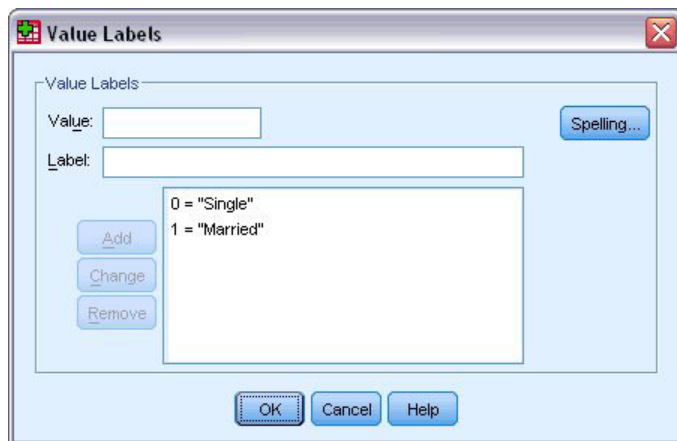


Figura 22. Cuadro de diálogo Etiquetas de valor

5. Escriba 1 en el campo Valor y, a continuación, escriba Casado en el campo Etiqueta.
6. Pulse en **Añadir** y a continuación, pulse en **Aceptar** para guardar los cambios y volver al Editor de datos.

Estas etiquetas se pueden visualizar también en la Vista de datos, lo que puede facilitar la lectura de los datos.

7. Pulse en la pestaña **Vista de datos**, que se encuentra en la parte inferior de la ventana Editor de datos.
8. Elija en los menús:

Ver > Etiquetas de valores

Las etiquetas aparecerán ahora en una lista cuando introduzca valores en el Editor de datos. Esta configuración tiene la ventaja de sugerir una respuesta válida y proporcionar una contestación más descriptiva.

Si el elemento de menú Etiquetas de valor ya está activo (aparece una marca de verificación junto al mismo), al volver a elegir **Etiquetas de valor** se *desactivará* la presentación de las etiquetas de valor.

Tratamiento de los datos perdidos

Los datos perdidos o no válidos son generalmente tan comunes que no se deben ignorar. Puede que las personas que han respondido a una encuesta se nieguen a contestar ciertas preguntas, no sepan la respuesta o contesten de forma inesperada. Si no filtra ni identifica estos datos, puede que el análisis no proporcione resultados exactos.

Los datos numéricos, campos de datos vacíos o campos con entradas no válidas se convierten a valores perdidos del sistema, lo que se puede identificar mediante un punto individual.

El motivo por el que se pierde un valor puede ser importante para el análisis. Por ejemplo, puede resultar útil realizar una distinción entre los encuestados que se negaron a contestar una pregunta y los encuestados que no contestaron porque la pregunta no era aplicable.

Valores perdidos de una variable numérica

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. Pulse en la casilla *Perdido* de la fila *edad* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Valores perdidos.

Es este cuadro de diálogo puede especificar hasta tres valores perdidos diferentes, o puede especificar un rango de valores más un valor de tipo discreto adicional.



Figura 23. Cuadro de diálogo Valores perdidos

3. Seleccione **Valores perdidos discretos**.
4. Escriba 999 en el primer cuadro de texto u deje los otros dos cuadros de texto en blanco.
5. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.
Ahora que se ha añadido el valor de datos perdidos se puede aplicar una etiqueta a ese valor.
6. Pulse en la casilla *Valores* de la fila *edad* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Etiquetas de valor.
7. Escriba 999 en el campo Valor.
8. Escriba Sin respuesta en el campo Etiqueta de valor.
9. Pulse en **Añadir** para que se añada esta etiqueta al archivo de datos.
10. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.

Valores perdidos de una variable de cadena

Los valores perdidos en variables de cadena se tratan de forma parecida a los valores perdidos en variables numéricas. Sin embargo, a diferencia de las variables numéricas, los campos vacíos de las variables de cadena no se consideran perdidos del sistema. En lugar de eso, se interpretan como una cadena vacía.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. Pulse en la casilla *Perdido* de la fila *sexo* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Valores perdidos.
3. Seleccione **Valores perdidos discretos**.
4. Escriba SR en el primer cuadro de texto.

Los valores perdidos de las variables de cadena distinguen entre mayúsculas y minúsculas. De este modo, un valor de *sr* no se considerará un valor perdido.

5. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.
Ahora puede añadir una etiqueta en el valor perdido.
6. Pulse en la casilla *Valores* de la fila *sexo* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Etiquetas de valor.
7. Escriba SR en el campo Valor.
8. Escriba Sin respuesta en el campo Etiqueta de valor.
9. Pulse en **Añadir** para que se añada esta etiqueta al proyecto.
10. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.

Capítulo 4. Examen de estadísticos de resumen para variables individuales

Esta sección describe medidas de resumen sencillas y la influencia del nivel de medición de una variable en los tipos de estadísticos que se deben utilizar. Se utilizará el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, "Archivos muestrales", en la página 81 para obtener más información.

Nivel de medición

Hay diferentes medidas de resumen adecuadas a diferentes tipos de datos dependiendo del nivel de medición:

Catagóricas. Datos con un número limitado de valores o categorías distintas (por ejemplo, género o estado civil). También se hace referencia a estos datos como **datos cualitativos**. Las variables catagóricas pueden ser variables de cadena (alfanuméricas) o variables numéricas que utilizan códigos numéricos para representar a las categorías (por ejemplo, 0 = *Soltero/a* y 1 = *Casado/a*). Hay dos tipos básicos de datos catagóricos:

- **Nominal.** Datos catagóricos en los que las categorías no tienen un orden inherente. Por ejemplo, una categoría laboral de *ventas* no es mayor o menor que una categoría laboral de *marketing* o *investigación*.
- **Ordinal.** Datos catagóricos en los que las categorías tienen un orden significativo, pero sin una distancia medible entre las categorías. Por ejemplo, hay un orden para los valores *alto*, *medio* y *bajo* pero no se puede calcular la "distancia" entre los valores.

Escalas. Datos medidos en una escala de intervalo o de razón en los que los valores de los datos indican el orden de los valores y la distancia entre ellos. Por ejemplo, un salario de 72.195€ es superior a un salario de 52.398€ y la distancia entre ambos valores es 19.797€. También se hace referencia a estos datos como **datos cuantitativos** o **continuos**.

Medidas de resumen para datos catagóricos

Para los datos catagóricos, la medida de resumen más habitual es el número o el porcentaje de casos de cada categoría. La **moda** es la categoría que contiene el mayor número de casos. Para los datos ordinales, la **mediana** (el valor por debajo y por encima del cual se encuentran la mitad de los casos) también puede ser una medida de resumen útil si hay un gran número de categorías.

El procedimiento Frecuencias genera tablas de frecuencias que muestran el número y el porcentaje de los casos de cada valor observado de una variable.

1. Elija en los menús:

Analizar > Estadísticos descriptivos > Frecuencias...

Nota: esta característica requiere la opción Statistics Base.

2. Seleccione *Tiene PDA [pda]* y *Tiene TV [tv]* y muévalas hasta la lista Variable(s).

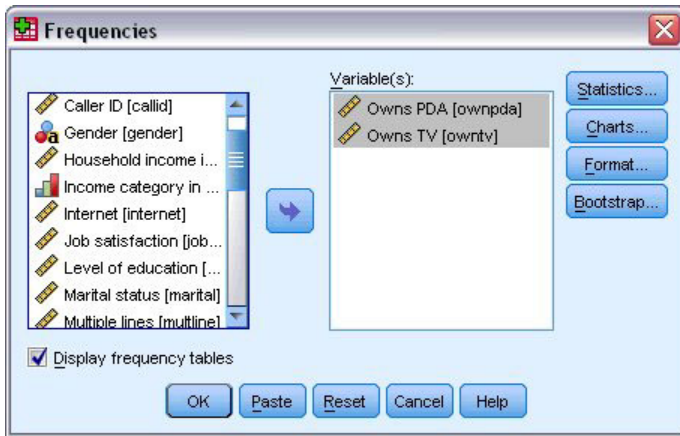


Figura 24. Variables categóricas seleccionadas para el análisis

3. Pulse en **Aceptar** para iniciar el procedimiento.

Frequency Table

Owns PDA

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	5093	79.6	79.6	79.6
	Yes	1307	20.4	20.4	100.0
Total		6400	100.0	100.0	

Owns TV

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	63	1.0	1.0	1.0
	Yes	6337	99.0	99.0	100.0
Total		6400	100.0	100.0	

Figura 25. Tablas de frecuencias

Las tablas de frecuencias aparecen en la ventana Visor. Las tablas de frecuencias revelan que sólo el 20,4% de las personas poseen una agenda digital, pero que casi todos poseen un televisor (99,0%). Aunque estos datos tal vez no constituyan grandes revelaciones, puede ser interesante descubrir algo más acerca del pequeño grupo de personas que no poseen televisores.

Gráficos para datos categóricos

Puede visualizar gráficamente la información en una tabla de frecuencias con un gráfico de barras o con un gráfico circular.

1. Vuelva a abrir el cuadro de diálogo Frecuencias. (Las dos variables deben estar seleccionadas aún).
Puede utilizar el botón Recuperar cuadros de diálogo de la barra de herramientas para regresar rápidamente a los procedimientos utilizados recientemente.



Figura 26. Botón Recuperar cuadros de diálogo

2. Pulse en **Gráficos**.
3. Seleccione **Gráficos de barras** y, a continuación, pulse en **Continuar**.
4. Pulse en **Aceptar** en el cuadro de diálogo principal para ejecutar el procedimiento.

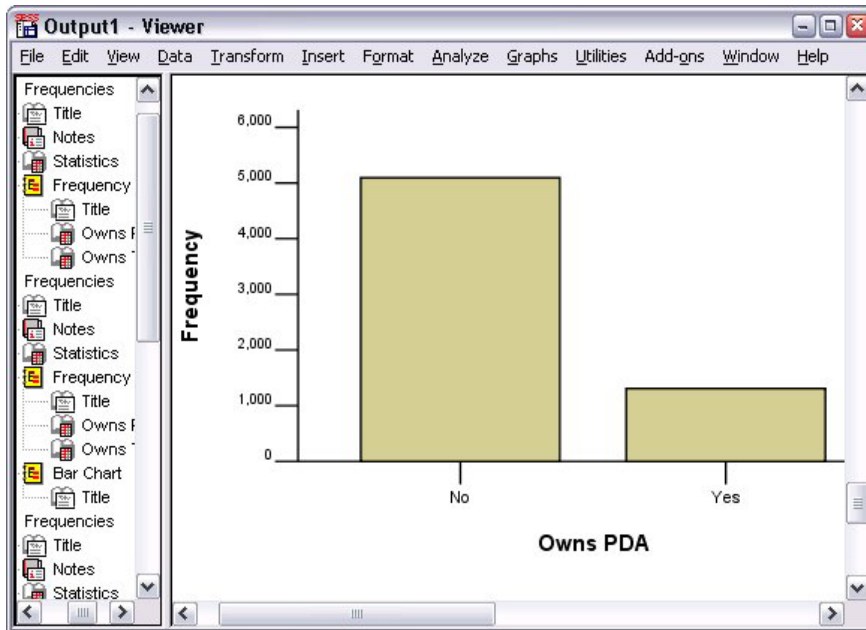


Figura 27. Gráfico de barras

Además de las tablas de frecuencias, la misma información aparece ahora en forma de gráficos de barras, con lo que es fácil ver que la mayoría de las personas no poseen agendas digitales y en cambio casi todos poseen un televisor.

Medidas de resumen para variables de escala

Hay muchas medidas de resumen disponibles para variables de escala, incluyendo:

- **Medidas de tendencia central.** Las medidas de tendencia central más comunes son la **media** (media aritmética) y la **mediana** (valor por debajo y por encima del cual se encuentran la mitad de los casos).
- **Medidas de dispersión.** Los estadísticos que miden la dispersión o variación en los datos incluyen la desviación estándar, mínimo y máximo.

1. Vuelva a abrir el cuadro de diálogo Frecuencias.
2. Pulse en **Restablecer** para borrar los ajustes anteriores.
3. Seleccione *Ingresos del hogar en miles [ingres]* y muévala hasta la lista Variable(s).
4. Pulse en **Estadísticos**.
5. Seleccione **Media, Mediana, Desv. estándar, Mínimo y Máximo**.
6. Pulse en **Continuar**.

- Anule la selección de **Mostrar tablas de frecuencias** en el cuadro de diálogo principal. (Por lo general, las tablas de frecuencias no son útiles para variables de escala ya que debe haber casi tantos valores distintos como casos en el archivo de datos).
- Pulse en **Aceptar** para iniciar el procedimiento.

La tabla Frecuencias: Estadísticos aparece en la ventana Visor.

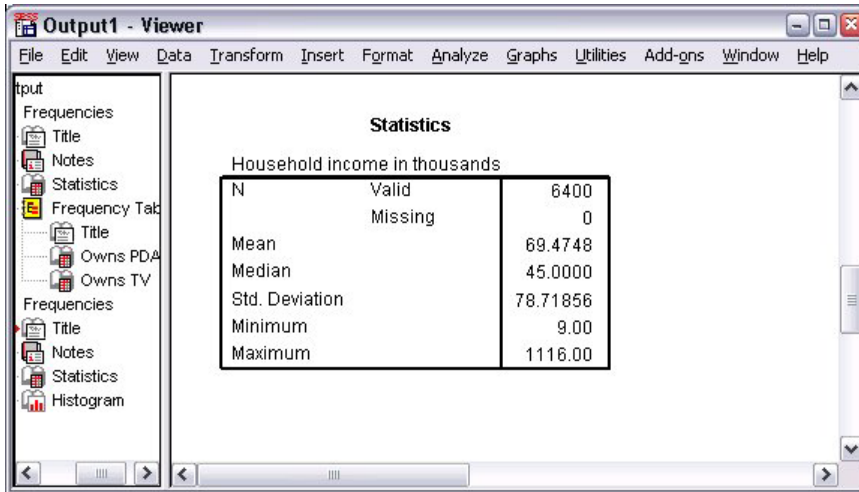


Figura 28. Tabla Frecuencias: Estadísticos

En este ejemplo, hay una importante diferencia entre la media y la mediana. La media es casi 25.000 mayor que la mediana, lo que indica que los valores no se distribuyen normalmente. Puede comprobar visualmente la distribución con un histograma.

Histogramas para variables de escala

- Vuelva a abrir el cuadro de diálogo Frecuencias.
- Pulse en **Gráficos**.
- Seleccione **Histogramas** y **Con curva normal**.
- Pulse en **Continuar** y, a continuación, pulse en **Aceptar** en el cuadro de diálogo principal para ejecutar el procedimiento.

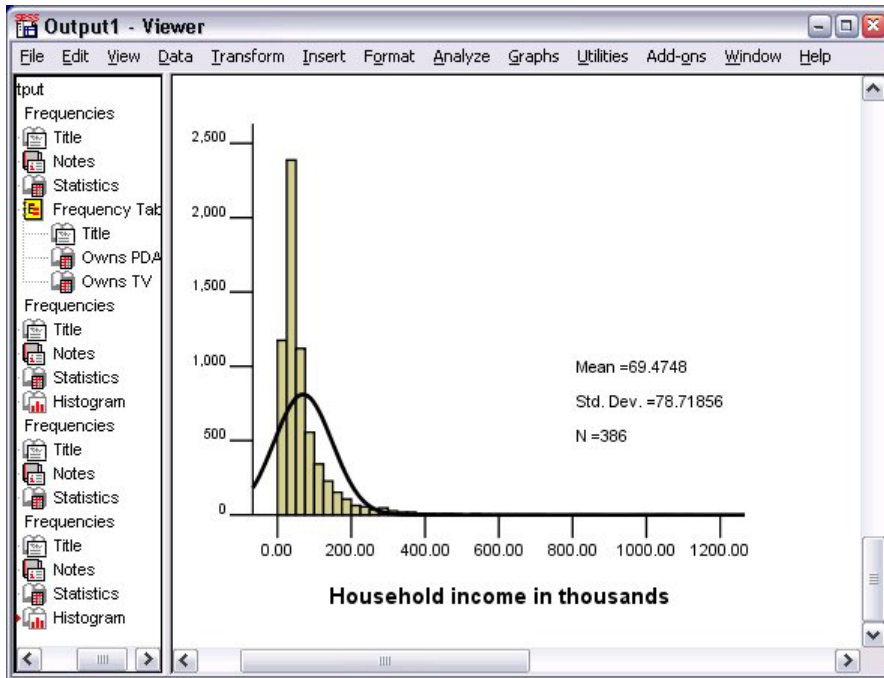


Figura 29. Histograma

La gran mayoría de casos se agrupan en la parte inferior de la escala, estando incluidos la mayoría de ellos por debajo de 100.000. Sin embargo, hay algunos casos en el rango 500.000 y más allá (demasiado escasos para que sean visibles sin modificar el histograma). Estos valores altos para sólo unos pocos casos tienen un efecto importante sobre la media y muy escaso o casi nulo sobre la mediana, lo que hace que la mediana sea un indicador más exacto de la tendencia central en este ejemplo.

Capítulo 5. Creación y edición de gráficos

Puede crear y editar una gran variedad de tipos de gráficos. En este capítulo, crearemos y editaremos gráficos de barras. Puede aplicar los principios a cualquier tipo de gráfico.

Procedimiento básico de creación de gráficos

Para mostrar el procedimiento básico de creación de gráficos, crearemos un gráfico de barras de los ingresos medios para los diferentes niveles de satisfacción laboral. En este ejemplo se utiliza el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

1. Elija en los menús:

Gráficos > Generador de gráficos...

El cuadro de diálogo Generador de gráficos es una ventana interactiva que le permite ver una vista previa del aspecto que tendrá un gráfico al generarlo.

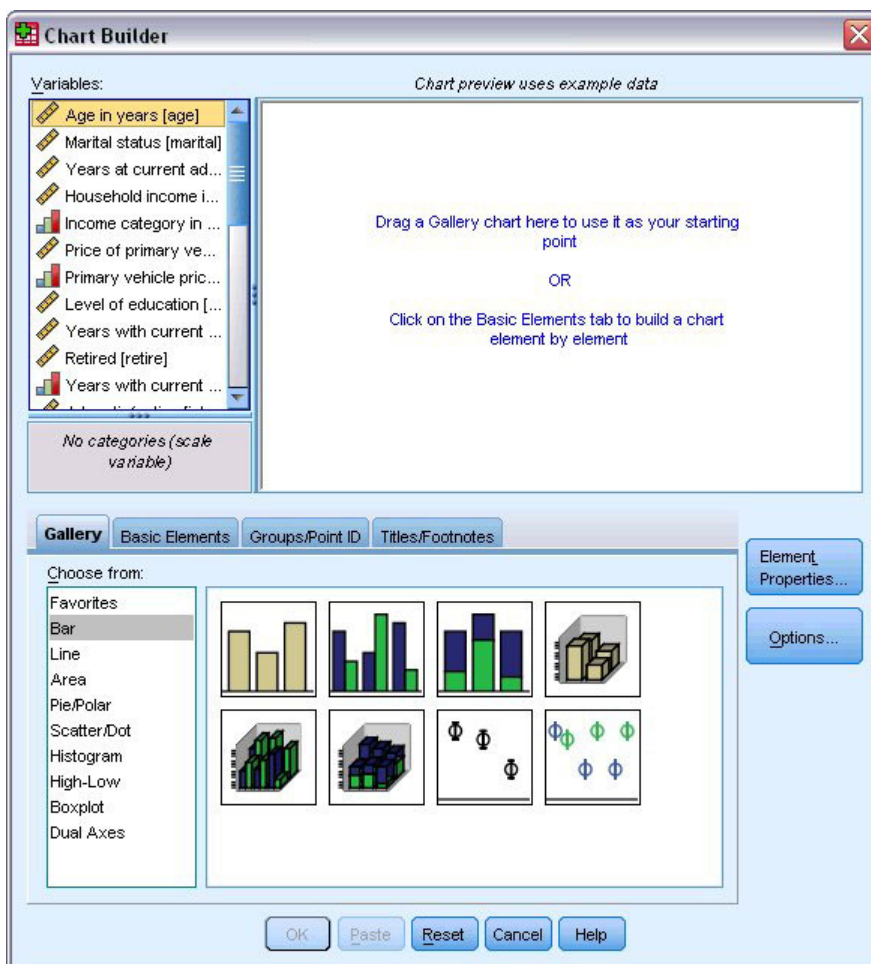


Figura 30. Cuadro de diálogo Generador de gráficos

Uso de la galería del generador de gráficos

1. Pulse en la pestaña **Galería** si no está seleccionada.

La galería incluye un gran número de gráficos predefinidos, organizados por tipo de gráfico. La pestaña Elementos básicos también ofrece los elementos básicos (como ejes y elementos gráficos) necesarios para crear gráficos a partir de cero, pero es más fácil utilizar la galería.

2. Pulse en la pestaña **Barra** si no está seleccionada.

Los iconos que representan los gráficos de barras que hay disponibles en la galería aparecen en el cuadro de diálogo. Las imágenes deberían proporcionar suficiente información para identificar el tipo de gráfico concreto. Si necesita más información, también puede ver una ayuda contextual con una descripción del gráfico si detiene el cursor sobre un icono.

3. Arrastre el icono correspondiente al gráfico de barras simple al "lienzo", que es la zona grande que hay encima de la galería. El generador de gráficos muestra una vista previa del gráfico en el lienzo. Observe que los datos utilizados para dibujar el gráfico no son los datos reales. Son datos de ejemplo.

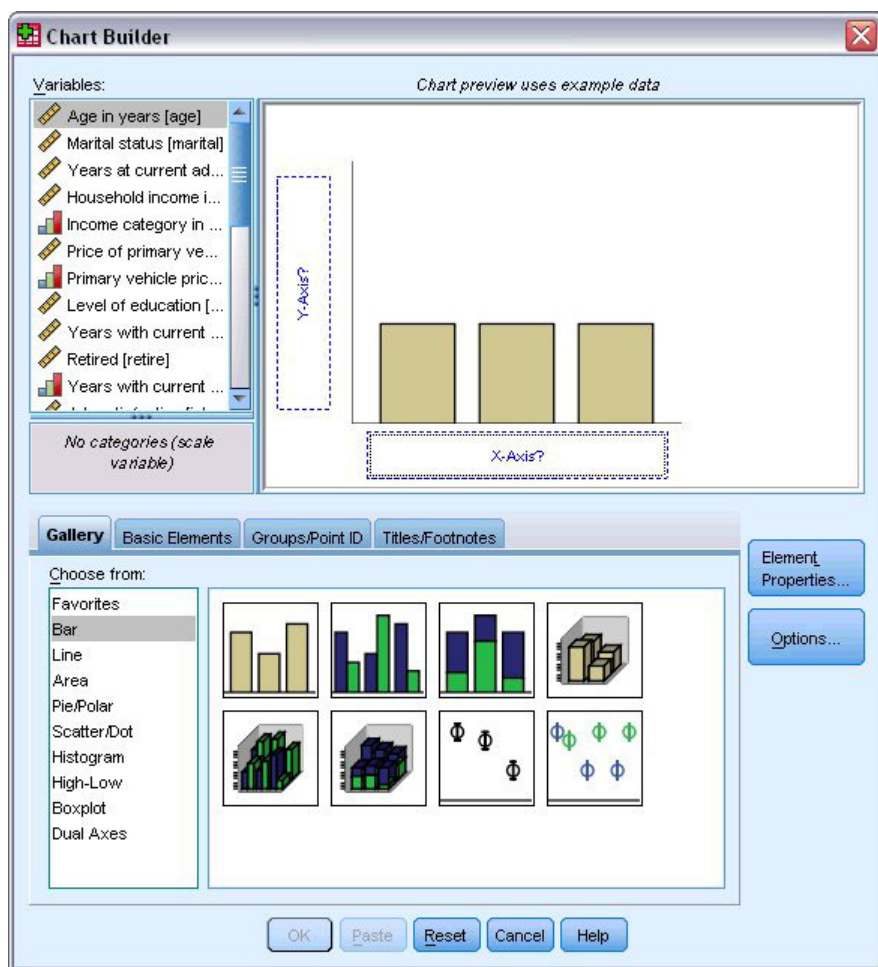


Figura 31. Gráfico de barras en el lienzo del generador de gráficos

Definición de variables y estadísticos

Aunque hay un gráfico en el lienzo, no está completo ya que no hay ninguna variable ni estadístico que controle la altura de las barras ni se ha especificado la variable que corresponde a cada barra. No se puede tener un gráfico sin variables ni estadísticos. Puede añadir variables arrastrándolas desde la lista de variables, situada a la izquierda del lienzo.

El nivel de medición de una variable es importante en el generador de gráficos. Vamos a utilizar la variable *Satisfacción laboral* en el eje X. Sin embargo, el icono (que tiene el aspecto de una regla) que hay junto a la variable indica que el nivel de medición se ha definido como escala. Para crear el gráfico correcto, deberá utilizar un nivel de medición categórico. En vez de retroceder y cambiar el nivel de medición en la Vista de variables, puede cambiar temporalmente el nivel de medición en el generador de gráficos.

1. Pulse con el botón derecho del ratón en *Satisfacción laboral* en la lista de variables y elija **Ordinal**. El nivel de medición Ordinal es adecuado ya que las categorías de *Satisfacción laboral* se pueden ordenar por nivel de satisfacción. Observe que el icono cambiará al modificar el nivel de medición.
2. Ahora arrastre *Satisfacción laboral* de la lista de variables a la zona de colocación del eje X.

La zona de colocación del eje *y* de forma predeterminada tiene el estadístico *Recuento*. Si desea utilizar otro estadístico (como el porcentaje o la media), puede cambiarlo fácilmente. No utilizará ninguno de estos estadísticos en este ejemplo, pero examinaremos el proceso necesario por si necesita cambiar este estadístico en otro momento.

3. Pulse en **Propiedades del elemento** para acceder a la ventana Propiedades del elemento.

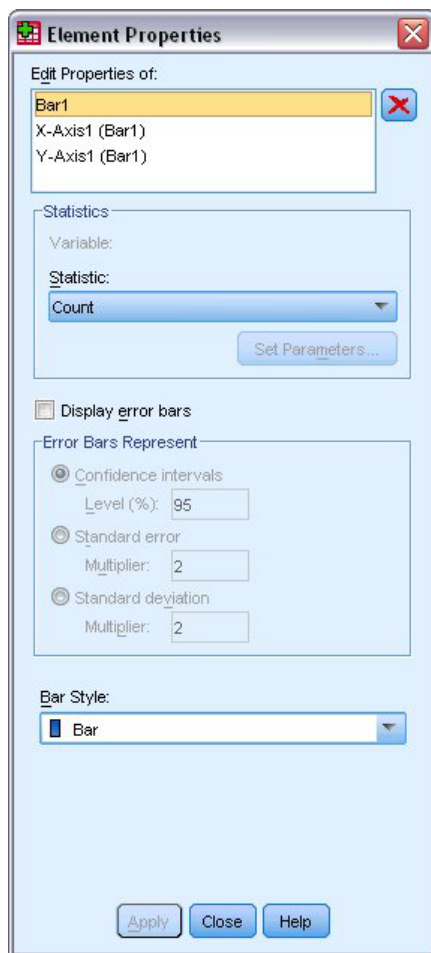


Figura 32. Ventana Propiedades del elemento

La ventana Propiedades del elemento permite cambiar las propiedades de los diferentes elementos del gráfico. Entre éstos se incluyen los elementos gráficos (como las barras del gráfico de barras) y los ejes del gráfico. Seleccione uno de los elementos de la lista Editar propiedades de para cambiar las propiedades asociadas a dicho elemento. Observe también la *X* roja que aparece en la parte derecha

de la lista. Este botón permite eliminar un elemento gráfico del lienzo. Al estar seleccionado **Barra1**, las propiedades que aparecen se aplican a los elementos gráficos, concretamente al elemento gráfico de barra.

La lista desplegable Estadístico muestra los estadísticos concretos que están disponibles. Normalmente, para todos los tipos de gráficos están disponibles los mismos estadísticos. No olvide que para algunos estadísticos es necesario que la zona de colocación del eje *y* contenga una variable.

4. Vuelva al cuadro de diálogo Generador de gráficos y arrastre *Ingresos del hogar en miles* de la lista de variables a la zona de colocación del eje *y*. Al ser la variable del eje *Y* escalar y la variable del eje *X* escalar (ordinal es un tipo de nivel de medición categórico), en la zona de colocación del eje *Y* aparece de forma predeterminada el estadístico *Media*. Estas son las variables y los estadísticos que desea, por lo que no es necesario cambiar las propiedades de los elementos.

Adición de texto

También puede añadir títulos y notas al pie al gráfico.

1. Pulse en la pestaña **Títulos/notas al pie**.
2. Seleccione **Título 1**.

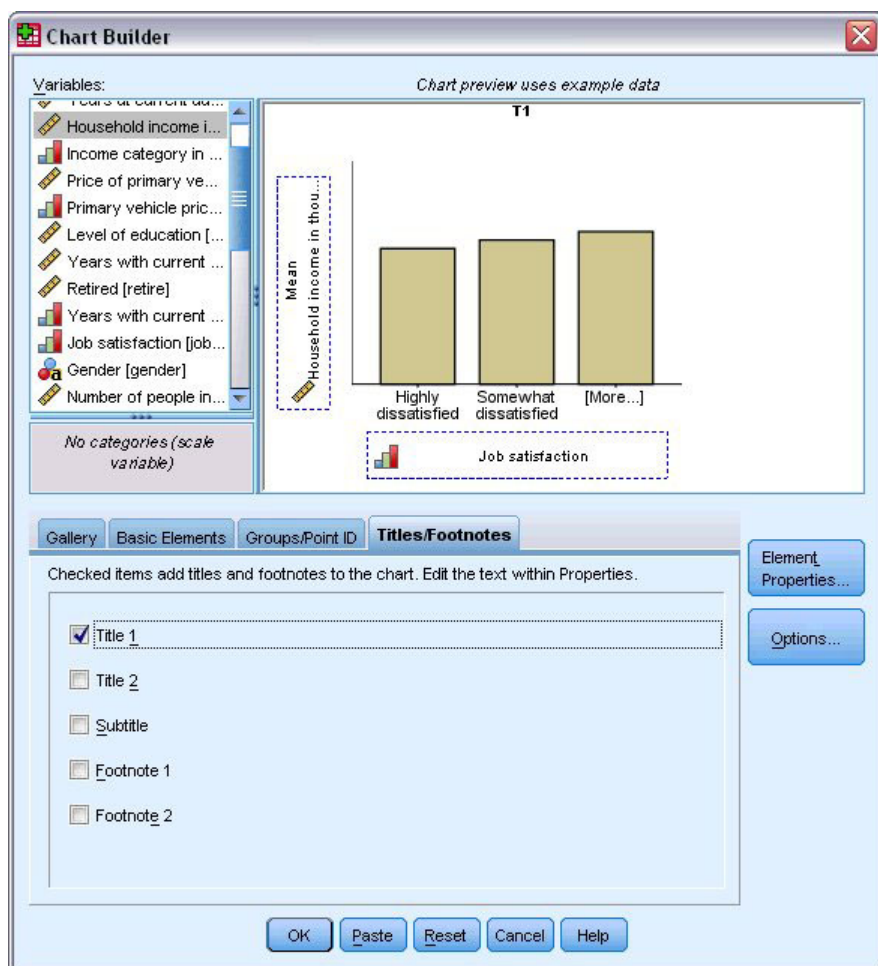


Figura 33. Título 1 mostrado en el lienzo

El título aparecerá en el lienzo con la etiqueta **T1**.

3. En la ventana Propiedades del elemento, seleccione **Título 1** en la lista Editar propiedades de.
4. En el cuadro de texto Contenido, escriba Ingresos por satisfacción laboral. Éste es el texto que mostrará el título.

5. Pulse en **Aplicar** para guardar el texto. Aunque el texto no se verá en el generador de gráficos, aparecerá al generar el gráfico.

Creación del gráfico

1. Pulse en **Aceptar** para crear el gráfico de barras.

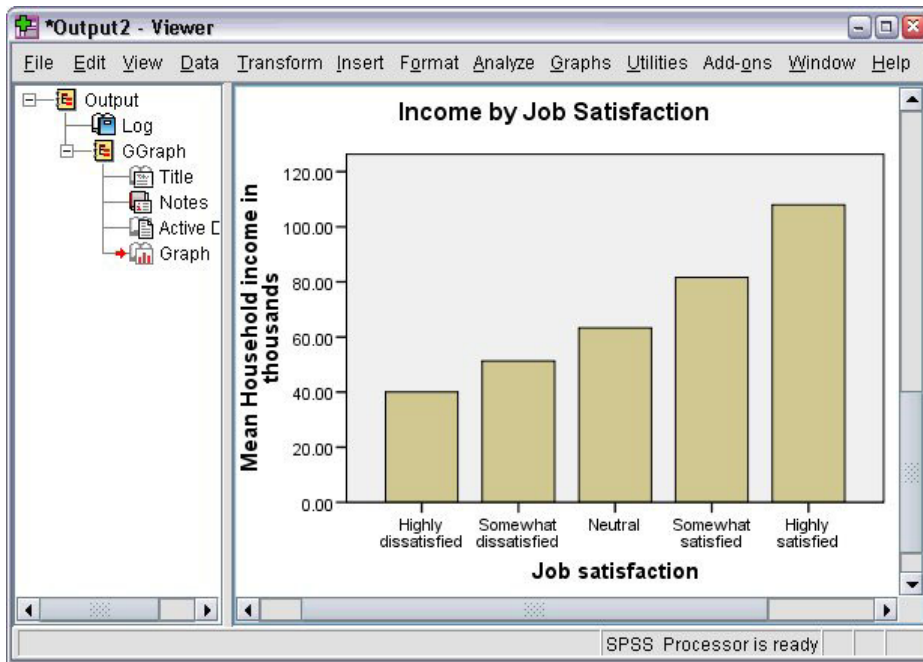


Figura 34. Gráfico de barras

El gráfico de barras muestra que los encuestados que están más satisfechos con su trabajo suelen tener ingresos más altos.

Capítulo 6. Cómo trabajar con resultados

Los resultados de ejecutar un procedimiento estadístico se muestran en el Visor. Los resultados obtenidos pueden ser tablas estadísticas, diagramas, gráficos o texto, dependiendo de las opciones tomadas al ejecutar el procedimiento. Esta sección utiliza los archivos *viewertut.spv* y *demo.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Uso del Visor

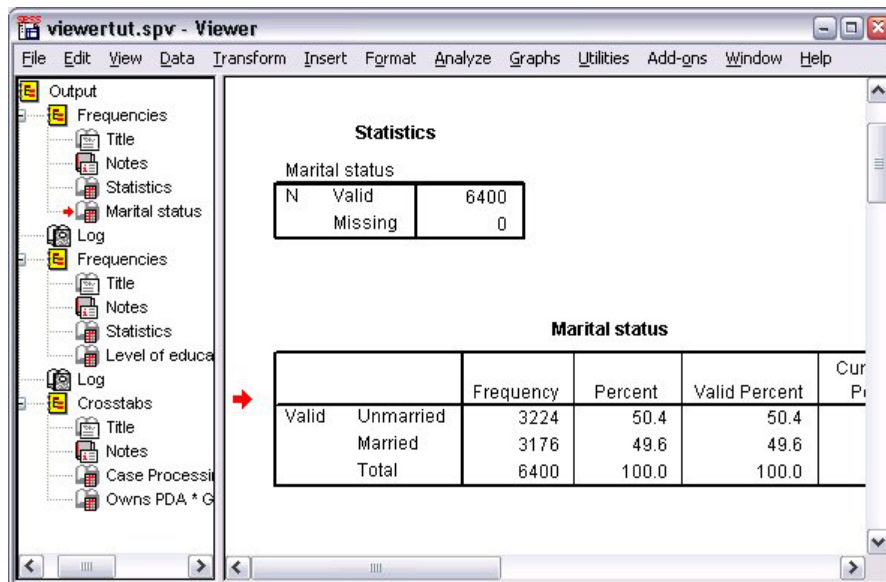


Figura 35. Visor

La ventana Visor se divide en dos paneles. El **panel de titulares** contiene un esquema de titulares con toda la información almacenada en el Visor. El **panel de contenidos** contiene tablas estadísticas, gráficos y resultados de texto.

Utilice las barras de desplazamiento para desplazarse a través de los contenidos de la ventana, tanto vertical como horizontalmente. Para simplificar el acceso, pulse en un elemento en el panel de titulares para visualizarlo en el panel de contenidos.

1. Pulse y arrastre el borde derecho del panel de titulares para modificar su anchura.
Un icono de un libro abierto en el panel de titulares indica que está visible en el Visor, aunque es posible que actualmente no esté en la parte visible del panel de contenidos.
2. Para ocultar una tabla o un gráfico, pulse dos veces en su icono de libro del panel de titulares.
El icono de libro abierto cambiará a un icono de libro cerrado, indicando que la información asociada está oculta.
3. Pulse dos veces en el icono de libro cerrado para volver a mostrar los resultados ocultos.
También puede ocultar los resultados de un procedimiento estadístico determinado o los resultados del Visor.
4. Para ocultar todos los resultados, pulse la casilla con el signo menos (-) a la izquierda del procedimiento que desea ocultar o en la casilla que está junto al elemento situado más arriba en los titulares.

Los titulares se contraen, indicando visualmente que dichos resultados están ocultos.

También puede cambiar el orden en el que se muestran los resultados.

5. En el panel de titulares, pulse en los elementos que desea mover.
6. Arrastre los elementos seleccionados a una nueva ubicación en los titulares.

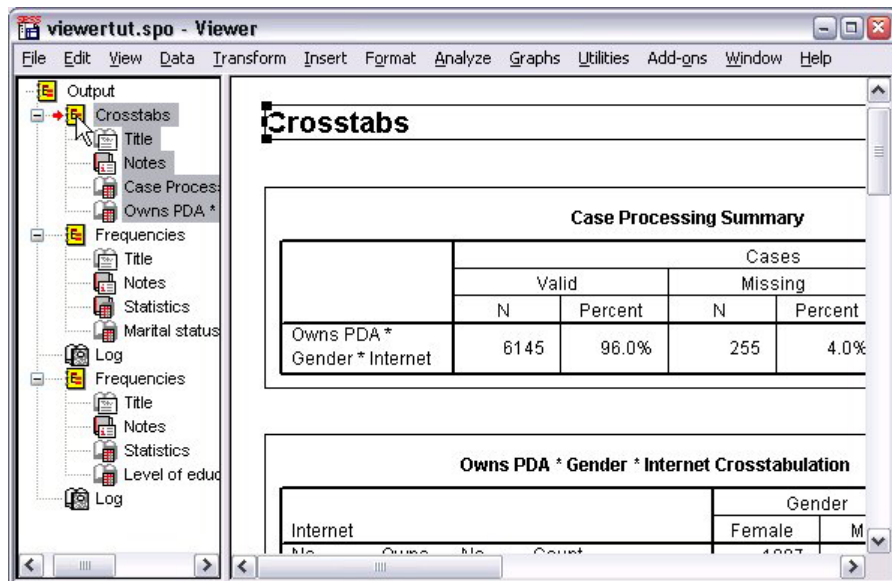


Figura 36. Resultados reordenados en el Visor

También se pueden mover los elementos de los resultados pulsándolos y arrastrándolos en el panel de contenidos.

Uso del Editor de tablas pivote

Los resultados de la mayoría de los procedimientos estadísticos se muestran en las **tablas pivote**.

Acceso a las definiciones de resultados

Muchos términos estadísticos se muestran en los resultados. Se puede acceder directamente a las definiciones de dichos términos en el propio Visor.

1. Pulse dos veces en la tabla *Tabulación cruzada Tiene PDA * Sexo * Internet*.
2. Pulse con el botón derecho en *Recuento esperado* y elija **¿Qué es esto?** en el menú emergente.

La definición se mostrará en una ventana emergente.

Owns PDA * Gender * Internet Crosstabulation

Internet			Gender		Total	
			Female	Male		
No	Owns PDA	No	Count	1897	1962	3859
			Expected Count	1903.4	1955.6	3859.0
		% within Owns PDA	50.3%	49.7%	100.0%	
	Total	Count	2224	2285	4509	
		Expected Count	2224.0	2285.0	4509.0	
Yes	Owns PDA	No	Count	513	512	1025
			Expected Count	512.5	512.5	1025.0
		% within Owns PDA	50.0%	50.0%	100.0%	
	Yes	Count	305	306	611	
		Expected Count	305.5	305.5	611.0	

The number of cases that would be expected in the cell if the row and column variables are statistically independent or unrelated to one another.

Figura 37. Definición emergente

Tablas de pivotado

Es posible que las tablas generadas de forma predeterminada no muestren la información con la claridad o nitidez deseadas. Con las tablas pivote, se pueden transponer filas y columnas ("voltear" la tabla), ajustar el orden de los datos en una tabla y modificar la tabla de muchas otras maneras. Por ejemplo, puede cambiar una tabla corta y ancha en una larga y estrecha, simplemente transponiendo las filas y las columnas. La modificación del diseño de la tabla no afecta a los resultados. En cambio, es una forma de mostrar la información de una manera distinta o más adecuada.

1. Si no está activada, pulse dos veces en la tabulación cruzada *Tiene PDA * Género* Internet* para activarla.
2. Si la ventana Paneles de pivotado no está visible, seleccione en los menús:

Pivote > Bandejas de pivote

Los Paneles de pivotado proporcionan una forma de mover los datos entre columnas, filas y capas.

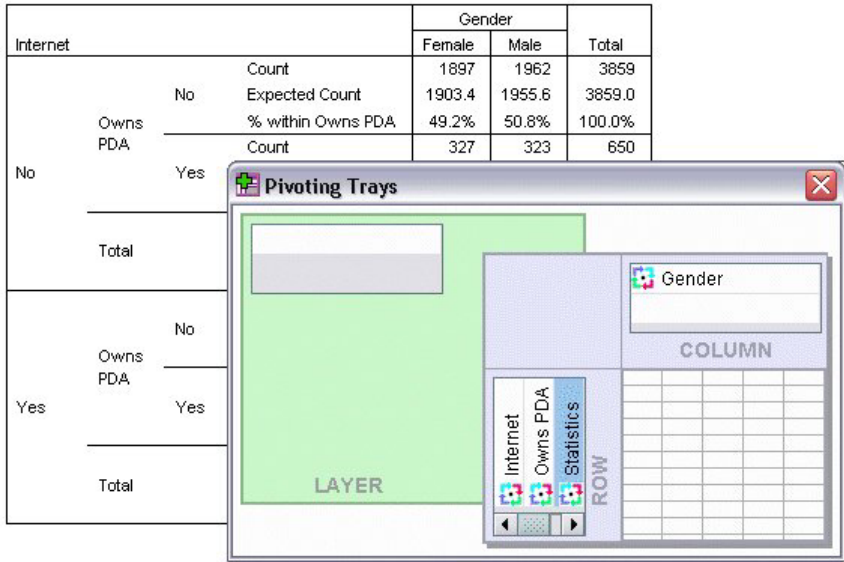


Figura 38. Bandejas de pivote

3. Arrastre el elemento *Estadísticos* de la dimensión de filas a la dimensión de columnas, debajo de *Género*. La tabla se reconfigurará de forma inmediata para reflejar los cambios.
El orden que presentan los elementos en el panel de pivotado refleja el orden de elementos en la tabla.
4. Arrastre y suelte el elemento *Tiene PDA* antes que el elemento *Internet* en la dimensión de filas para invertir el orden de estas dos filas.

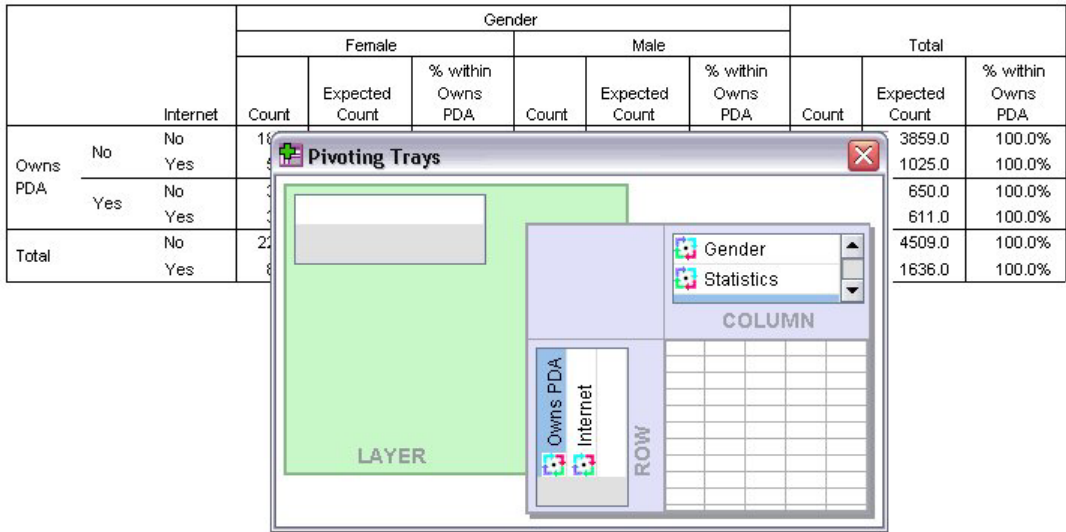


Figura 39. Intercambiar filas

Creación y visualización de capas

Las capas pueden ser de utilidad para tablas grandes con categorías de información anidadas. Al crear capas, se simplifica el aspecto de la tabla, haciendo más sencilla su lectura.

1. Arrastre el elemento *Género* desde la dimensión de columnas hasta la dimensión de capas.

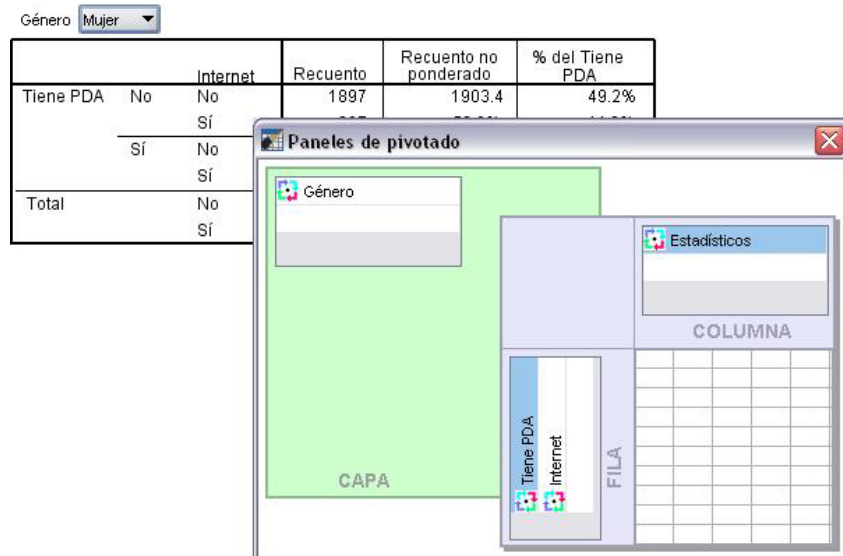


Figura 40. Icono de pivotado Género en la dimensión Capa

Para mostrar una capa distinta, seleccione una categoría de la lista desplegable en la tabla.

Edición de tablas

A menos que se haya tomado la molestia de crear un Aspecto de tabla personalizado, las tablas pivote se crearán con el formato estándar. Se puede modificar el formato de cualquier texto dentro de una tabla. Entre los formatos que se pueden cambiar se incluyen el nombre de la fuente, el tamaño de la fuente, el estilo de la fuente (negrita o cursiva) y el color.

1. Pulse dos veces en la tabla *Nivel educativo*.
2. Si la barra de herramientas Formato no está visible, seleccione en los menús:
Ver > Barra de herramientas
3. Pulse en el texto del título, *Nivel educativo*.
4. En la lista desplegable de los tamaños de fuente de la barra de herramientas, elija **12**.
5. Para cambiar el color del texto del título, pulse en la herramienta de color del texto y seleccione un color nuevo.

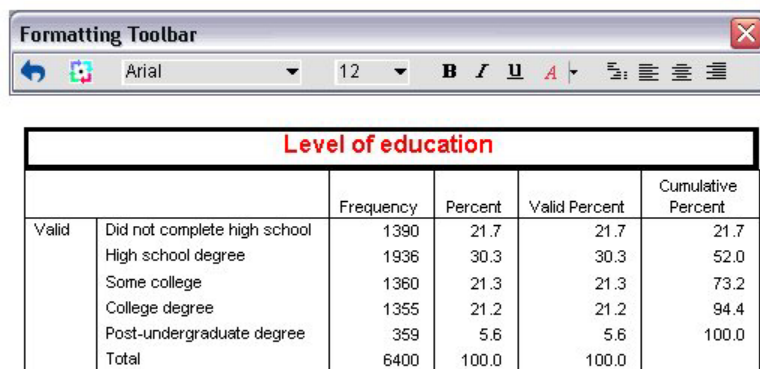


Figura 41. Nuevo formato del Texto del título en la tabla pivote

También puede editar el contenido de las tablas y etiquetas. Por ejemplo, puede cambiar el título de esta tabla.

6. Pulse dos veces en el título.
7. Escriba Nivel educativo para la nueva etiqueta.

Note: si se modifican los valores de una tabla, no se volverán a calcular los valores totales ni los demás estadísticos.

Ocultación de filas y columnas

Es posible que algunos de los datos que se muestran en las tablas no sean útiles o pueden complicar la tabla de forma innecesaria. Afortunadamente, se pueden ocultar filas y columnas completas sin perder ningún dato.

1. Si no está activada, pulse dos veces en la tabla *Nivel educativo* para activarla.
2. Pulse en la etiqueta de columna *Porcentaje válido* para seleccionarla.
3. En el menú Edición o el menú emergente que aparece al pulsar el botón derecho, seleccione: **Seleccionar > Casillas de etiquetas y datos**
4. En el menú Ver, seleccione **Ocultar**, en el menú emergente que aparece al pulsar el botón derecho, elija **Ocultar categoría**.

La columna quedará oculta pero no se eliminará.

		Frecuencia	Porcentaje	Porcentaje acumulado
Válidos	No completó el bachillerato	1390	21,7	21,7
	Bachillerato	1936	30,3	52,0
	Universitarios parciales	1360	21,3	73,2
	Universitarios	1355	21,2	94,4
	Post-universitarios	359	5,6	100,0
	Total	6400	100,0	

Figura 42. Columna *Porcentaje válido* oculta en una tabla

Para volver a mostrar la columna:

5. Elija en los menús:

Ver > Mostrar todo

Las filas se pueden ocultar y mostrar de la misma manera que las columnas.

Cambio de formato de presentación de datos

Puede cambiar fácilmente el formato de presentación de los datos en las tablas pivote.

1. Si no está activada, pulse dos veces en la tabla *Nivel educativo* para activarla.
2. Pulse en la etiqueta de columna *Porcentaje* para seleccionarla.
3. En el menú Edición o el menú emergente que aparece al pulsar el botón derecho, seleccione: **Seleccione > Datos Cells**
4. En el menú Edición o el menú emergente que aparece al pulsar el botón derecho, seleccione **Propiedades de casilla**.
5. Pulse en la pestaña **Valor de formato**.
6. Escriba 0 en el campo Decimales para ocultar todas las separaciones decimales de esta columna.

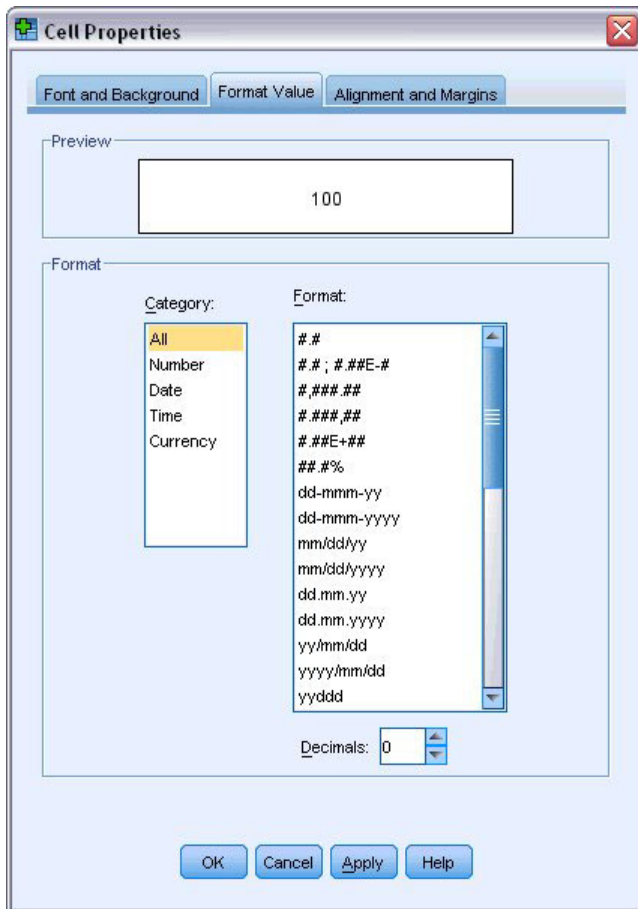


Figura 43. Pestaña Valor de formato del cuadro de diálogo Propiedades de casilla

También se puede modificar el tipo de datos y el formato en este cuadro de diálogo.

7. Seleccione el tipo que desee de la lista Categoría y a continuación, el formato para dicho tipo de la lista Formato.
8. Pulse en **Aceptar** o en **Aplicar** para aplicar los cambios.

		Frecuencia	Porcentaje	Porcentaje acumulado
Válidos	No completó el bachillerato	1390	22	21,7
	Bachillerato	1936	30	52,0
	Universitarios parciales	1360	21	73,2
	Universitarios	1355	21	94,4
	Post-universitarios	359	6	100,0
	Total	6400	100	

Figura 44. Decimales ocultos en la columna Porcentaje

Los decimales quedarán ocultos en la columna *Porcentaje*.

TableLook

El formato de las tablas es una parte muy importante a la hora de proporcionar resultados claros, concisos y con significado. Si la tabla es difícil de leer, es posible que la información contenida en dicha tabla no se entienda fácilmente.

Utilización de formatos predefinidos

1. Pulse dos veces en la tabla *Estado civil*.
2. Elija en los menús:

Formato > Aspectos de tabla...

El cuadro de diálogo Aspectos de la tabla recoge una variedad de estilos predefinidos. Seleccione un estilo de la lista para obtener una vista previa en la ventana Muestra situada a la derecha.

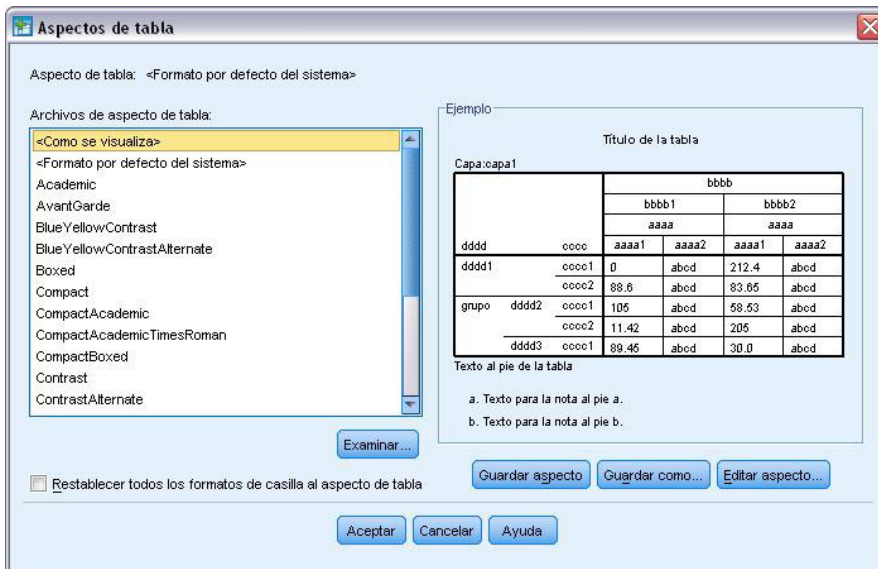


Figura 45. Cuadro de diálogo TableLooks

Puede utilizar un estilo tal cual o puede editar un estilo existente para que se ajuste mejor a sus necesidades.

3. Para utilizar un estilo existente, seleccione uno y pulse en **Aceptar**.

Personalización de estilos de aspectos de tabla

Puede personalizar un formato para que se ajuste a sus necesidades específicas. Se pueden personalizar casi todos los aspectos de una tabla, desde el color de fondo hasta los estilos de los bordes.

1. Pulse dos veces en la tabla *Estado civil*.
2. Elija en los menús:
Formato > Aspectos de tabla...
3. Seleccione el estilo que más se aproxime al formato que quiere y pulse en **Editar aspecto**.
4. Pulse en la pestaña **Formatos de casilla** para ver las opciones de formato.

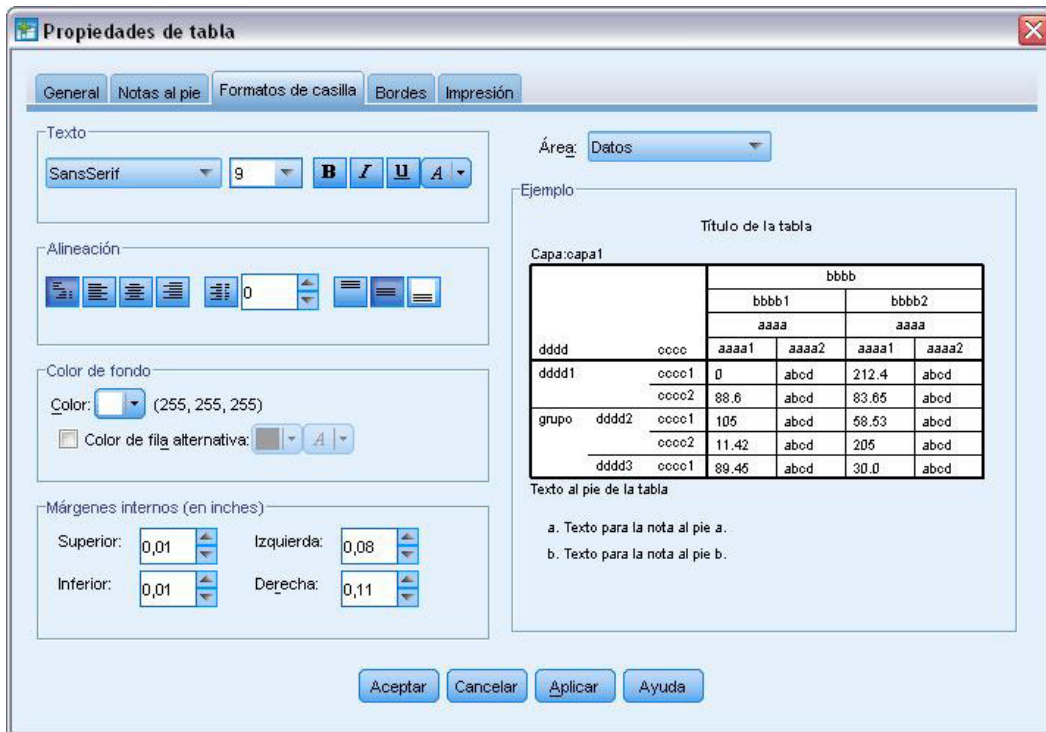


Figura 46. Cuadro de diálogo Propiedades de tabla

Entre las opciones de formato se incluyen el nombre de la fuente, el tamaño de la fuente, el estilo y el color. Entre las opciones adicionales se incluyen la alineación, el color del texto y del fondo, y los tamaños de los márgenes.

La ventana Muestra situada a la derecha proporciona una vista previa de cómo afectan a la tabla los cambios de formato. Cada área de la tabla puede tener un estilo de formato diferente. Por ejemplo, probablemente no querrá que el título tenga el mismo formato que los datos. Para seleccionar un área de la tabla para editar, puede seleccionar el área por el nombre en la lista desplegable Área o pulsar en el área que desee cambiar en la ventana Muestra.

5. Seleccione **Datos** de la lista desplegable Área.
6. Seleccione un color nuevo de la paleta desplegable Fondo.
7. A continuación, seleccione un nuevo color de texto.

La ventana Muestra ofrece el nuevo estilo mostrar.

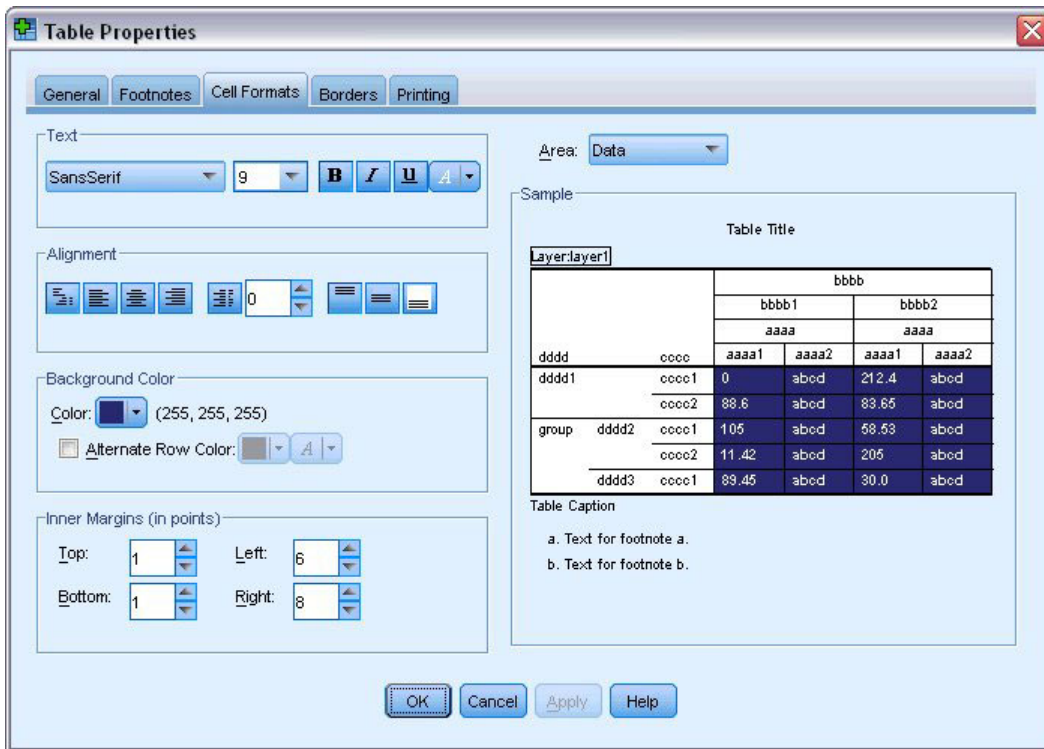


Figura 47. Cambio de los formatos de casilla de la tabla

8. Pulse en **Aceptar** para volver al cuadro de diálogo Aspectos de tabla.
El nuevo estilo se puede guardar, lo que facilita su aplicación en futuras tablas.
9. Pulse en **Guardar como**.
10. Acceda al directorio de destino e introduzca un nombre para el nuevo estilo en el cuadro de diálogo Nombre de archivo.
11. Pulse en **Guardar**.
12. Pulse en **Aceptar** para aplicar los cambios y volver al Visor.

La tabla contendrá el formato personalizado especificado.

The screenshot shows a PivotTable titled "Marital status" within a window named "viewertut.spv". The table displays the following data:

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Unmarried	3224	50.4	50.4	50.4
Married	3176	49.6	49.6	100.0
Total	6400	100.0	100.0	

Figura 48. Aspecto de tabla personalizado

Cambio de los formatos de tablas predeterminadas

Aunque se puede modificar el formato de una tabla tras su creación, puede ser más eficaz cambiar el aspecto de tabla predeterminada de manera que no se tenga que cambiar el formato cada vez que se cree una tabla.

Para cambiar el estilo del aspecto de tabla predeterminada para las tablas pivote, elija en los menús:

Editar > Opciones...

1. Pulse en la pestaña **Tablas pivote** en el cuadro de diálogo Opciones.

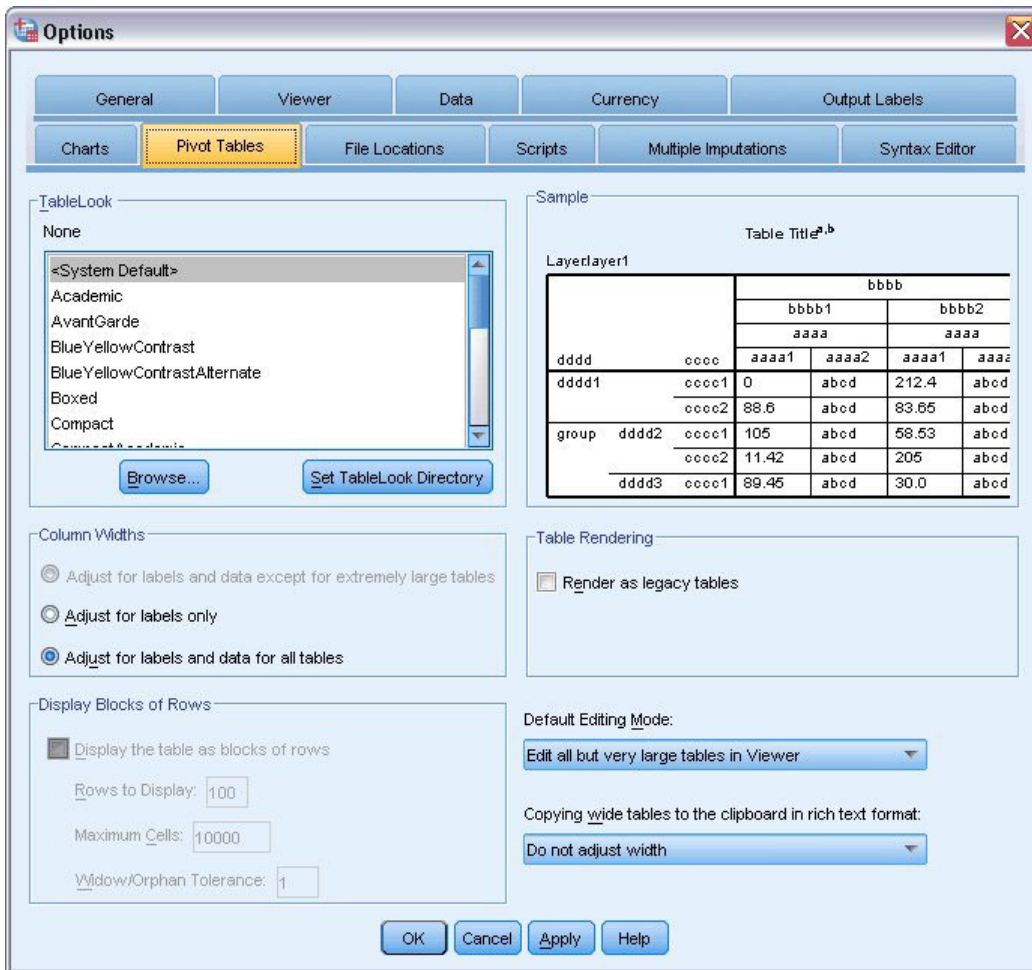


Figura 49. Cuadro de diálogo Opciones

2. Seleccione el estilo de aspecto de tabla que desea utilizar para todas las tablas nuevas.
La ventana Muestra situada a la derecha muestra una vista previa de cada aspecto de tabla.
3. Pulse en **Aceptar** para guardar la configuración y cerrar el cuadro de diálogo.

Todas las tablas que se creen después de la modificación del aspecto de tabla predeterminada cumplirán de forma automática las nuevas reglas de formato.

Personalización de las opciones de presentación inicial

Las opciones de presentación inicial incluyen la alineación de objetos en el Visor, si de forma predeterminada se muestran u ocultan los objetos y la anchura de la ventana del Visor. Para cambiar estas opciones:

1. Elija en los menús:
Editar > Opciones...
2. Pulse en la pestaña **Visor**.

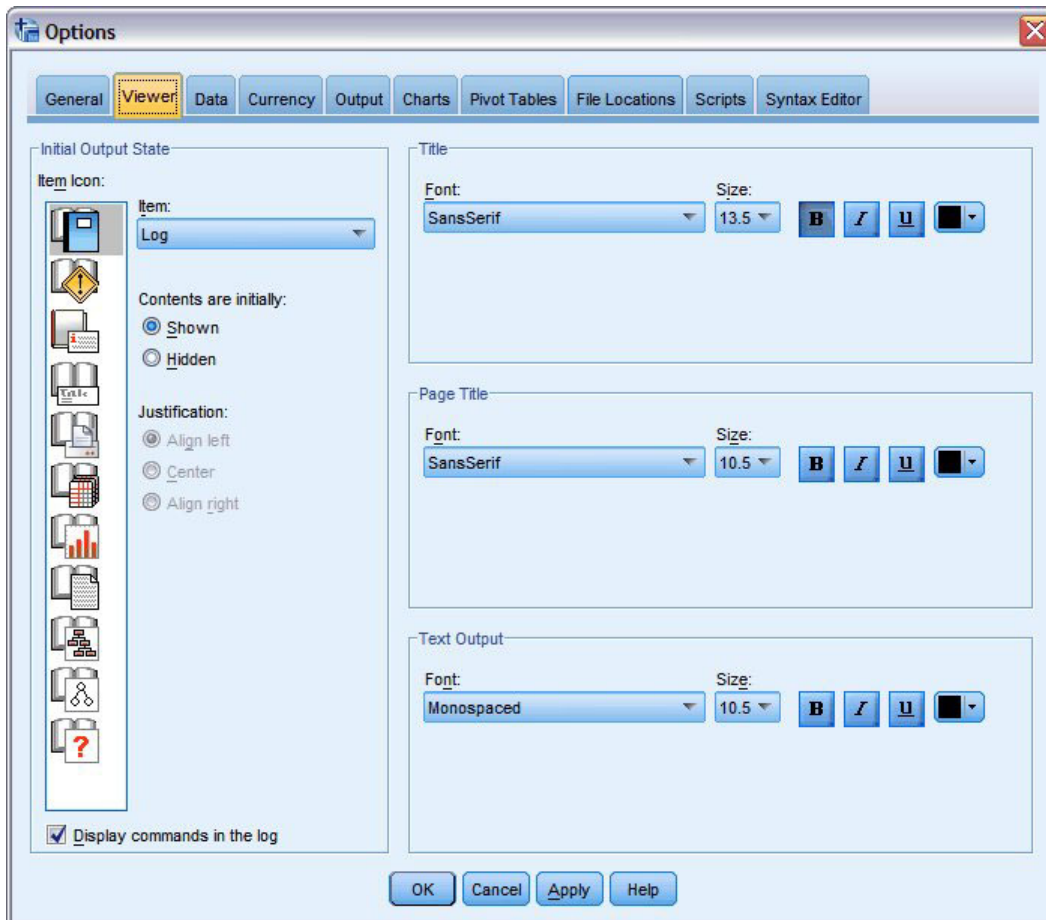


Figura 50. Opciones del Visor

Las opciones se aplican objeto por objeto. Por ejemplo, puede personalizar cómo se muestran los gráficos sin cambiar la forma de presentación de las tablas. Sólo tiene que seleccionar el objeto que desea personalizar y realizar los cambios.

3. Pulse en el icono **Título** para mostrar sus selecciones.
4. Pulse en **Centrar** para mostrar todos los títulos en el centro (horizontal) del Visor.

También se pueden ocultar elementos, como el registro y los mensajes de advertencia, que normalmente tienden a colapsar los resultados. Al pulsar dos veces en un icono, se cambia de forma automática la propiedad de presentación de dicho objeto.

5. Pulse dos veces en el icono **Advertencias** para ocultar los mensajes de advertencias de los resultados.
6. Pulse en **Aceptar** para guardar los cambios y cerrar el cuadro de diálogo.

Mostrar las etiquetas de variable y de valor

En la mayoría de los casos la presentación de etiquetas de variables y valores es más efectiva que mostrar el nombre de la variable o el valor actual de datos. Sin embargo, pueden darse casos en los que desee mostrar tanto los nombres como las etiquetas.

1. Elija en los menús:
Editar > Opciones...
2. Pulse en la pestaña **Etiquetas de los resultados**.

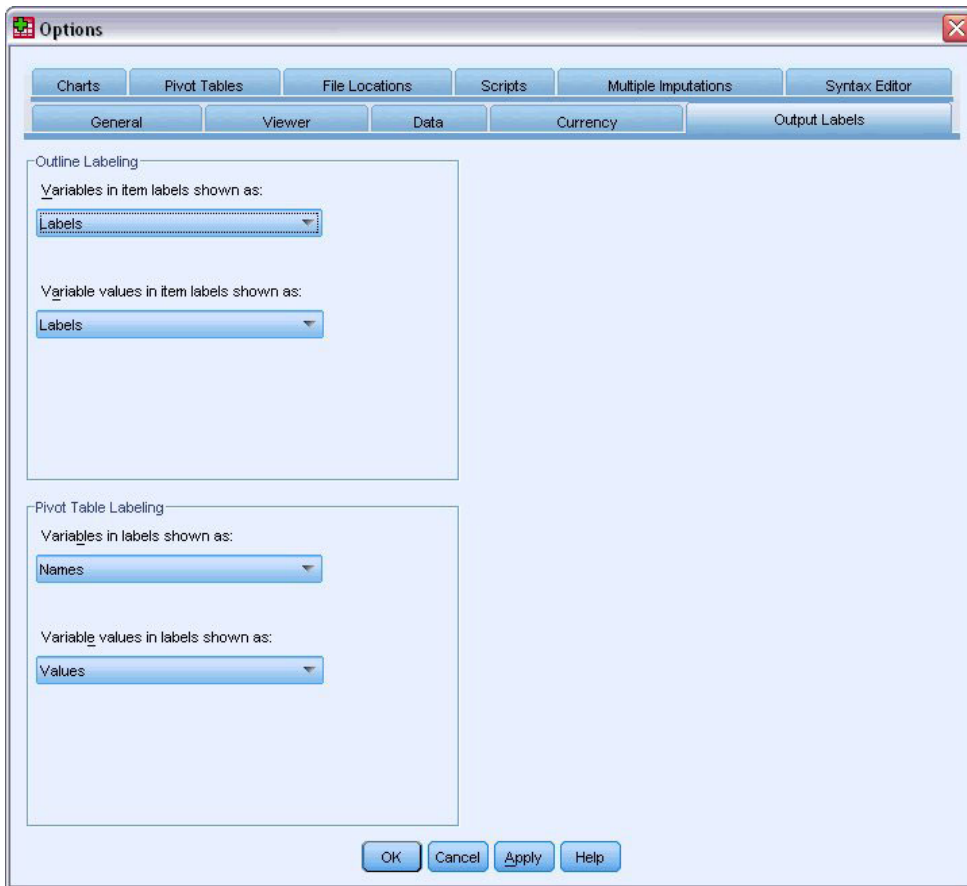


Figura 51. Selecciones de etiquetado de tablas pivote

Puede especificar distintas selecciones para los paneles de contenidos y titulares. Por ejemplo, para mostrar etiquetas en el titular y nombres de variable y valores de datos en los contenidos:

3. En el grupo Etiquetado de tablas pivote, seleccione **Nombres** de la lista desplegable de variables en etiquetas para mostrar los nombres de variables en vez de las etiquetas.
4. A continuación seleccione **Valores** de la lista desplegable Los valores de las variables se muestran en las etiquetas como para mostrar los valores de datos en lugar de las etiquetas.

Las siguientes tablas generadas en la sesión reflejarán estos cambios.

		marital			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3224	50.4	50.4	50.4
	1	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

Figura 52. Nombres y valores de variables mostrados

Uso de los resultados en otras aplicaciones

Los resultados se pueden utilizar en muchas aplicaciones. Por ejemplo, es posible que desee incluir una tabla o un diagrama en una presentación o un informe.

Los siguientes ejemplos son específicos para Microsoft Word, pero funcionarán de forma similar en otras aplicaciones de procesador de textos.

Pegado de resultados como tablas de Word

Las tablas pivote se pueden pegar en Word como si fuesen tablas nativas de Word. Se mantendrán todos los atributos de las tablas, como colores y tamaños de fuentes. Como la tabla se pegará con el formato de tabla de Word, podrá editarla como cualquier otra tabla.

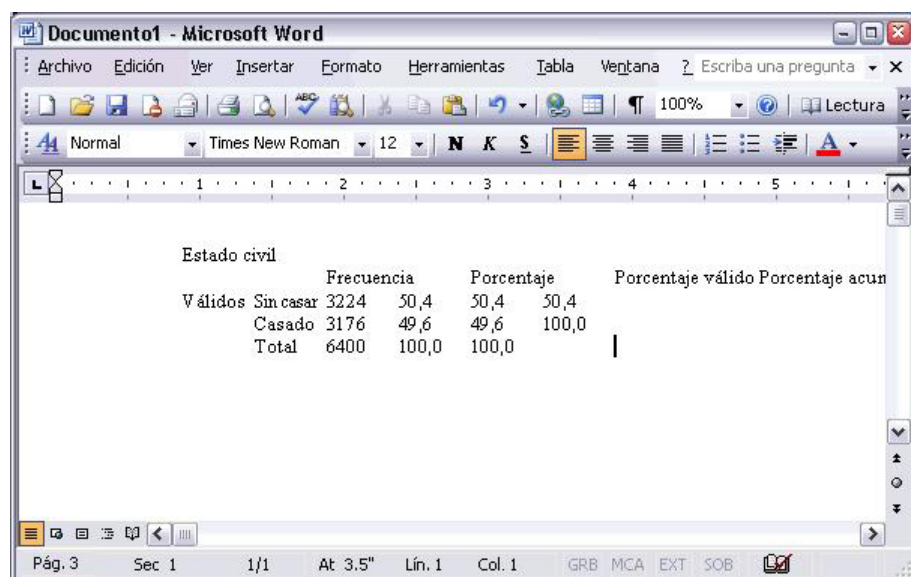
1. Pulse en una tabla del Visor para activarla.
2. Elija en los menús:
Editar > Copiar
3. Abra la aplicación de procesador de texto.
4. Desde el menú del procesador de texto, elija:
Editar > Pegar Especial...
5. Seleccione **Texto enriquecido (RTF)** en el cuadro de diálogo Pegado especial.
6. Pulse en **Aceptar** para pegar los resultados en el documento actual.

La tabla aparecerá en el documento. Puede aplicar un formato personalizado, edita los datos y cambiar el tamaño de la tabla para que se ajuste a sus necesidades.

Pegado de resultados como texto

Las tablas pivote se pueden copiar a otras aplicaciones como texto sin formato. Con este método no se conservarán los estilos de formato, pero se pueden editar los datos pegados en la tabla en la aplicación de destino.

1. Pulse en una tabla del Visor para activarla.
2. Elija en los menús:
Editar > Copiar
3. Abra la aplicación de procesador de texto.
4. Desde el menú del procesador de texto, elija:
Editar > Pegar Especial...
5. Seleccione **texto sin formato** en el cuadro de diálogo Pegado especial.
6. Pulse en **Aceptar** para pegar los resultados en el documento actual.



Estado civil		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos Sin casar		3224	50,4	50,4	50,4
Casado		3176	49,6	49,6	100,0
Total		6400	100,0	100,0	

Figura 53. Tabla pivote mostrada en Word

Cada columna de la tabla aparece separada por tabulaciones. Puede cambiar la anchura de las columnas ajustando las tabulaciones en la aplicación de procesador de textos.

Exportación de resultados a archivos de Microsoft Word, PowerPoint y Excel

Puede exportar resultados a un archivo de Microsoft Word, PowerPoint o Excel. Puede exportar los elementos seleccionados o todos los elementos en el Visor. Esta sección utiliza los archivos *msouttut.sps* y *demo.sav*. Consulte el tema Capítulo 10, "Archivos muestrales", en la página 81 para obtener más información.

Nota: la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

En el panel de titulares del Visor puede seleccionar objetos específicos que quiere exportar, exportar todos los objetos o todos los elementos visibles.

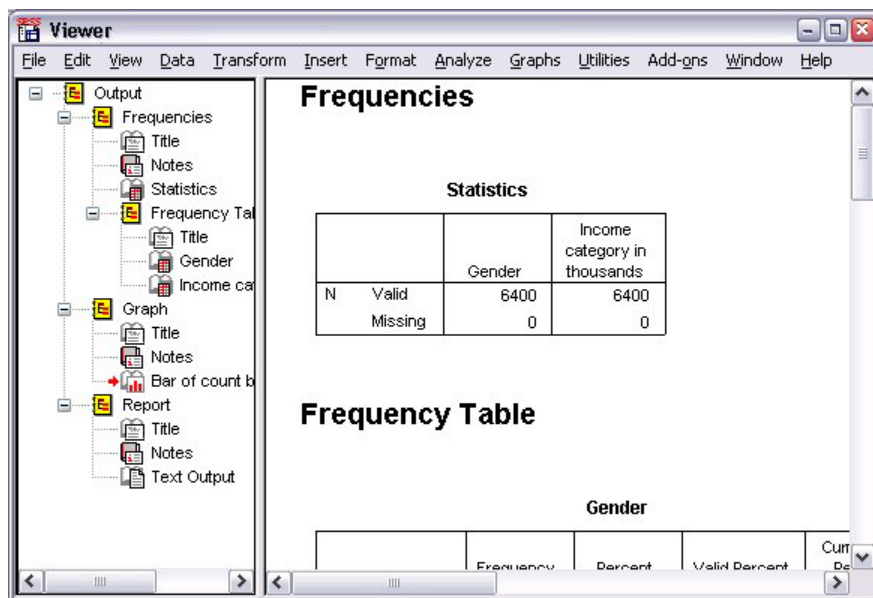


Figura 54. Visor

1. Seleccione en los menús del Visor:

Archivo > Exportar...

En lugar de exportar todos los objetos del Visor, puede optar por exportar únicamente objetos visibles (libros abiertos en el panel de titulares) o sólo los seleccionados en el panel de titulares. Si no seleccionó ningún elemento del panel de titulares, no dispondrá de la opción de exportar objetos seleccionados.

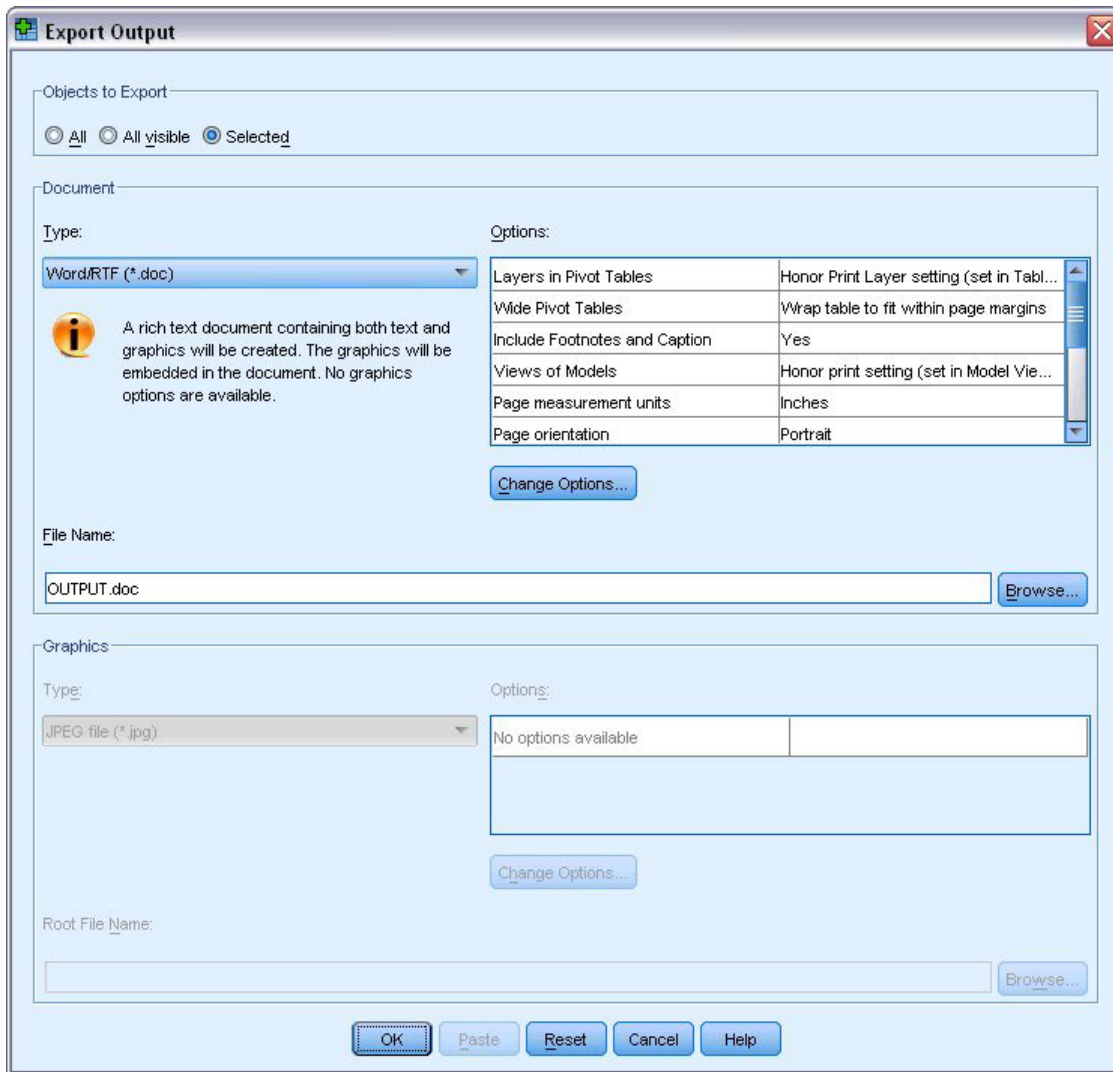


Figura 55. Cuadro de diálogo Exportar resultados

2. En el grupo Objetos para exportar, seleccione **Todos**.
3. Seleccione **Archivo Word/RTF (*.doc)** de la lista desplegable Tipo.
4. Pulse **Aceptar** para generar el archivo de Word.

Al abrir el archivo resultante en Word, podrá observar cómo se exportan los archivos. Las notas, que no son objetos visibles, aparecen en Word por haberse seleccionado la exportación de todos los objetos.

Las tablas pivote pasan a ser tablas de Word, conservando todo el formato original de la tabla pivote inicial, incluyendo fuentes, colores, bordes, etcétera.

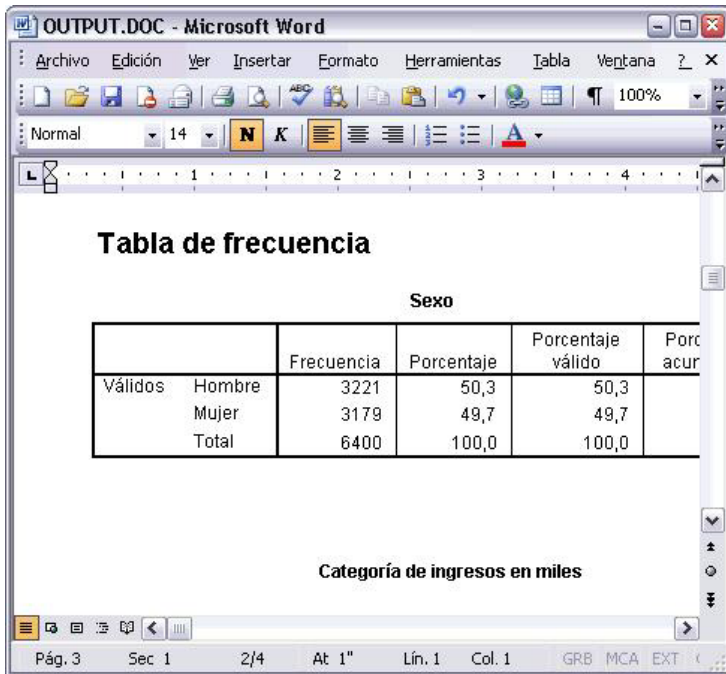


Figura 56. Tablas pivote en Word

Los gráficos se incluyen en el documento de Word como imágenes de gráficos.

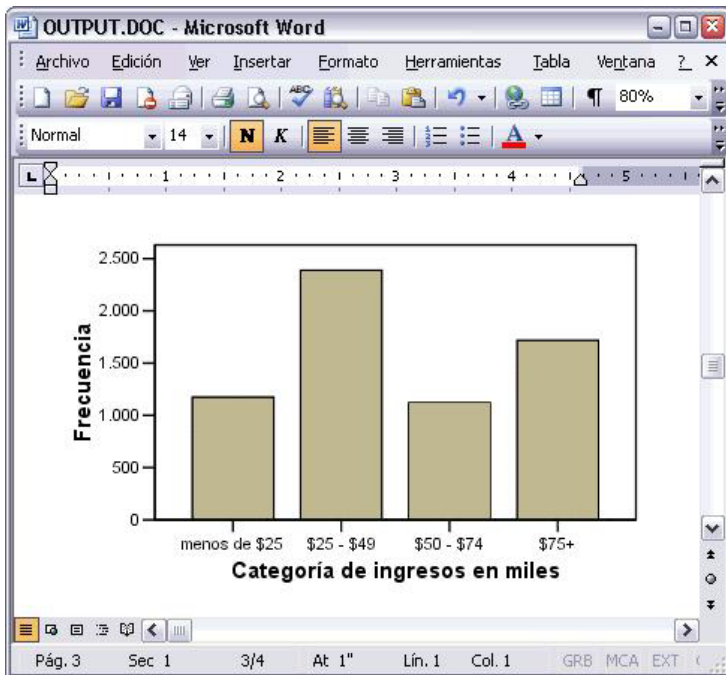


Figura 57. Gráficos en Word

Los resultados de texto aparecen en la misma fuente utilizada para el objeto de texto en el Visor. Para que aparezcan correctamente alineados, los resultados de texto deben utilizar una fuente de paso fijo (monoespaciada).

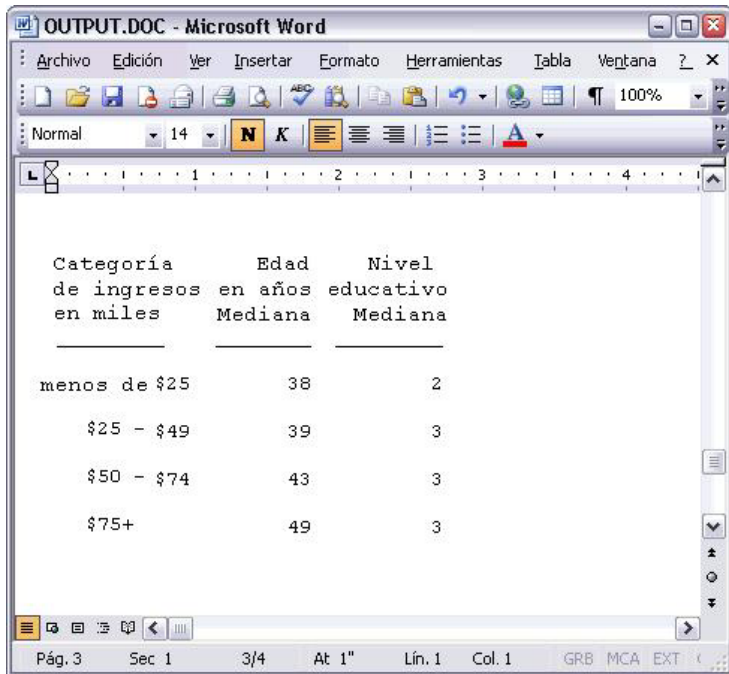


Figura 58. Resultados de texto en Word.

Si exporta a un archivo de PowerPoint, cada elemento exportado se colocará en una diapositiva diferente. Las tablas pivote pasan a ser tablas de PowerPoint, conservando el formato original de la tabla pivote inicial, incluyendo fuentes, colores, bordes, etcétera.

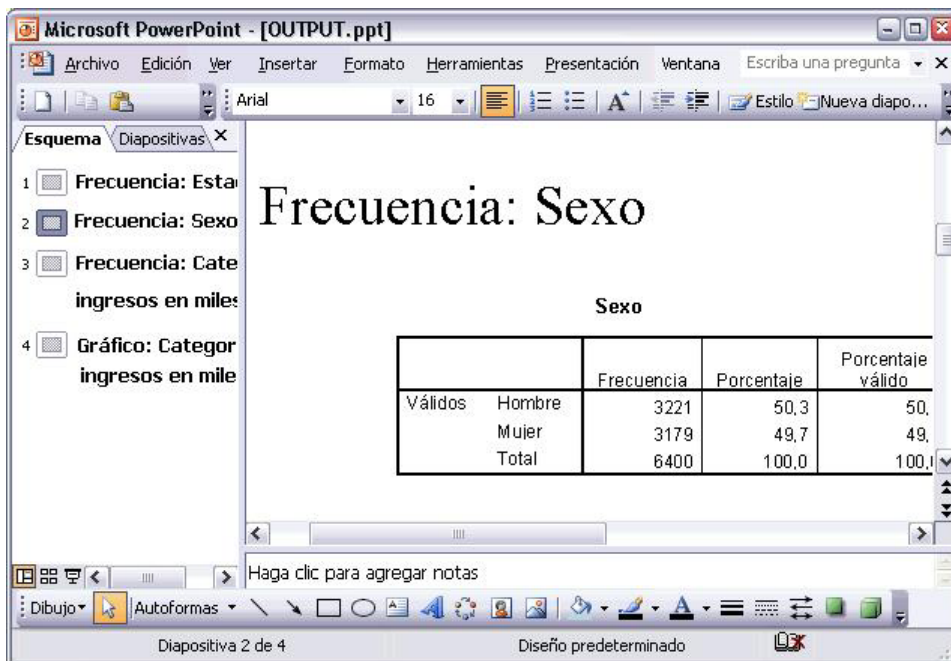


Figura 59. Tablas pivote en PowerPoint

Los gráficos seleccionados para su exportación a PowerPoint se incrustan en el archivo de PowerPoint.

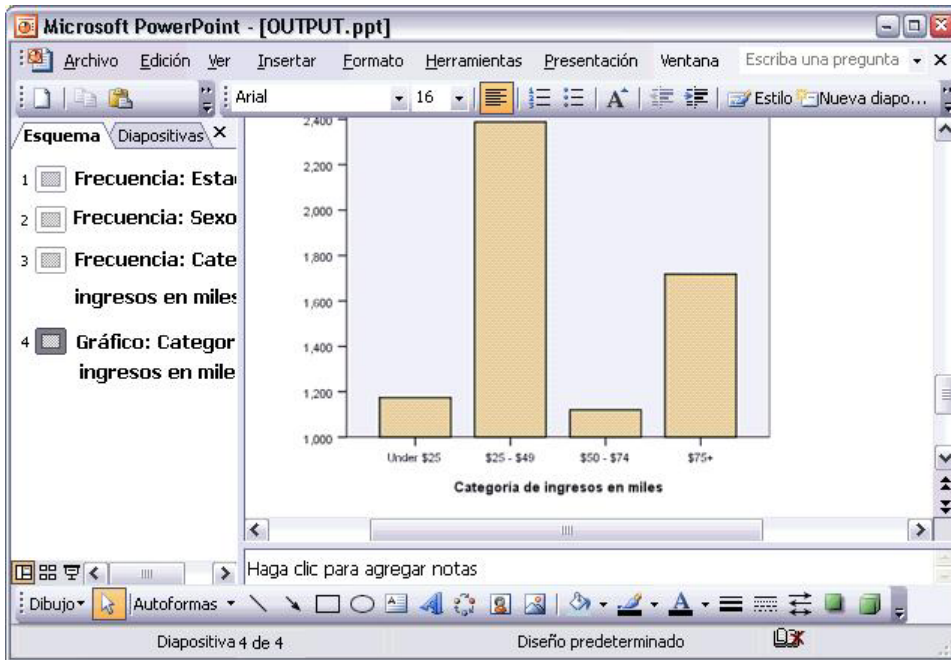


Figura 60. Gráficos en PowerPoint

Nota: la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

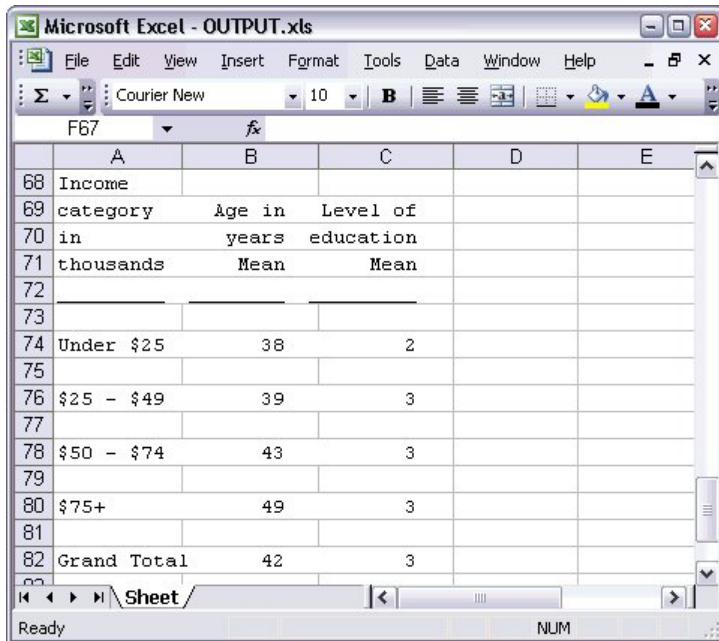
Si desea exportar a un archivo Excel, los resultados de la exportación serán distintos.

Gender					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	3,179	49.7	49.7	49.7
	Male	3,221	50.3	50.3	100.0
	Total	6,400	100.0	100.0	
Income category in thousands					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under \$25	1,174	18.3	18.3	18.3
	\$25 - \$49	2,388	37.3	37.3	55.7
	\$50 - \$74	1,120	17.5	17.5	73.2
	\$75+	1,718	26.8	26.8	100.0
	Total	6,400	100.0	100.0	

Figura 61. Output.xls en Excel

Las filas, columnas y casillas de las tablas pivote se convierten en filas, columnas y casillas de Excel.

Cada línea del resultado de texto constituye una fila del archivo de Excel y se incluye todo su contenido en una sola casilla.



	A	B	C	D	E
68	Income				
69	category	Age in	Level of		
70	in	years	education		
71	thousands	Mean	Mean		
72					
73					
74	Under \$25	38	2		
75					
76	\$25 - \$49	39	3		
77					
78	\$50 - \$74	43	3		
79					
80	\$75+	49	3		
81					
82	Grand Total	42	3		

Figura 62. Resultados de texto en Excel

Exportación de resultados a PDF

Puede exportar los elementos seleccionados o todos los elementos en el Visor a un archivo PDF (formato de documento portátil).

1. Elija en los menús de la ventana Visor que contiene el resultado que desea exportar a PDF:
Archivo > Exportar...
2. En el cuadro de diálogo Exportar resultados, en la lista desplegable Formato de exportación, Tipo de archivo elija **Formato de documento portátil**.

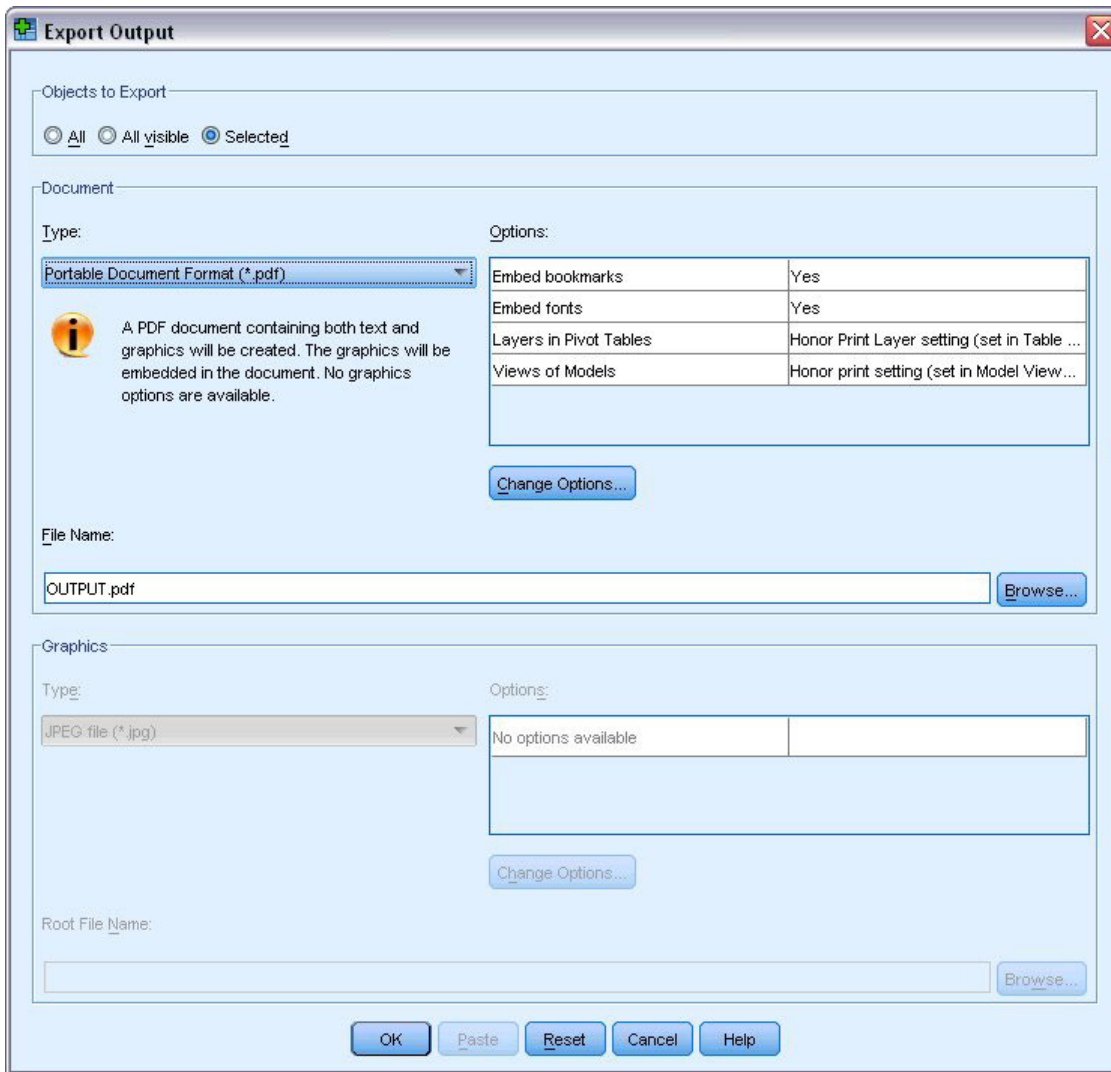


Figura 63. Cuadro de diálogo Exportar resultados

- El panel de titulares del documento del Visor se convertirá en marcadores en el archivo PDF para que sea más fácil desplazarse por él.
- El tamaño de la página, la orientación, los márgenes, el contenido y la presentación de los encabezados y pies de página, así como el tamaño del gráfico impreso en los documentos PDF está controlado por las opciones de preparación de página (menú Archivo, Preparar página en la ventana Visor).
- La resolución (PPP) del documento PDF es la configuración de la resolución actual para la impresora predeterminada o que esté seleccionada en ese momento (y que puede cambiarse mediante Preparar página). La resolución máxima es de 1200 PPP. Si el valor de la impresora es superior, la resolución del documento PDF será de 1200 PPP. *Nota:* los documentos de alta resolución pueden generar resultados de baja calidad si se imprimen en impresoras con menor resolución.

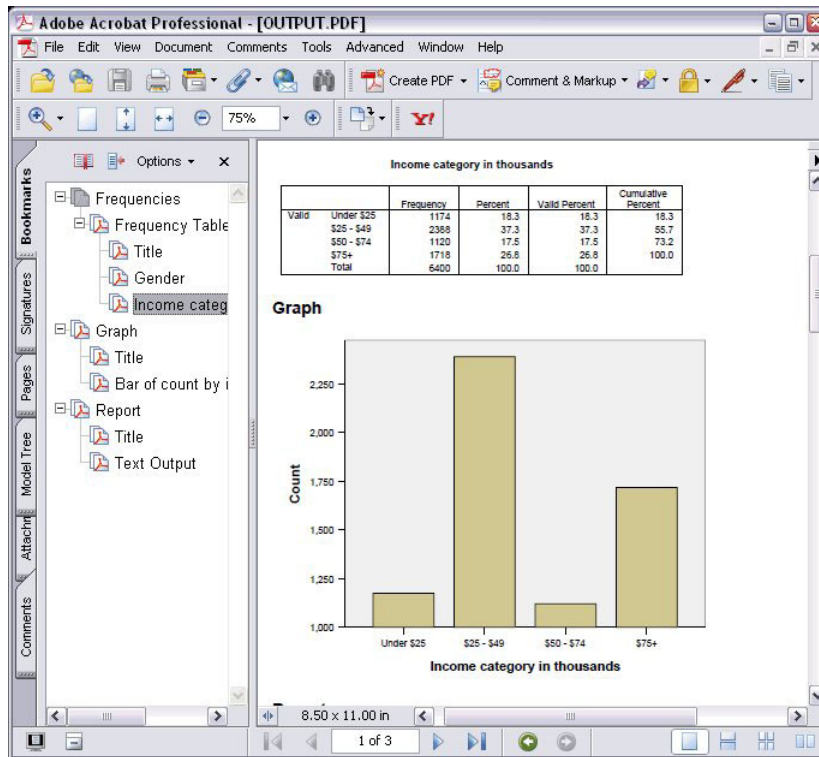


Figura 64. Archivo PDF con marcadores

Exportación de resultados a HTML

Puede exportar también los resultados a HTML (lenguaje de marcas de hipertexto). Al guardar como HTML, todos los resultados que no sean gráficos se exportarán a un único archivo HTML.

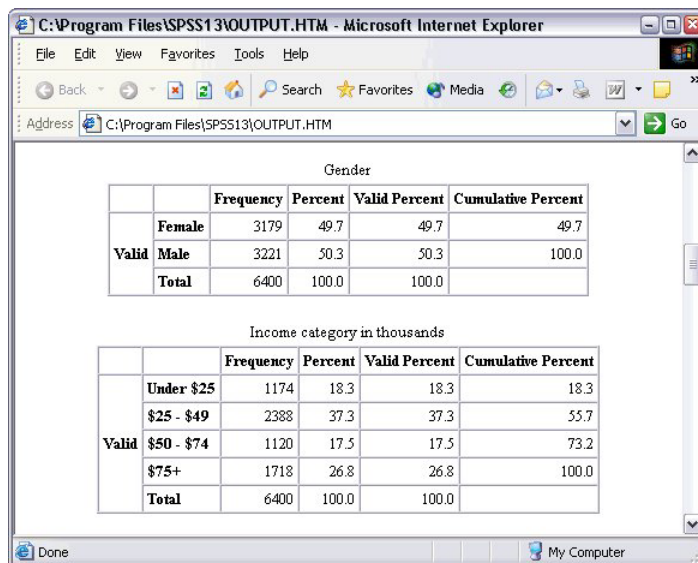


Figura 65. Archivo Output.htm en navegador Web

Al exportar a HTML, es posible también exportar gráficos, pero no a un único archivo.

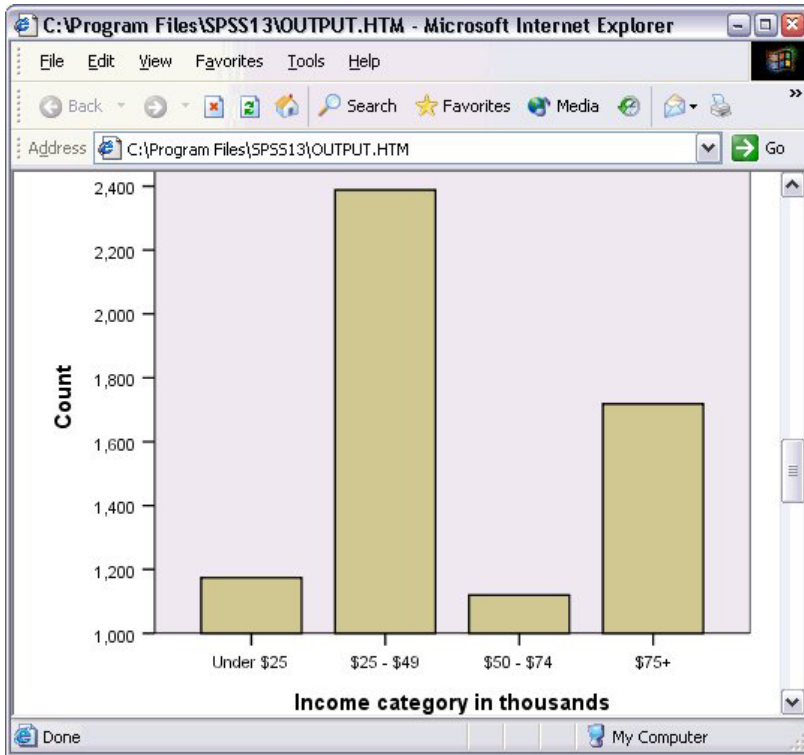


Figura 66. Gráfico en HTML

Cada gráfico se guardará como un archivo en un formato especificado por el usuario y las referencias a dichos archivos gráficos se incluirán en el documento HTML. Existe también una opción para exportar todos los gráficos (o los gráficos seleccionados) a archivos gráficos independientes.

Capítulo 7. Trabajo con sintaxis

Puede ahorrar y automatizar muchas tareas comunes mediante el eficaz lenguaje de comandos. 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.

Un archivo de sintaxis de comandos es simplemente un archivo de texto que contiene comandos de IBM SPSS Statistics. Puede abrir una ventana de sintaxis y escribir directamente los comandos, pero con frecuencia es más sencillo permitir que los cuadros de diálogo se ocupen automáticamente de parte o todo este trabajo.

En los ejemplos de este capítulo se utiliza el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Nota: en la versión para estudiantes no está disponible la sintaxis de comandos.

Pegado de la sintaxis

La forma más sencilla de crear sintaxis consiste en utilizar el botón Pegar que se encuentra en la mayoría de los cuadros de diálogo.

1. Abra el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.
2. Elija en los menús:
Analizar > Estadísticos descriptivos > Frecuencias...
3. Seleccione *Estado civil [ecivil]* y muévela hasta la lista Variable(s).
4. Pulse en **Gráficos**.
5. En el cuadro de diálogo Gráficos, seleccione **Gráficos de barras**.
6. En el grupo Valores del gráfico, seleccione **Porcentajes**.
7. Pulse **Continuar**. Pulse **Pegar** para copiar la sintaxis creada como resultado de las selecciones del cuadro de diálogo en el Editor de sintaxis.

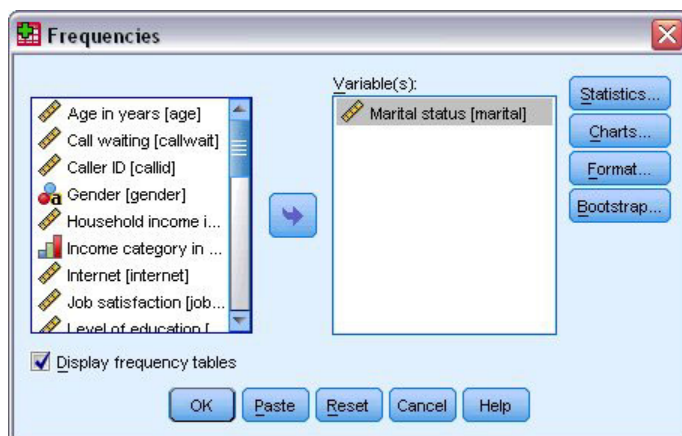


Figura 67. Cuadro de diálogo Frecuencias

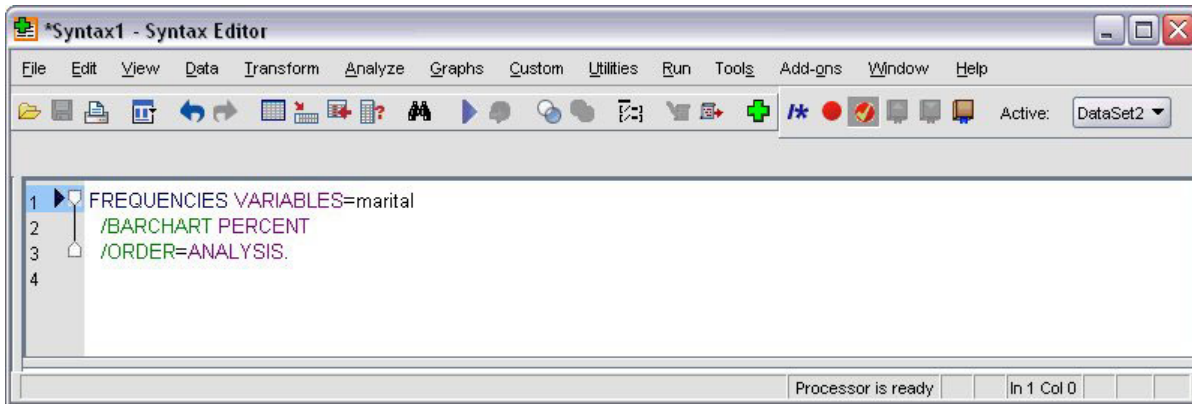


Figura 68. Sintaxis de frecuencias

8. Para ejecutar la sintaxis que se muestra en ese momento, elija en los menús:
Ejecutar > Selección

Edición de la sintaxis

En la ventana de sintaxis, puede editar la sintaxis. Por ejemplo, puede modificar el subcomando /BARCHART para que muestre frecuencias en lugar de porcentajes. (Cada subcomando se indica con una barra inclinada.) Si conoce la palabra clave para mostrar frecuencias, puede introducirlas directamente. Si no conoce la palabra clave, puede obtener una lista de palabras claves disponibles para el subcomando situando el cursor en cualquier parte siguiendo el nombre del subcomando y pulsando Ctrl+barra espaciadora. Muestra el control autocompletar para el subcomando.

Eliminar la palabra clave PERCENT del subcomando BARCHART.

Pulse Ctrl-Barra de espacio.

1. Pulse en el elemento etiquetado **FREC** para mostrar las frecuencias. Al pulsar en un elemento en el control autocompletar lo insertará en el lugar donde se encuentre el cursor actualmente.
De forma predeterminada, el control autocompletar mostrará una lista de términos disponibles conforme escribe. Por ejemplo, le gustaría incluir un gráfico circular en el gráfico de barras. El gráfico circular se especifica por medio de un subcomando independiente.
2. Pulse Intro tras la palabra clave **FREC** y escriba una barra inclinada para indicar el inicio de un subcomando.

El Editor de sintaxis muestra la lista de subcomandos para el comando actual.

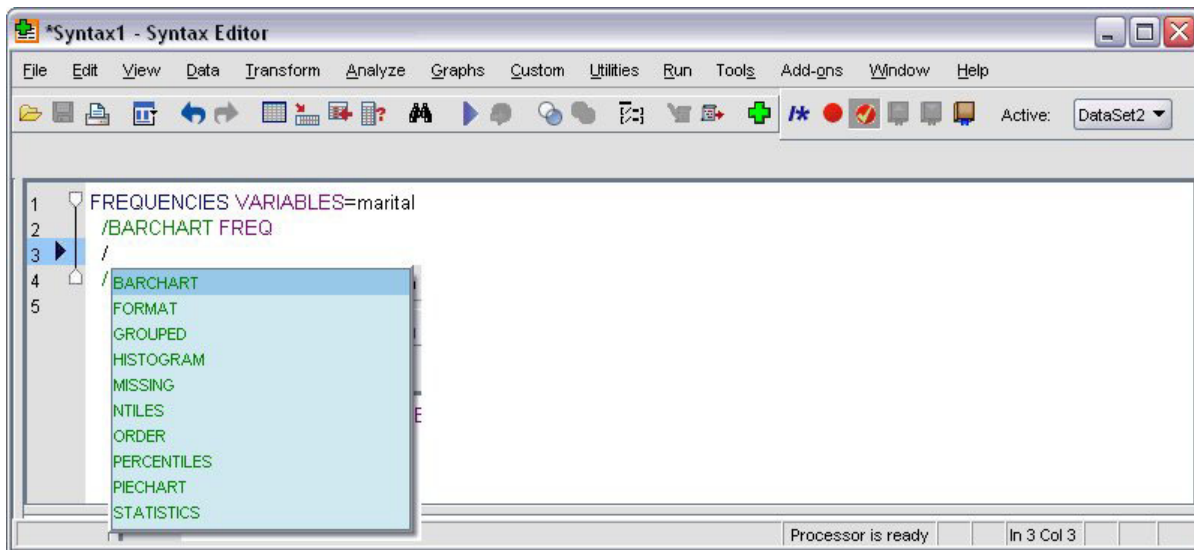


Figura 69. El control autocompletar muestra los subcomandos.

Para obtener más ayuda detallada acerca del comando actual, pulse la tecla F1. Al hacerlo, aparecerá directamente la información de referencia de sintaxis de comandos correspondiente al comando actual.

Puede que haya notado que el texto que aparece en la ventana de sintaxis tiene color. La codificación de colores le permite identificar rápidamente los términos no reconocidos, si sólo los términos reconocidos tienen color. Por ejemplo, usted escribe incorrectamente el subcomando `FORMAT0` como `FRMAT0`. Los subcomandos se colorean de verde de forma predeterminada, pero el texto `FRMAT0` aparecerá sin colorear si no se ha reconocido.

Apertura y ejecución de un archivo de sintaxis

1. Para abrir un archivo de sintaxis almacenado, elija en los menús:
Archivo > Abrir > Sintaxis...
Aparece un cuadro de diálogo estándar para abrir archivos.
2. Seleccione un archivo de sintaxis. Si no se muestra ningún archivo de sintaxis, compruebe que la opción **Sintaxis (*.sps)** esté seleccionada como el tipo de archivos que desea ver.
3. Pulse en **Abrir**.
4. Utilice el menú **Ejecutar** de la ventana de sintaxis para ejecutar los comandos.

Si los comandos se aplican a un archivo de datos específico, deberá abrir el archivo de datos antes de ejecutar los comandos o bien deberá incluir un comando que lo abra. Puede pegar este tipo de comando desde los cuadros de diálogo que abren archivos de datos.

Utilización de puntos de separación

Los puntos de separación le permiten detener la ejecución de la sintaxis de comando en puntos específicos de la sintaxis y continuar con la ejecución cuando esté listo. Esto le permite visualizar los resultados o datos en el punto intermedio durante un trabajo de sintaxis o ejecutar la sintaxis de comandos que muestra la información sobre el estado actual de los datos, como `FREQUENCIES`. Los puntos de separación sólo se pueden establecer en el nivel de un comando, no en líneas específicas dentro de un comando.

Para insertar un punto de interrupción en un comando:

1. Pulse en cualquier lugar de la región que aparece a la izquierda del texto asociado al comando.

El punto de interrupción se representa mediante un círculo rojo en la región que aparece a la izquierda del texto de comando, en la misma línea que el nombre del comando, independientemente de donde pulse.

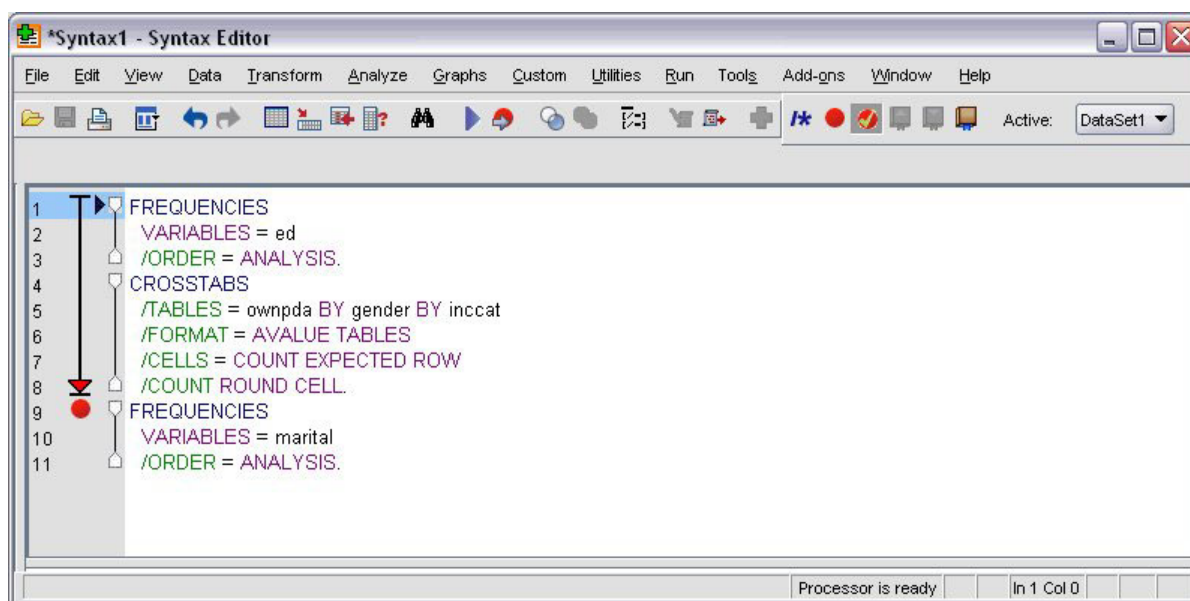


Figura 70. La ejecución se detiene en un punto de interrupción.

Cuando ejecute la sintaxis de comandos con puntos de interrupción, la ejecución se detiene antes de cada comando que contenga un punto de interrupción.

La flecha que apunta hacia abajo, a la izquierda del texto de comando muestra el progreso de la ejecución de la sintaxis. Abarca la región desde la primera ejecución de comando hasta la última y resulta especialmente útil cuando se ejecuta una sintaxis de comandos que contiene puntos de interrupción.

Para reanudar la ejecución que sigue a un punto de interrupción:

2. En los menús del Editor de sintaxis, elija:

Ejecutar > Continuar

Capítulo 8. Modificación de los valores de los datos

Puede que los datos con los que comience a trabajar no siempre estén organizados de la forma más conveniente para sus necesidades de análisis y de creación de informes. Por ejemplo, puede que desee:

- Crear una variable categórica a partir de una variable de escala.
- Combinar varias categorías de respuesta en una única categoría.
- Crear una nueva variable que sea la diferencia calculada entre dos variables existentes.
- Calcular la distancia temporal entre dos fechas.

En este capítulo se utiliza el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Creación de una variable categórica a partir de una variable de escala

Muchas de las variables categóricas del archivo de datos *demo.sav* se derivan, de hecho, de variables de escala de ese archivo de datos. Por ejemplo, la variable *catinc* consiste simplemente en la variable *ingres* agrupada en cuatro categorías. Esta variable categórica emplea los valores enteros 1–4 para representar las categorías de ingresos siguientes: menos que 25\$, entre 25\$ y 49\$, entre 50\$ y 74\$ y, 75\$ o superior.

Para crear la variable categórica *ingcat*:

1. Seleccione en los menús de la ventana Editor de datos:

Transformar > Agrupación visual...

En el cuadro de diálogo inicial Agrupación visual, puede seleccionar las variables de escala y ordinales para las cuales desee crear nuevas variables agrupadas en intervalos. La **agrupación** consiste simplemente en tomar dos o más valores contiguos y agruparlos en una misma categoría.

Debido a que la agrupación visual se basa en los valores reales del archivo de datos para ayudarle a realizar buenas elecciones de agrupación, antes necesita leer el archivo de datos. Puesto que dicha operación puede llevar algún tiempo si su archivo de datos contiene un gran número de casos, este cuadro de diálogo inicial le permite también limitar el número de casos que se debe leer (que se exploran). Este proceso no es preciso con nuestro archivo de datos de muestra. A pesar de que contiene más de 6.000 casos, no lleva demasiado tiempo el explorar esa cantidad de casos.

2. Arrastre y coloque *Ingresos del hogar en miles [ingres]* desde la lista Variables a la lista Variables para agrupar, y pulse en **Continuar**.

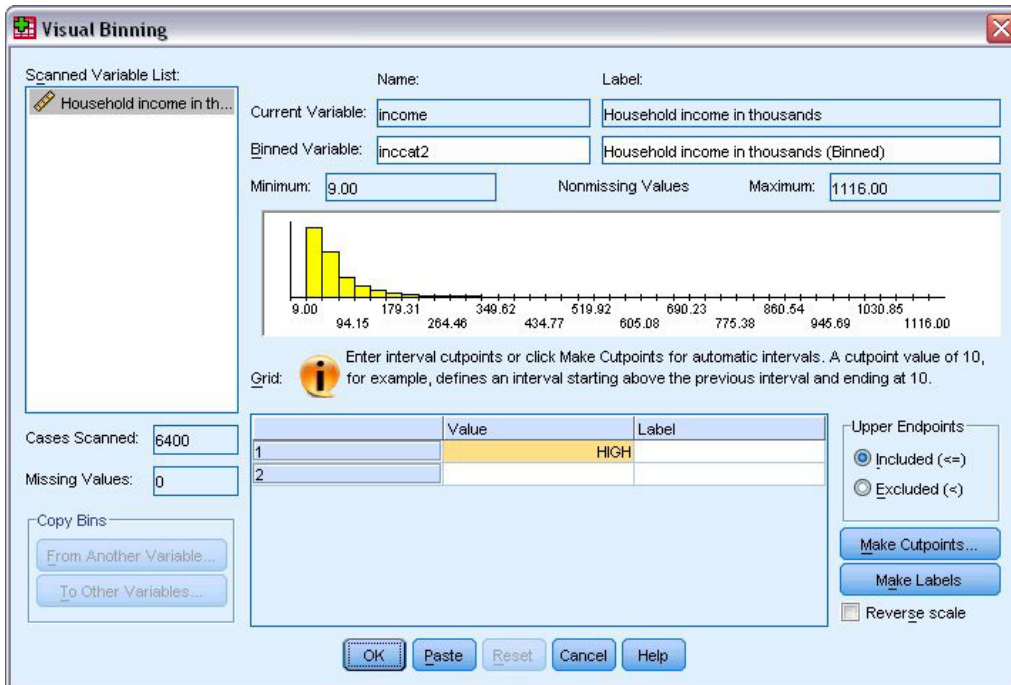


Figura 71. Cuadro de diálogo principal Agrupación visual

3. En el cuadro de diálogo principal Agrupación visual, seleccione *Ingresos del hogar en miles [ingres]* en la Lista de variables exploradas.

Un histograma muestra la distribución de la variable seleccionada (que, en este caso, es bastante asimétrica).

4. Escriba `ingcat2` como nuevo nombre para la variable agrupada y Categoría de ingresos [en miles] para la etiqueta de variable.
5. Pulse **Crear puntos de corte**.
6. Seleccione **Intervalos de igual amplitud**.
7. Introduzca 25 para el primer punto de corte, 3 para el número de puntos de corte y 25 para la amplitud.

El número de categorías agrupadas es superior al número de puntos de corte. Así que, en este ejemplo, la nueva variable agrupada contará con cuatro categorías (cada una de las tres primeras contendrá intervalos de 25 unidades [en miles] y la última, todos los valores por encima del de punto de corte mayor: 75 [en miles]).

8. Pulse en **Aplicar**.

Los valores que se muestran ahora en la rejilla, representan los puntos de corte definidos, que son los límites superiores de cada categoría. Las líneas verticales del histograma indican también la ubicación de los puntos de corte.

De forma predeterminada, los valores de dichos puntos de corte se incluyen en las categorías correspondientes. Por ejemplo, el primer valor 25 incluiría entonces todos los valores menores o iguales a 25. Pero, en este ejemplo, queremos contar con categorías que correspondan a cantidades menores que 25, 25–49, 50–74 y 75 o superiores.

9. En el grupo Límites superiores, seleccione **Excluidos (<)**.
10. A continuación pulse en **Crear etiquetas**.

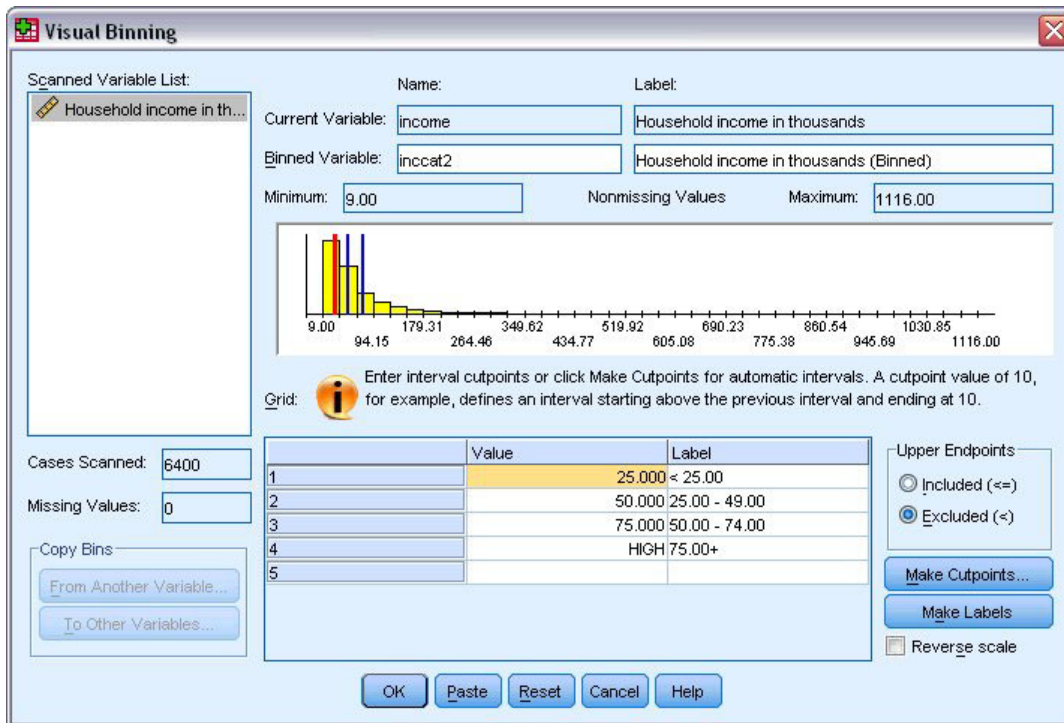


Figura 72. Etiquetas de valor generadas automáticamente.

Con ello se generan automáticamente etiquetas de valor descriptivas para cada categoría. Puesto que los valores reales asignados a la nueva variable agrupada son sólo números enteros consecutivos comenzando por el 1, las etiquetas de valor pueden resultar muy útiles.

También se pueden introducir o modificar manualmente los puntos de corte y las etiquetas en la rejilla, cambiar la ubicación de los puntos de corte arrastrando y colocando sus líneas correspondientes sobre el histograma y eliminar puntos de corte arrastrando sus líneas fuera del histograma.

11. Pulse en **Aceptar** para crear la nueva variable agrupada.

La nueva variable aparecerá en el Editor de datos. Puesto que la variable se añade al final de archivo, aparece en la columna que se encuentra más a la derecha en la Vista de datos y en la última fila de la Vista de variables.

Cálculo de nuevas variables

Es posible calcular nuevas variables basadas en ecuaciones muy complejas utilizando una amplia variedad de funciones matemáticas. En este ejemplo, sin embargo, sólo calcularemos una nueva variable que sea la diferencia entre los valores de dos variables existentes.

El archivo de datos *demo.sav* contiene una variable para la edad actual de los encuestados y otra variable para el número de años transcurridos en el trabajo actual. Sin embargo, no contiene una variable para la edad del encuestado en el momento que comenzó ese trabajo. Podemos crear una nueva variable que sea la diferencia calculada entre la edad actual y el número de años transcurridos en el trabajo actual, que deberá ser la edad aproximada a la que el encuestado comenzó ese trabajo.

1. Seleccione en los menús de la ventana Editor de datos:
Transformar > Calcular variable...
2. Para la Variable de destino introduzca `trabajo1`.
3. Seleccione *Edad en años [edad]* en la lista de variables de origen y pulse el botón de flecha para copiarla en la Expresión numérica.

- Pulse en el botón del símbolo menos (-) de la calculadora en el cuadro de diálogo (o la tecla menos del teclado).
- Seleccione *Años con empresa actual [empleo]* y pulse el botón de flecha para copiarla en la expresión.

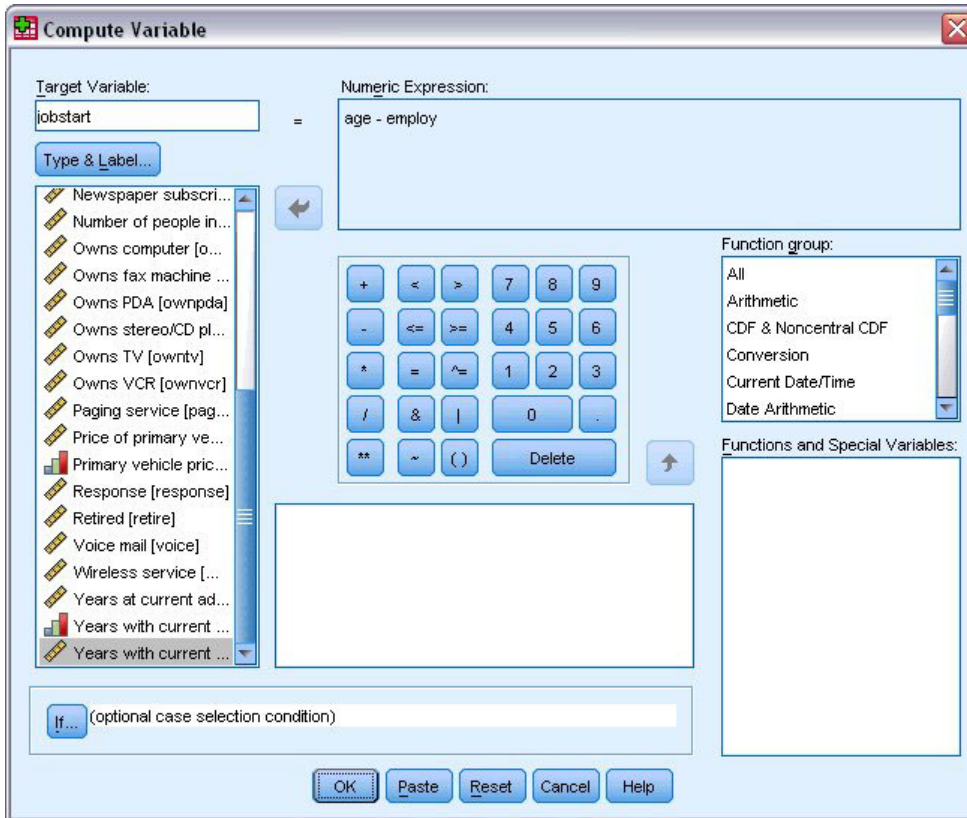


Figura 73. Cuadro de diálogo Calcular variable

Nota: esté atento al seleccionar la variable de empleo adecuada. Existe también una versión categórica recodificada de la variable, que *no* es la que usted desea. La expresión numérica debe ser *edad-empleo*, no *edad-empcat*.

- Pulse en **Aceptar** para calcular la nueva variable.

La nueva variable aparecerá en el Editor de datos. Puesto que la variable se añade al final de archivo, aparece en la columna que se encuentra más a la derecha en la Vista de datos y en la última fila de la Vista de variables.

Uso de funciones en las expresiones

Puede utilizar también funciones predefinidas en las expresiones. Hay más de 70 funciones preincorporadas:

- Funciones aritméticas
- Funciones estadísticas
- Funciones de distribución
- Funciones lógicas
- Funciones de agregación y extracción de fecha y hora
- Funciones de valores perdidos
- Funciones entre casos
- Funciones de cadena

Las funciones se organizan en grupos lógicos distintos (por ejemplo, un grupo para operaciones aritméticas y otro para calcular métricas estadísticas. Por comodidad, también se incluyen algunas variables de sistema de uso común (como *\$TIME*, para la fecha y hora actuales), organizadas en grupos de funciones.

Pegado de una función en una expresión

Para pegar una función en una expresión:

1. Coloque el cursor en el punto de la expresión en el que desea que aparezca la función.
2. Seleccione el grupo adecuado de la lista Grupo de funciones. El grupo con la etiqueta **Todo** contiene una lista de todas las funciones y variables de sistema disponibles.
3. Pulse dos veces en la función de las listas de funciones y variables especiales (o seleccione la función y pulse en la flecha que se encuentra sobre la lista Grupo de funciones).

Se insertará la función en la expresión. Si se resalta primero parte de la expresión y, a continuación, se inserta la función, se utilizará la parte resaltada de la expresión como primer argumento de la función.

Edición de una función de una expresión

La función no estará completa mientras que no se introduzcan los argumentos, que aparecen representados por interrogaciones en la función que se acaba de pegar. El número de interrogaciones indica el número mínimo de argumentos que son necesarios para completar la función.

1. Resalte las interrogaciones de la función pegada.
2. Introduzca los argumentos. Si los argumentos son nombres de variable, puede pegarlos desde la lista de variables.

Uso de Expresiones condicionales

Se pueden utilizar expresiones condicionales (también denominadas expresiones lógicas) para aplicar transformaciones a determinados subconjuntos de casos. Una expresión condicional devuelve el valor verdadero, falso o perdido para cada caso. Si el resultado de una expresión condicional es verdadero, la transformación se aplicará a ese caso. Si el resultado es falso o perdido, no se aplicará.

Para especificar una expresión condicional:

1. Pulse **Si** en el cuadro de diálogo Calcular variable. Se abrirá el cuadro de diálogo Si los casos.
2. Seleccione **Incluir si el caso satisface la condición**.
3. Introduzca la expresión condicional.

La mayoría de las expresiones condicionales contienen al menos un operador de relación como, por ejemplo, en:

`edad>=21`

o

`ingresos*3<100`

En el primer ejemplo, sólo aquellos casos que tengan un valor de 21 o mayor para la *Edad* [*edad*] seleccionada. En el segundo ejemplo, *Ingresos del hogar en miles* [*ingres*] multiplicado por 3 deberá ser inferior a 100 para que se pueda seleccionar un caso.

También se puede enlazar dos o más expresiones condicionales mediante operadores lógicos, como en:

`edad>=21 | ed>=4`

o

`ingresos*3<100 & educ=5`

En el primer ejemplo, se seleccionarán los casos que cumplan la condición de *Edad [edad]* o la condición de *Nivel educativo [educ]*. En el segundo ejemplo, las condiciones *Ingresos del hogar en miles [ingres]* y *Nivel educativo [educ]* deberán cumplirse ambas para que se pueda seleccionar un caso.

Trabajo con fechas y horas

Hay determinadas tareas que generalmente se llevan a cabo con fechas y horas que se pueden realizar fácilmente usando el Asistente para fecha y hora. Con este asistente, podrá:

- Crear una variable de fecha/hora a partir de una cadena que contiene una fecha o una hora.
- Crear una variable de fecha/hora fusionando variables que contengan partes diferentes de la fecha u hora.
- Añadir o quitar valores en variables de fecha/hora (lo que incluye la adición o sustracción de dos variables de fecha/hora).
- Extraer una parte de una variable de fecha/hora (por ejemplo, el día del mes de una variable de fecha/hora con el formato mm/dd/aaaa).

En los ejemplos de este apartado se utiliza el archivo de datos *upgrade.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Para usar el Asistente para fecha y hora:

1. Elija en los menús:

Transformar > Asistente para fecha y hora...



Figura 74. Pantalla principal del Asistente para fecha y hora

En la pantalla principal del Asistente para fecha y hora se presenta una serie de tareas generales. Las tareas no aplicables para los datos actuales se desactivan. Por ejemplo, el archivo de datos *upgrade.sav* no contiene variables de cadena, por lo que la tarea de creación de una variable de fecha a partir de una cadena se desactiva.

Si no está familiarizado con las fechas y las horas en IBM SPSS Statistics, puede seleccionar **Aprender cómo se representan las fechas y las horas** y pulsar en **Siguiente**. Se mostrará una pantalla con una breve descripción de las variables de fecha/hora y un enlace (a través del botón Ayuda) a información más detallada.

Cálculo de la distancia temporal entre dos fechas

Una de las tareas más comunes relacionadas con las fechas es el cálculo de la distancia temporal entre dos fechas. Como ejemplo, imagine que una compañía de software está interesada en analizar las compras de licencias actualizadas determinando el número de años que transcurren desde que cada cliente compró su última actualización. El archivo de datos *upgrade.sav* contiene una variable para la fecha en que cada cliente compró una actualización, pero no el número de años transcurridos desde entonces. Una nueva variable que represente la distancia temporal en años entre la fecha de la última actualización y la fecha de lanzamiento del nuevo producto proporciona la medida de esta cantidad.

Para calcular la distancia temporal entre dos fechas:

1. Seleccione **Realizar cálculos con fechas y horas** en la pantalla principal del Asistente para fecha y hora y pulse en **Siguiente**.
2. Seleccione **Calcular el número de unidades de tiempo entre dos fechas** y pulse en **Siguiente**.

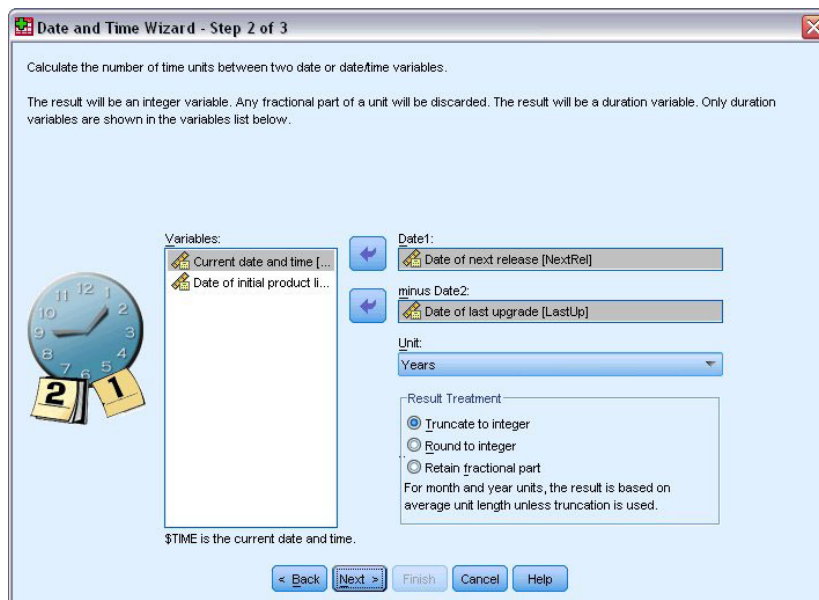


Figura 75. Cálculo de la distancia temporal entre dos fechas: paso 2

3. En el paso 2, seleccione *Fecha de la siguiente versión* en Fecha1.
4. Seleccione *Fecha de la última actualización* en Fecha2.
5. Seleccione **Años** para Unidad y **Truncar a entero** para Tratamiento resultante. (Estas selecciones se realizan de forma predeterminada.)
6. Pulse en **Siguiente**.
7. En paso 3, escriba *YearsLastUp* como nombre de la variable resultante. Las variables resultantes no puede tener el mismo nombre que una existente.

8. Escriba *Años desde última actualización* como etiqueta de la variable resultante. Las etiquetas de las variables resultantes son opcionales.
9. Mantenga las selecciones predeterminadas de **Crear la variable ahora** y pulse en **Finalizar** para crear la nueva variable.

La nueva variable, *AñosÚltimAct*, que se muestra en el Editor de datos, es el número entero de años transcurridos entre las dos fechas. Las partes fraccionales de un año se han omitido.

Adición de una duración a una fecha

Si lo desea, puede añadir o sustraer duraciones, como 10 días o 12 meses, a una fecha. Siguiendo con el ejemplo de la compañía de software del apartado anterior, imagine que se desea determinar la fecha en la que finaliza el contrato inicial de asistencia técnica de cada cliente. El archivo de datos *upgrade.sav* contiene una variable para el número de años de servicio de asistencia técnica contratado y otra para la fecha inicial de la compra. Así, es posible determinar la fecha final del servicio de asistencia inicial añadiendo años de asistencia a la fecha de compra.

Para añadir una duración a una fecha:

1. Seleccione **Realizar cálculos con fechas y horas** en la pantalla principal del Asistente para fecha y hora y pulse en **Siguiente**.
2. En el paso 1, seleccionar **Añadir o sustraer una duración a una fecha** y pulse en **Siguiente**.

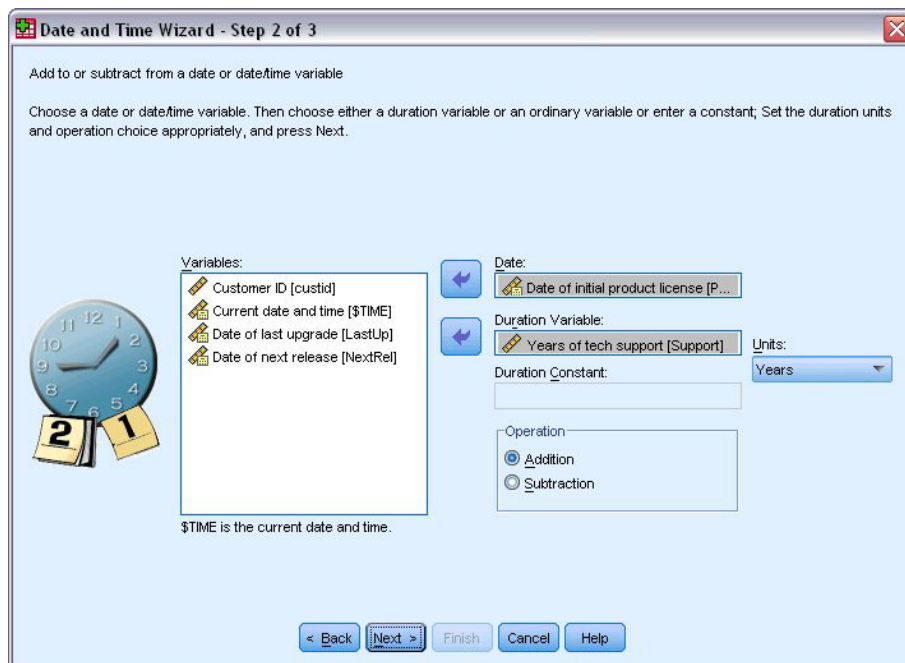


Figura 76. Adición de una duración a una fecha: paso 2

3. Seleccione *Fecha de la licencia de producto inicial* en Fecha.
4. En paso 2, seleccione *Años de asistencia técnica* como la variable de duración.
Como *Años de asistencia técnica* es una variable numérica simple, es preciso indicar las unidades que se deben usar al añadir esta variable como duración.
5. Seleccione **Años** en la lista desplegable Unidades.
6. Pulse en **Siguiente**.
7. En paso 3, escriba *SupEndDate* como nombre de la variable resultante. Las variables resultantes no puede tener el mismo nombre que una existente.

8. Escriba *Fecha de finalización de asistencia técnica* como etiqueta para la variable resultante. Las etiquetas de las variables resultantes son opcionales.
9. Pulse en **Finalizar** para crear la nueva variable.

La nueva variable aparecerá en el Editor de datos.

Capítulo 9. Ordenación y selección de datos

Los archivos de datos no siempre están organizados de la forma ideal para las necesidades específicas del usuario. Para preparar los datos para analizarlos posteriormente, puede transformar de muchas maneras diferentes el archivo, entre las que se incluyen:

- **Ordenar datos.** Puede ordenar los casos en función del valor de una o más variables.
- **Seleccionar subconjuntos de casos.** Puede restringir el análisis a un subconjunto de casos o efectuar análisis simultáneos de subconjuntos diferentes.

En los ejemplos de este capítulo se utiliza el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, “Archivos muestrales”, en la página 81 para obtener más información.

Ordenación de datos

Muchas veces resulta útil, y en ocasiones es incluso necesario, ordenar los casos (ordenar las filas del archivo de datos) para realizar determinados tipos de análisis.

Para cambiar el orden de la secuencia de casos en el archivo de datos de acuerdo con el valor de una o varias variables de ordenación:

1. Elija en los menús:

Datos > Ordenar casos...

Aparecerá el cuadro de diálogo Ordenar casos.

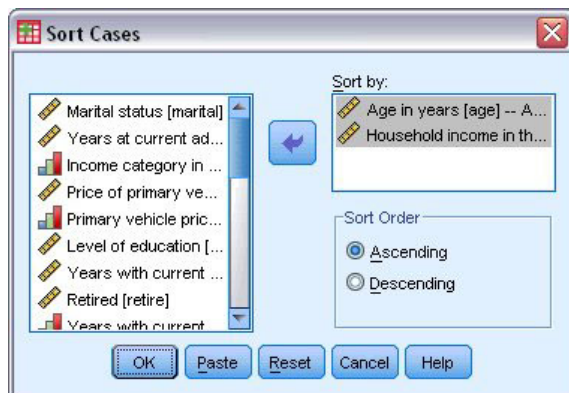


Figura 77. Cuadro de diálogo Ordenar casos

2. Añada las variables *Edad en años [edad]* e *Ingresos del hogar en miles [ingres]* a la lista Ordenar por.

Si selecciona varias variables de ordenación, los casos se ordenarán de acuerdo con el orden en que estas variables aparezcan en la lista Ordenar por. En este ejemplo, basado en las entradas de la lista Ordenar por, los casos se ordenarán por el valor de *Ingresos del hogar en miles [ingres]* dentro de las categorías de *Edad en años [edad]*. Para las variables de cadena, las letras mayúsculas preceden a las minúsculas correspondientes en orden (por ejemplo, el valor de cadena *Sí* es anterior *sí* en el orden de clasificación).

Procesamiento de archivos segmentados

Para dividir el archivo de datos en diferentes grupos:

1. Elija en los menús:

Datos > Dividir archivo...

Aparecerá el cuadro de diálogo Segmentar archivo.

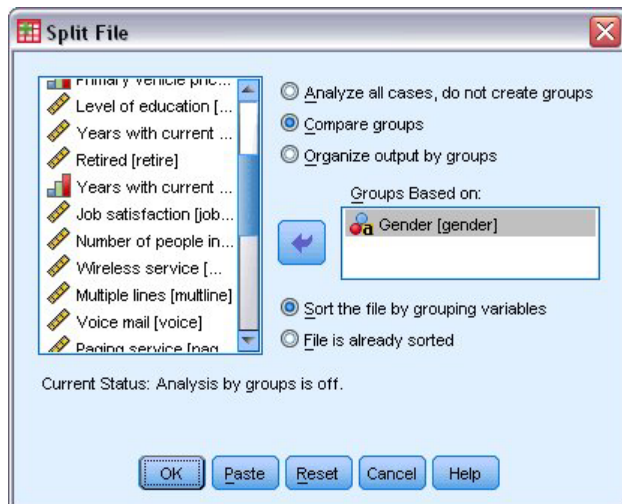


Figura 78. Cuadro de diálogo Segmentar archivo

2. Seleccione **Comparar los grupos** u **Organizar los resultados por grupos**. (Los ejemplos que siguen a estos pasos muestran las diferencias entre estas dos opciones.)
3. Seleccione *Sexo [sexo]* para dividir el archivo en grupos independientes para estas variables.

Puede utilizar variables numéricas, de cadena corta y de cadena larga como variables de agrupación. Se realizará un análisis distinto para cada uno de los subgrupos definidos por las variables de agrupación. Si selecciona varias variables de agrupación, los casos se agruparán de acuerdo con el orden en que estas variables aparezcan en la lista Grupos basados en.

Si selecciona **Comparar los grupos**, los resultados de todos los grupos de archivos segmentados se incluirán en las mismas tablas, como se muestra en la siguiente tabla de estadísticos de resumen generada por el procedimiento Frecuencias.

Estadísticos			
Ingresos del hogar en miles			
Hombre	N	Válidos	3221
		Perdidos	0
		Media	70,1608
		Mediana	45,0000
		Desv. típ.	81,56216
Mujer	N	Válidos	3179
		Perdidos	0
		Media	68,7798
		Mediana	44,0000
		Desv. típ.	75,73510

Figura 79. Salida del archivo segmentado con una sola tabla pivote

Si selecciona **Organizar los resultados por grupos** y ejecuta el procedimiento Frecuencias, se crean dos tablas pivote: una tabla para mujeres y una tabla para hombres. :

Estadísticos ^a		
Ingresos del hogar en miles		
N	Válidos	3179
	Perdidos	0
	Media	68,7798
	Mediana	44,0000
	Desv. típ.	75,73510

a. Género = Mujer

Figura 80. Salida del archivo segmentado con tabla pivote para mujeres

Estadísticos ^a		
Ingresos del hogar en miles		
N	Válidos	3221
	Perdidos	0
	Media	70,1608
	Mediana	45,0000
	Desv. típ.	81,56216

a. Género = Hombre

Figura 81. Salida del archivo segmentado con tabla pivote para hombres

Ordenación de casos para el procesamiento de archivos segmentados

El procedimiento Segmentar archivo crea un nuevo subgrupo cada vez que se encuentra un valor diferente para una de las variables de agrupación. Por tanto, es importante ordenar los casos de acuerdo con los valores de las variables de agrupación antes de invocar el procesamiento de archivos segmentados.

De forma predeterminada, el procedimiento Segmentar archivo ordena automáticamente el archivo de datos de acuerdo con los valores de las variables de agrupación. Si el archivo ya está ordenado con el orden correcto, puede ahorrar tiempo de procesamiento si selecciona **El archivo ya está ordenado**.

Activación y desactivación del procesamiento de archivos segmentados

Tras invocar el procesamiento de archivos segmentados, éste permanecerá en vigor durante el resto de la sesión, a menos que lo desactive.

- **Analizar todos los casos.** Esta opción desactiva el procesamiento de archivos segmentados.
- **Comparar grupos y Organizar los resultados por grupos.** Esta opción activa el procesamiento de archivos segmentados.

Mientras permanezca activado el procesamiento de archivos segmentados, aparecerá el mensaje **Segmentar archivo activado** en la barra de estado situada en la parte inferior de la ventana de la aplicación.

Selección de subconjuntos de casos

Puede limitar el análisis a un determinado subconjunto de acuerdo con criterios que pueden incluir 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

Para seleccionar un subconjunto de casos para su análisis:

1. Seleccione en los menús:
Datos > Seleccionar casos...

Se abrirá el cuadro de diálogo Seleccionar casos.

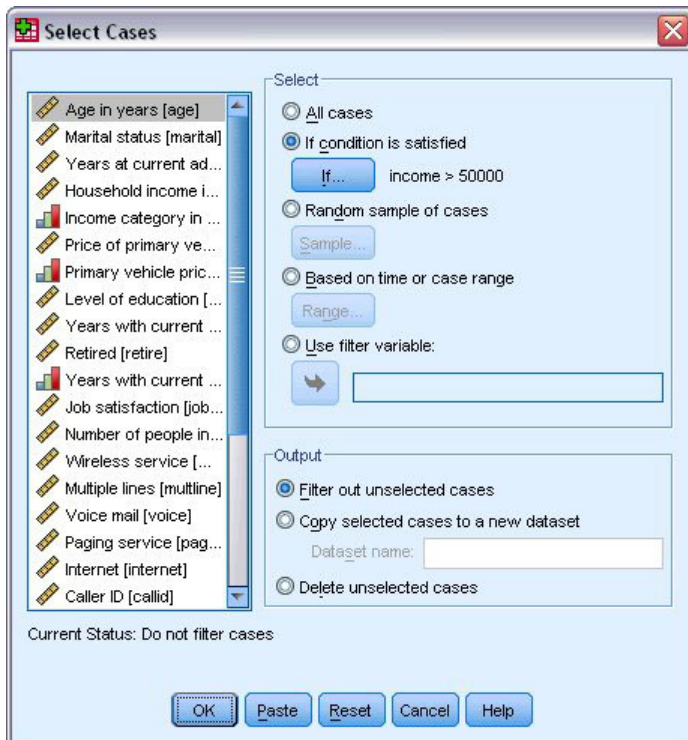


Figura 82. Cuadro de diálogo Seleccionar casos

Para seleccionar subconjuntos de casos basados en una expresión condicional

Para seleccionar los casos basados en una expresión condicional:

1. Seleccione **Si se satisface la condición** y pulse en **Si** en el cuadro de diálogo Seleccionar casos.

Se abrirá el cuadro de diálogo Seleccionar casos: Si.

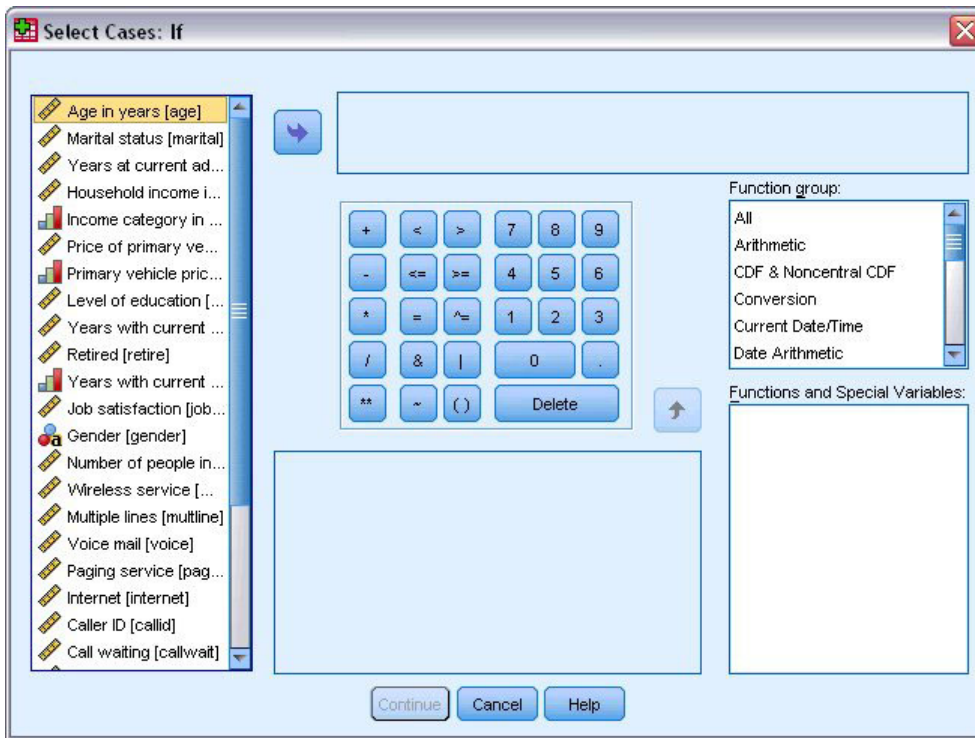


Figura 83. Cuadro de diálogo Seleccionar casos: Si

En la expresión condicional se pueden utilizar nombres de variables existentes, constantes, operadores aritméticos, operadores lógicos, operadores de relación y funciones. Puede escribir y editar la expresión en el cuadro de texto de la misma manera que cualquier otro texto de una ventana de resultados. También puede utilizar el teclado de calculadora, la lista de variables y la lista de funciones para pegar elementos en la expresión. Consulte el tema “Uso de Expresiones condicionales” en la página 67 para obtener más información.

Para seleccionar una muestra aleatoria de casos

Para obtener una muestra aleatoria:

1. Seleccione **Muestra aleatoria de casos** en el cuadro de diálogo Seleccionar casos.
2. Pulse en **Muestra**.

Se abrirá el cuadro de diálogo Seleccionar casos: Muestra aleatoria.

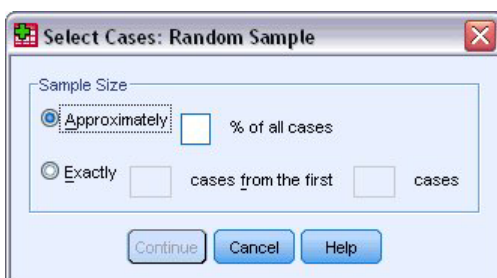


Figura 84. Cuadro de diálogo Seleccionar casos: Muestra aleatoria

Puede elegir una de las siguientes alternativas para el tamaño de la muestra:

- **Aproximadamente.** Un porcentaje especificado por el usuario. Esta opción generará una muestra aleatoria que contendrá aproximadamente el porcentaje especificado de los casos.

- **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.

Selección de un rango de tiempo o un rango de casos

Para seleccionar un rango de casos basados en fechas, horas o números de observación (filas):

1. Seleccione **Basándose en el rango del tiempo o de los casos** y pulse en **Rango** en el cuadro de diálogo Seleccionar casos.

Se abrirá el cuadro de diálogo Seleccionar casos: Rango, en el que podrá seleccionar un rango de números de observación (filas)

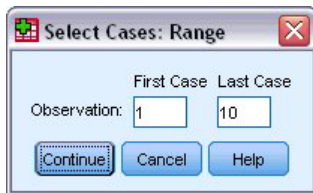


Figura 85. Cuadro de diálogo Seleccionar casos: Rango

- **Primer caso.** Introduzca los valores de la fecha y hora iniciales del rango. Si no se ha definido ninguna variable de fecha, introduzca el número de la observación inicial (número de fila del Editor de datos, a no ser que se haya activado Segmentar archivo). Si no se especifica un valor Último caso, se seleccionarán todos los casos posteriores a la fecha u hora iniciales hasta el final de la serie temporal.
- **Último caso.** Introduzca los valores de la fecha y hora finales del rango. Si no se ha definido ninguna variable de fecha, introduzca el número de la observación final (número de fila del Editor de datos, a no ser que se haya activado Segmentar archivo). Si no se especifica un valor Primer caso, se seleccionarán todos los casos desde el inicio de la serie temporal hasta la fecha u hora finales.

Para datos de serie temporal en los que se hayan definido variables de fecha, puede seleccionar un rango de fechas u horas de acuerdo con las variables de fecha definidas. Cada caso representa observaciones realizadas en un momento diferente y el archivo está ordenado en orden cronológico.



Figura 86. Cuadro de diálogo Seleccionar casos: Rango (series temporales)

Para generar variables de fecha para datos de serie temporal:

2. Elija en los menús:

Datos > Definir fechas...

Tratamiento de los casos no seleccionados

Puede elegir una de las siguientes alternativas para tratar los casos no seleccionados:

- **Descartar casos no seleccionados.** Los casos no seleccionados no se incluyen en el análisis, pero se conservan en el conjunto de datos. Podrá utilizar los casos no seleccionados más adelante en la sesión,

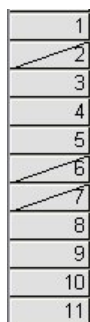
si desactiva el filtrado. Si selecciona una muestra aleatoria o si selecciona los casos mediante una expresión condicional, se generará una variable con el nombre *filter_\$* que tendrá el valor 1 para los casos seleccionados y el valor 0 para los casos no seleccionados.

- **Copiar casos seleccionados a un nuevo conjunto de datos.** Los casos seleccionados se copiarán a un nuevo conjunto de datos, lo que mantendrá inalterado el conjunto de datos original. Los casos no seleccionados no se incluirán en el nuevo conjunto de datos y se mantendrán en su estado original en el conjunto de datos original.
- **Eliminar casos no seleccionados.** Los casos no seleccionados se eliminarán del conjunto de datos. Sólo se pueden recuperar los casos eliminados saliendo del archivo sin guardar ningún cambio y abriéndolo de nuevo. La eliminación de los casos será permanente si se guardan los cambios en el archivo de datos.

Nota: Si elimina los casos no seleccionados y guarda el archivo, no será posible recuperar estos casos.

Estado de selección de casos

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 atravesando el número de fila.



1
2
3
4
5
6
7
8
9
10
11

Figura 87. Estado de selección de casos

Capítulo 10. Archivos muestrales

Los archivos muestrales instalados con el producto se encuentran en el subdirectorío *Samples* del directorío de instalación. Existe otra carpeta con los subdirectoríos de muestra de cada uno de los idiomas siguientes: inglés, francés, alemán, italiano, japonés, coreano, polaco, ruso, chino simplificado, español y chino tradicional.

No todos los archivos muestrales están disponibles en todos los idiomas. Si un archivo muestral no está disponible en un idioma, esa carpeta de idioma contendrá una versión en inglés del archivo muestral.

Descripciones

A continuación, se describen brevemente los archivos muestrales usados en varios ejemplos que aparecen a lo largo de la documentación.

- **accidents.sav.** Archivo de datos hipotéticos sobre una compañía de seguros que estudia los factores de riesgo de edad y género que influyen en los accidentes de automóviles de una región determinada. Cada caso corresponde a una clasificación cruzada de categoría de edad y género.
- **adl.sav.** Archivo de datos hipotéticos relativo a los esfuerzos para determinar las ventajas de un tipo propuesto de tratamiento para pacientes que han sufrido un derrame cerebral. Los médicos dividieron de manera aleatoria a pacientes (mujeres) que habían sufrido un derrame cerebral en dos grupos. El primer grupo recibió el tratamiento físico estándar y el segundo recibió un tratamiento emocional adicional. Tres meses después de los tratamientos, se puntuaron las capacidades de cada paciente para realizar actividades cotidianas como variables ordinales.
- **advert.sav.** Archivo de datos hipotéticos sobre las iniciativas de un minorista para examinar la relación entre el dinero invertido en publicidad y las ventas resultantes. Para ello, se han recopilado cifras de ventas anteriores y los costes de publicidad asociados.
- **aflatoxin.sav.** Archivo de datos hipotéticos sobre las pruebas realizadas en las cosechas de maíz con relación a la aflatoxina, un veneno cuya concentración varía ampliamente en los rendimientos de cultivo y entre los mismos. Un procesador de grano ha recibido 16 muestras de cada uno de los 8 rendimientos de cultivo y ha medido los niveles de aflatoxinas en partes por millón (PPM).
- **anorectic.sav.** Mientras trabajaban en una sintomatología estandarizada del comportamiento anoréxico/bulímico, los investigadores ¹ realizaron un estudio de 55 adolescentes con trastornos de la alimentación conocidos. Cada paciente fue examinado cuatro veces durante cuatro años, lo que representa un total de 220 observaciones. En cada observación, se puntuó a los pacientes por cada uno de los 16 síntomas. Faltan las puntuaciones de los síntomas para el paciente 71 en el tiempo 2, el paciente 76 en el tiempo 2 y el paciente 47 en el tiempo 3, lo que nos deja 217 observaciones válidas.
- **bankloan.sav.** Archivo de datos hipotéticos sobre las iniciativas de un banco para reducir la tasa de moras de créditos. El archivo contiene información financiera y demográfica de 850 clientes anteriores y posibles clientes. Los primeros 700 casos son clientes a los que anteriormente se les ha concedido un préstamo. Al menos 150 casos son posibles clientes cuyos riesgos de crédito el banco necesita clasificar como positivos o negativos.
- **bankloan_binning.sav.** Archivo de datos hipotéticos que contiene información financiera y demográfica sobre 5.000 clientes anteriores.

1. Van der Ham, T., J. J. Meulman, D. C. Van Strien, and H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, 363-368.

- **behavior.sav.** En un ejemplo clásico ², se pidió a 52 estudiantes que valoraran las combinaciones de 15 situaciones y 15 comportamientos en una escala de 10 puntos que oscila entre 0 = "extremadamente apropiado" y 9 = "extremadamente inapropiado". Los valores promediados respecto a los individuos se toman como disimilaridades.
- **behavior_ini.sav.** Este archivo de datos contiene una configuración inicial para una solución bidimensional de *behavior.sav*.
- **brakes.sav.** Archivo de datos hipotéticos sobre el control de calidad de una fábrica que produce frenos de disco para automóviles de alto rendimiento. El archivo de datos contiene las mediciones del diámetro de 16 discos de cada una de las 8 máquinas de producción. El diámetro objetivo para los frenos es de 322 milímetros.
- **breakfast.sav.** En un estudio clásico ³, se pidió a 21 estudiantes de administración de empresas de la Wharton School y sus cónyuges que ordenaran 15 elementos de desayuno por orden de preferencia, de 1="más preferido" a 15="menos preferido". Sus preferencias se registraron en seis escenarios distintos, de "Preferencia global" a "Aperitivo, con bebida sólo".
- **breakfast-overall.sav.** Este archivo de datos sólo contiene las preferencias de elementos de desayuno para el primer escenario, "Preferencia global".
- **broadband_1.sav** Archivo de datos hipotéticos que contiene el número de suscriptores, por región, a un servicio de banda ancha nacional. El archivo de datos contiene números de suscriptores mensuales para 85 regiones durante un período de cuatro años.
- **broadband_2.sav** Este archivo de datos es idéntico a *broadband_1.sav* pero contiene datos para tres meses adicionales.
- **car_insurance_claims.sav.** Un conjunto de datos presentados y analizados en otro lugar ⁴ estudia las reclamaciones por daños en vehículos. La cantidad de reclamaciones media se puede modelar como si tuviera una distribución Gamma, mediante una función de enlace inversa para relacionar la media de la variable dependiente con una combinación lineal de la edad del asegurado, el tipo de vehículo y la antigüedad del vehículo. El número de reclamaciones presentadas se puede utilizar como peso de escalamiento.
- **car_sales.sav.** Este archivo de datos contiene estimaciones de ventas, precios de lista y especificaciones físicas hipotéticas de varias marcas y modelos de vehículos. Los precios de lista y las especificaciones físicas se han obtenido de *edmunds.com* y de sitios de fabricantes.
- **car_sales_uprepared.sav.** Ésta es una versión modificada de *car_sales.sav* que no incluye ninguna versión transformada de los campos.
- **carpet.sav** En un ejemplo muy conocido de análisis conjunto⁵, una compañía interesada en sacar al mercado un nuevo limpiador de alfombras desea examinar la influencia de cinco factores sobre la preferencia del consumidor: diseño del producto, marca comercial, precio, sello de *Buen producto para el hogar* y garantía de devolución de importe. Hay tres niveles de factores para el diseño del producto, cada uno con una diferente colocación del cepillo del aplicador; tres nombres comerciales (*K2R*, *Glory* y *Bissell*); tres niveles de precios; y dos niveles (no o sí) para los dos últimos factores. Diez consumidores clasificaron 22 perfiles definidos por estos factores. La variable *Preferencia* contiene el rango de las clasificaciones medias de cada perfil. Las clasificaciones inferiores corresponden a preferencias elevadas. Esta variable refleja una medida global de la preferencia de cada perfil.
- **carpet_prefs.sav** Este archivo de datos se basa en el mismo ejemplo que el descrito para *carpet.sav*, pero contiene las clasificaciones reales recogidas de cada uno de los 10 consumidores. Se pidió a los consumidores que clasificaran los 22 perfiles de los productos empezando por el menos preferido. Las variables desde *PREF1* hasta *PREF22* contienen los ID de los perfiles asociados, como se definen en *carpet_plan.sav*.

2. Price, R. H., and D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579-586.

3. Green, P. E., and V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

4. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

5. Green, P. E., and Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

- **catalog.sav** Este archivo de datos contiene cifras de ventas mensuales hipotéticas de tres productos vendidos por una compañía de venta por catálogo. También se incluyen datos para cinco variables predictoras posibles.
- **catalog_seasfac.sav** Este archivo de datos es igual que *catalog.sav*, con la excepción de que incluye un conjunto de factores estacionales calculados a partir del procedimiento Descomposición estacional junto con las variables de fecha que lo acompañan.
- **cellular.sav** Archivo de datos hipotéticos sobre las iniciativas de una compañía de telefonía móvil para reducir el abandono de clientes. Las puntuaciones de propensión al abandono de clientes se aplican a las cuentas, oscilando de 0 a 100. Las cuentas con una puntuación de 50 o superior pueden estar buscando otros proveedores.
- **ceramics.sav** Archivo de datos hipotéticos sobre las iniciativas de un fabricante para determinar si una nueva aleación de calidad tiene una mayor resistencia al calor que una aleación estándar. Cada caso representa una prueba independiente de una de las aleaciones; la temperatura a la que registró el fallo del rodamiento.
- **cereal.sav** Archivo de datos hipotéticos sobre una encuesta realizada a 880 personas sobre sus preferencias en el desayuno, teniendo también en cuenta su edad, sexo, estado civil y si tienen un estilo de vida activo o no (en función de si practican ejercicio al menos dos veces a la semana). Cada caso representa un encuestado diferente.
- **clothing_defects.sav** Archivo de datos hipotéticos sobre el proceso de control de calidad en una fábrica de prendas. Los inspectores toman una muestra de prendas de cada lote producido en la fábrica, y cuentan el número de prendas que no son aceptables.
- **coffee.sav** Este archivo de datos pertenece a las imágenes percibidas de seis marcas de café helado ⁶. Para cada uno de los 23 atributos de imagen de café helado, los encuestados seleccionaron todas las marcas que quedaban descritas por el atributo. Las seis marcas se denotan AA, BB, CC, DD, EE y FF para mantener la confidencialidad.
- **contacts.sav** Archivo de datos hipotéticos sobre las listas de contactos de un grupo de representantes de ventas de ordenadores de empresa. Cada uno de los contactos está categorizado por el departamento de la compañía en el que trabaja y su categoría en la compañía. Además, también se registran los importes de la última venta realizada, el tiempo transcurrido desde la última venta y el tamaño de la compañía del contacto.
- **creditpromo.sav** Archivo de datos hipotéticos sobre las iniciativas de unos almacenes para evaluar la eficacia de una promoción de tarjetas de crédito reciente. Para este fin, se seleccionaron aleatoriamente 500 titulares. La mitad recibieron un anuncio promocionando una tasa de interés reducida sobre las ventas realizadas en los siguientes tres meses. La otra mitad recibió un anuncio estacional estándar.
- **customer_dbase.sav** Archivo de datos hipotéticos sobre las iniciativas de una compañía para usar la información de su almacén de datos para realizar ofertas especiales a los clientes con más probabilidades de responder. Se seleccionó un subconjunto de la base de clientes aleatoriamente a quienes se ofrecieron las ofertas especiales y sus respuestas se registraron.
- **customer_information.sav** Archivo de datos hipotéticos que contiene la información de correo del cliente, como el nombre y la dirección.
- **customer_subset.sav** Un subconjunto de 80 casos de *customer_dbase.sav*.
- **debate.sav** Archivos de datos hipotéticos sobre las respuestas emparejadas de una encuesta realizada a los asistentes a un debate político antes y después del debate. Cada caso corresponde a un encuestado diferente.
- **debate_aggregate.sav** Archivo de datos hipotéticos que agrega las respuestas de *debate.sav*. Cada caso corresponde a una clasificación cruzada de preferencias antes y después del debate.
- **demo.sav** Archivos de datos hipotéticos sobre una base de datos de clientes adquirida con el fin de enviar por correo ofertas mensuales. Se registra si el cliente respondió a la oferta, junto con información demográfica diversa.

6. Kennedy, R., C. Riquier, and B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56-70.

- **demo_cs_1.sav.** Archivo de datos hipotéticos sobre el primer paso de las iniciativas de una compañía para recopilar una base de datos de información de encuestas. Cada caso corresponde a una ciudad diferente, y se registra la identificación de la ciudad, la región, la provincia y el distrito.
- **demo_cs_2.sav.** Archivo de datos hipotéticos sobre el segundo paso de las iniciativas de una compañía para recopilar una base de datos de información de encuestas. Cada caso corresponde a una unidad familiar diferente de las ciudades seleccionadas en el primer paso, y se registra la identificación de la unidad, la subdivisión, la ciudad, el distrito, la provincia y la región. También se incluye la información de muestreo de las primeras dos etapas del diseño.
- **demo_cs.sav.** Archivo de datos hipotéticos que contiene información de encuestas recopilada mediante un diseño del muestreo complejo. Cada caso corresponde a una unidad familiar distinta, y se recopila información demográfica y de muestreo diversa.
- **dmdata.sav.** Es un archivo de datos hipotético que contiene información demográfica y de compra para una empresa de marketing directo. *dmdata2.sav* contiene información para un subconjunto de contactos que recibieron un envío de correos de prueba y *dmdata3.sav* contiene información sobre el resto de contactos que no recibieron el envío de correos de prueba.
- **dietstudy.sav.** Este archivo de datos hipotéticos contiene los resultados de un estudio sobre la "dieta Stillman"⁷. Cada caso corresponde a un sujeto distinto y registra sus pesos antes y después de la dieta en libras y niveles de triglicéridos en mg/100 ml.
- **dvdplayer.sav.** Archivo de datos hipotéticos sobre el desarrollo de un nuevo reproductor de DVD. El equipo de marketing ha recopilado datos de grupo de enfoque mediante un prototipo. Cada caso corresponde a un usuario encuestado diferente y registra información demográfica sobre los encuestados y sus respuestas a preguntas acerca del prototipo.
- **german_credit.sav.** Este archivo de datos se toma del conjunto de datos "German credit" de las Repository of Machine Learning Databases⁸ de la Universidad de California, Irvine.
- **grocery_1month.sav.** Este archivo de datos hipotéticos es el archivo de datos *grocery_coupons.sav* con las compras semanales "acumuladas" para que cada caso corresponda a un cliente diferente. Algunas de las variables que cambiaban semanalmente desaparecen de los resultados, y la cantidad gastada registrada se convierte ahora en la suma de las cantidades gastadas durante las cuatro semanas del estudio.
- **grocery_coupons.sav.** Archivo de datos hipotéticos que contiene datos de encuestas recopilados por una cadena de tiendas de alimentación interesada en los hábitos de compra de sus clientes. Se sigue a cada cliente durante cuatro semanas, y cada caso corresponde a un cliente-semana distinto y registra información sobre dónde y cómo compran los clientes, incluida la cantidad que invierten en comestibles durante esa semana.
- **guttman.sav.** Bell⁹ presentó una tabla para ilustrar posibles grupos sociales. Guttman¹⁰ utilizó parte de esta tabla, en la que se cruzaron cinco variables que describían elementos como la interacción social, sentimientos de pertenencia a un grupo, proximidad física de los miembros y grado de formalización de la relación con siete grupos sociales teóricos, incluidos multitudes (por ejemplo, las personas que acuden a un partido de fútbol), espectadores (por ejemplo, las personas que acuden a un teatro o de una conferencia), públicos (por ejemplo, los lectores de periódicos o los espectadores de televisión), muchedumbres (como una multitud pero con una interacción mucho más intensa), grupos primarios (íntimos), grupos secundarios (voluntarios) y la comunidad moderna (confederación débil que resulta de la proximidad cercana física y de la necesidad de servicios especializados).

7. Rickman, R., N. Mitchell, J. Dingman, and J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228:, 54-58.

8. Blake, C. L., and C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

9. Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

10. Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469-506.

- **health_funding.sav.** Archivo de datos hipotéticos que contiene datos sobre inversión en sanidad (cantidad por 100 personas), tasas de enfermedad (índice por 10.000 personas) y visitas a centros de salud (índice por 10.000 personas). Cada caso representa una ciudad diferente.
- **hivassay.sav.** Archivo de datos hipotéticos sobre las iniciativas de un laboratorio farmacéutico para desarrollar un ensayo rápido para detectar la infección por VIH. Los resultados del ensayo son ocho tonos de rojo con diferentes intensidades, donde los tonos más oscuros indican una mayor probabilidad de infección. Se llevó a cabo una prueba de laboratorio de 2.000 muestras de sangre, de las cuales una mitad estaba infectada con el VIH y la otra estaba limpia.
- **hourlywagedata.sav.** Archivo de datos hipotéticos sobre los salarios por horas de enfermeras de puestos de oficina y hospitales y con niveles distintos de experiencia.
- **insurance_claims.sav.** Éste es un archivo de datos hipotéticos sobre una compañía de seguros que desea generar un modelo para señalar las reclamaciones sospechosas y potencialmente fraudulentas. Cada caso representa una reclamación diferente.
- **insure.sav.** Archivo de datos hipotéticos sobre una compañía de seguros que estudia los factores de riesgo que indican si un cliente tendrá que hacer una reclamación a lo largo de un contrato de seguro de vida de 10 años. Cada caso del archivo de datos representa un par de contratos (de los que uno registró una reclamación y el otro no), agrupados por edad y sexo.
- **judges.sav.** Archivo de datos hipotéticos sobre las puntuaciones concedidas por jueces cualificados (y un aficionado) a 300 actuaciones gimnásticas. Cada fila representa una actuación diferente; los jueces vieron las mismas actuaciones.
- **kinship_dat.sav.** Rosenberg y Kim ¹¹ comenzaron a analizar 15 términos de parentesco (tía, hermano, primo, hija, padre, nieta, abuelo, abuela, nieto, madre, sobrino, sobrina, hermana, hijo, tío). Le pidieron a cuatro grupos de estudiantes universitarios (dos masculinos y dos femeninos) que ordenaran estos grupos según las similitudes. A dos grupos (uno masculino y otro femenino) se les pidió que realizaran la ordenación dos veces, pero que la segunda ordenación la hicieran según criterios distintos a los de la primera. Así, se obtuvo un total de seis "orígenes". Cada origen se corresponde con una matriz de proximidades de 15 x 15, cuyas casillas son iguales al número de personas de un origen menos el número de veces que se partitionaron los objetos en ese origen.
- **kinship_ini.sav.** Este archivo de datos contiene una configuración inicial para una solución tridimensional de *kinship_dat.sav*.
- **kinship_var.sav.** Este archivo de datos contiene variables independientes *sexo*, *gener*(ación), y *grado* (de separación) que se pueden usar para interpretar las dimensiones de una solución para *kinship_dat.sav*. Concretamente, se pueden usar para restringir el espacio de la solución a una combinación lineal de estas variables.
- **marketvalues.sav.** Archivo de datos sobre las ventas de casas en una nueva urbanización de Algonquin, Ill., durante los años 1999 y 2000. Los datos de estas ventas son públicos.
- **nhis2000_subset.sav.** La National Health Interview Survey (NHIS, encuesta del Centro Nacional de Estadísticas de Salud de EE.UU.) es una encuesta detallada realizada entre la población civil de Estados Unidos. Las encuestas se realizaron en persona a una muestra representativa de las unidades familiares del país. Se recogió tanto la información demográfica como las observaciones acerca del estado y los hábitos de salud de los integrantes de cada unidad familiar. Este archivo de datos contiene un subconjunto de información de la encuesta de 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Archivo de datos y documentación de uso público. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/. Fecha de acceso: 2003.

11. Rosenberg, S., and M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

- **ozono.sav.** Los datos incluyen 330 observaciones de seis variables meteorológicas para pronosticar la concentración de ozono a partir del resto de variables. Los investigadores anteriores^{12, 13} han encontrado que no hay linealidad entre estas variables, lo que dificulta los métodos de regresión estándar.
- **pain_medication.sav.** Este archivo de datos hipotéticos contiene los resultados de una prueba clínica sobre medicación antiinflamatoria para tratar el dolor artrítico crónico. Resulta de particular interés el tiempo que tarda el fármaco en hacer efecto y cómo se compara con una medicación existente.
- **patient_los.sav.** Este archivo de datos hipotéticos contiene los registros de tratamiento de pacientes que fueron admitidos en el hospital ante la posibilidad de sufrir un infarto de miocardio (IM o "ataque al corazón"). Cada caso corresponde a un paciente distinto y registra diversas variables relacionadas con su estancia hospitalaria.
- **patlos_sample.sav.** Este archivo de datos hipotéticos contiene los registros de tratamiento de una muestra de pacientes que recibieron trombolíticos durante el tratamiento del infarto de miocardio (IM o "ataque al corazón"). Cada caso corresponde a un paciente distinto y registra diversas variables relacionadas con su estancia hospitalaria.
- **poll_cs.sav.** Archivo de datos hipotéticos sobre las iniciativas de los encuestadores para determinar el nivel de apoyo público a una ley antes de una asamblea legislativa. Los casos corresponden a votantes registrados. Cada caso registra el condado, la población y el vecindario en el que vive el votante.
- **poll_cs_sample.sav.** Este archivo de datos hipotéticos contiene una muestra de los votantes enumerados en *poll_cs.sav*. La muestra se tomó según el diseño especificado en el archivo de plan *poll_csplan* y este archivo de datos registra las probabilidades de inclusión y las ponderaciones muestrales. Sin embargo, tenga en cuenta que debido a que el plan muestral hace uso de un método de probabilidad proporcional al tamaño (PPS), también existe un archivo que contiene las probabilidades de selección conjunta (*poll_jointprob.sav*). Las variables adicionales que corresponden a los datos demográficos de los votantes y sus opiniones sobre la propuesta de ley se recopilaron y añadieron al archivo de datos después de tomar la muestra.
- **property_assess.sav.** Archivo de datos hipotéticos sobre las iniciativas de un asesor del condado para mantener actualizada la evaluación de los valores de las propiedades utilizando recursos limitados. Los casos corresponden a las propiedades vendidas en el condado el año anterior. Cada caso del archivo de datos registra la población en que se encuentra la propiedad, el último asesor que visitó la propiedad, el tiempo transcurrido desde la última evaluación, la valoración realizada en ese momento y el valor de venta de la propiedad.
- **property_assess_cs.sav.** Archivo de datos hipotéticos sobre las iniciativas de un asesor de un estado para mantener actualizada la evaluación de los valores de las propiedades utilizando recursos limitados. Los casos corresponden a propiedades del estado. Cada caso del archivo de datos registra el condado, la población y el vecindario en el que se encuentra la propiedad, el tiempo transcurrido desde la última evaluación y la valoración realizada en ese momento.
- **property_assess_cs_sample.sav** Este archivo de datos hipotéticos contiene una muestra de las propiedades recogidas en *property_assess_cs.sav*. La muestra se tomó en función del diseño especificado en el archivo de plan *property_assess_csplan*, y este archivo de datos registra las probabilidades de inclusión y las ponderaciones muestrales. La variable adicional *Valor actual* se recopiló y añadió al archivo de datos después de tomar la muestra.
- **recidivism.sav.** Archivo de datos hipotéticos sobre las iniciativas de una agencia de orden público para comprender los índices de reincidencia en su área de jurisdicción. Cada caso corresponde a un infractor anterior y registra su información demográfica, algunos detalles de su primer delito y, a continuación, el tiempo transcurrido desde su segundo arresto, si ocurrió en los dos años posteriores al primer arresto.
- **recidivism_cs_sample.sav.** Archivo de datos hipotéticos sobre las iniciativas de una agencia de orden público para comprender los índices de reincidencia en su área de jurisdicción. Cada caso corresponde

12. Breiman, L., and J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580-598.

13. Hastie, T., and R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

a un delincuente anterior, puesto en libertad tras su primer arresto durante el mes de junio de 2003 y registra su información demográfica, algunos detalles de su primer delito y los datos de su segundo arresto, si se produjo antes de finales de junio de 2006. Los delincuentes se seleccionaron de una muestra de departamentos según el plan de muestreo especificado en *recidivism_cs.csplan*. Como este plan utiliza un método de probabilidad proporcional al tamaño (PPS), también existe un archivo que contiene las probabilidades de selección conjunta (*recidivism_cs_jointprob.sav*).

- **rfm_transactions.sav.** Archivo de datos hipotéticos que contiene datos de transacciones de compra, incluida la fecha de compra, los artículos adquiridos y el importe de cada transacción.
- **salesperformance.sav.** Archivo de datos hipotéticos sobre la evaluación de dos nuevos cursos de formación de ventas. Sesenta empleados, divididos en tres grupos, reciben formación estándar. Además, el grupo 2 recibe formación técnica; el grupo 3, un tutorial práctico. Cada empleado se sometió a un examen al final del curso de formación y se registró su puntuación. Cada caso del archivo de datos representa a un alumno distinto y registra el grupo al que fue asignado y la puntuación que obtuvo en el examen.
- **satisf.sav.** Archivo de datos hipotéticos sobre una encuesta de satisfacción llevada a cabo por una empresa minorista en cuatro tiendas. Se encuestó a 582 clientes en total y cada caso representa las respuestas de un único cliente.
- **screws.sav** Este archivo de datos contiene información acerca de las características de tornillos, pernos, clavos y tacos ¹⁴.
- **shampoo_ph.sav.** Archivo de datos hipotéticos sobre el control de calidad en una fábrica de productos para el cabello. Se midieron seis lotes de resultados distintos en intervalos regulares y se registró su pH. El intervalo objetivo es 4,5–5,5.
- **ships.sav.** Un conjunto de datos presentados y analizados en otro lugar ¹⁵ sobre los daños en los cargueros producidos por las olas. Los recuentos de incidentes se pueden modelar como si ocurrieran con una tasa de Poisson dado el tipo de barco, el período de construcción y el período de servicio. Los meses de servicio agregados para cada casilla de la tabla formados por la clasificación cruzada de factores proporcionan valores para la exposición al riesgo.
- **site.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía para seleccionar sitios nuevos para sus negocios en expansión. Se ha contratado a dos consultores para evaluar los sitios de forma independiente, quienes, además de un informe completo, han resumido cada sitio como una posibilidad "buena", "media" o "baja".
- **smokers.sav.** Este archivo de datos es un resumen de la encuesta sobre toxicomanía 1998 National Household Survey of Drug Abuse y es una muestra de probabilidad de unidades familiares americanas. (<http://dx.doi.org/10.3886/ICPSR02934>) Así, el primer paso de un análisis de este archivo de datos debe ser ponderar los datos para reflejar las tendencias de población.
- **stocks.sav** Este archivo de datos hipotéticos contiene precios de acciones y volumen de un año.
- **stroke_clean.sav.** Este archivo de datos hipotéticos contiene el estado de una base de datos médica después de haberla limpiado mediante los procedimientos de la opción Preparación de datos.
- **stroke_invalid.sav.** Este archivo de datos hipotéticos contiene el estado inicial de una base de datos médica que incluye contiene varios errores de entrada de datos.
- **stroke_survival.** Este archivo de datos hipotéticos registra los tiempos de supervivencia de los pacientes que finalizan un programa de rehabilitación tras un ataque isquémico. Tras el ataque, la ocurrencia de infarto de miocardio, ataque isquémico o ataque hemorrágico se anotan junto con el momento en el que se produce el evento registrado. La muestra está truncada a la izquierda ya que únicamente incluye a los pacientes que han sobrevivido al final del programa de rehabilitación administrado tras el ataque.
- **stroke_valid.sav.** Este archivo de datos hipotéticos contiene el estado de una base de datos médica después de haber comprobado los valores mediante el procedimiento Validar datos. Sigue conteniendo casos potencialmente anómalos.

14. Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.

15. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

- **survey_sample.sav.** Este archivo de datos contiene datos de encuestas, incluyendo datos demográficos y diferentes medidas de actitud. Se basa en un subconjunto de variables de NORC General Social Survey de 1998, aunque algunos valores de datos se han modificado y que existen variables ficticias adicionales se han añadido para demostraciones.
- **telco.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía de telecomunicaciones para reducir el abandono de clientes en su base de clientes. Cada caso corresponde a un cliente distinto y registra diversa información demográfica y de uso del servicio.
- **telco_extra.sav.** Este archivo de datos es similar al archivo de datos *telco.sav*, pero las variables de meses con servicio y gasto de clientes transformadas logarítmicamente se han eliminado y sustituido por variables de gasto del cliente transformadas logarítmicamente tipificadas.
- **telco_missing.sav.** Este archivo de datos es un subconjunto del archivo de datos *telco.sav*, pero algunos valores de datos demográficos se han sustituido con valores perdidos.
- **testmarket.sav.** Archivo de datos hipotéticos sobre los planes de una cadena de comida rápida para añadir un nuevo artículo a su menú. Hay tres campañas posibles para promocionar el nuevo producto, por lo que el artículo se presenta en ubicaciones de varios mercados seleccionados aleatoriamente. Se utiliza una promoción diferente en cada ubicación y se registran las ventas semanales del nuevo artículo durante las primeras cuatro semanas. Cada caso corresponde a una ubicación semanal diferente.
- **testmarket_1month.sav.** Este archivo de datos hipotéticos es el archivo de datos *testmarket.sav* con las ventas semanales "acumuladas" para que cada caso corresponda a una ubicación diferente. Como resultado, algunas de las variables que cambiaban semanalmente desaparecen y las ventas registradas se convierten en la suma de las ventas realizadas durante las cuatro semanas del estudio.
- **tree_car.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de precios de compra de vehículos.
- **tree_credit.sav** Archivo de datos hipotéticos que contiene datos demográficos y de historial de créditos bancarios.
- **tree_missing_data.sav** Archivo de datos hipotéticos que contiene datos demográficos y de historial de créditos bancarios con un elevado número de valores perdidos.
- **tree_score_car.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de precios de compra de vehículos.
- **tree_textdata.sav.** Archivo de datos sencillos con dos variables diseñadas principalmente para mostrar el estado predeterminado de las variables antes de realizar la asignación de nivel de medición y etiquetas de valor.
- **tv-survey.sav.** Archivo de datos hipotéticos sobre una encuesta dirigida por un estudio de TV que está considerando la posibilidad de ampliar la emisión de un programa de éxito. Se preguntó a 906 encuestados si verían el programa en distintas condiciones. Cada fila representa un encuestado diferente; cada columna es una condición diferente.
- **ulcer_recurrence.sav.** Este archivo contiene información parcial de un estudio diseñado para comparar la eficacia de dos tratamientos para prevenir la reaparición de úlceras. Constituye un buen ejemplo de datos censurados por intervalos y se ha presentado y analizado en otro lugar ¹⁶.
- **ulcer_recurrence_recoded.sav.** Este archivo reorganiza la información de *ulcer_recurrence.sav* para permitir modelar la probabilidad de eventos de cada intervalo del estudio en lugar de sólo la probabilidad de eventos al final del estudio. Se ha presentado y analizado en otro lugar ¹⁷.
- **verd1985.sav.** Archivo de datos sobre una encuesta ¹⁸. Se han registrado las respuestas de 15 sujetos a 8 variables. Se han dividido las variables de interés en tres grupos. El conjunto 1 incluye *edad* y *ecivil*, el conjunto 2 incluye *mascota* y *noticia*, mientras que el conjunto 3 incluye *música* y *vivir*. Se escala *mascota* como nominal múltiple y *edad* como ordinal; el resto de variables se escalan como nominal simple.

16. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

17. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

18. Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

- **virus.sav.** Archivo de datos hipotéticos sobre las iniciativas de un proveedor de servicios de Internet (ISP) para determinar los efectos de un virus en sus redes. Se ha realizado un seguimiento (aproximado) del porcentaje de tráfico de correos electrónicos infectados en sus redes a lo largo del tiempo, desde el momento en que se descubre hasta que la amenaza se contiene.
- **wheeze_steubenville.sav.** Subconjunto de un estudio longitudinal de los efectos sobre la salud de la polución del aire en los niños ¹⁹. Los datos contienen medidas binarias repetidas del estado de las sibilancias en niños de Steubenville, Ohio, con edades de 7, 8, 9 y 10 años, junto con un registro fijo de si la madre era fumadora durante el primer año del estudio.
- **workprog.sav.** Archivo de datos hipotéticos sobre un programa de obras del gobierno que intenta colocar a personas desfavorecidas en mejores trabajos. Se siguió una muestra de participantes potenciales del programa, algunos de los cuales se seleccionaron aleatoriamente para entrar en el programa, mientras que otros no siguieron esta selección aleatoria. Cada caso representa un participante del programa diferente.
- **worldsales.sav** Este archivo de datos hipotéticos contiene ingresos por ventas por continente y producto.

19. Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, and B. G. Ferris Jr. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, 366-374.

Avisos

Esta información se ha desarrollado para productos y servicios ofrecidos en EE.UU.

Es posible que IBM no ofrezca los productos, servicios y las características que se tratan en este documento en otros países. Consulte a su representante local de IBM para obtener información sobre los productos y servicios disponibles actualmente en su localidad. Las referencias hechas a un producto, programa o servicio de IBM no pretenden establecer ni implicar que sólo se pueda utilizar dicho producto, programa o servicio de IBM. En su lugar se puede utilizar cualquier producto, programa o servicio funcionalmente equivalente que no vulnere ningún derecho de propiedad intelectual de IBM. Sin embargo, es responsabilidad del usuario evaluar y comprobar el funcionamiento de todo producto, programa o servicio que no sea de IBM.

IBM puede tener patentes o solicitudes de patentes pendientes que cubran la materia descrita en este documento. Este documento no le otorga ninguna licencia para estas patentes. Puede enviar preguntas acerca de las licencias, por escrito, a:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
EE.UU.

Para consultas de licencias relacionadas con información de doble byte (DBCS), póngase en contacto con el Departamento de Propiedad intelectual de IBM de su país o envíe sus consultas, por escrito, a:

Licencia de propiedad intelectual
Ley de propiedad intelectual y legal
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japón

El párrafo siguiente no se aplica al Reino Unido ni a ningún otro país donde estas disposiciones sean incompatibles con la legislación local: INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN TAL CUAL, SIN NINGÚN TIPO DE GARANTÍA, EXPLÍCITAS NI IMPLÍCITAS, INCLUYENDO PERO NO LIMITÁNDOSE A ELLAS, LAS GARANTÍAS IMPLÍCITAS DE NO VULNERACIÓN, COMERCIALIZACIÓN O IDONEIDAD PARA UN PROPÓSITO DETERMINADO. Algunos estados no permiten la renuncia a expresar o a garantías implícitas en determinadas transacciones, por lo tanto, esta declaración no se aplique a usted.

Esta información puede incluir imprecisiones técnicas o errores tipográficos. Periódicamente, se efectúan cambios en la información aquí y estos cambios se incorporarán en nuevas ediciones de la publicación. IBM puede realizar mejoras y/o cambios en los productos y/o programas descritos en esta publicación en cualquier momento, sin previo aviso.

Cualquier referencia a sitios Web que no sean de IBM en esta información sólo es ofrecida por comodidad y de ningún modo sirve como aprobación de esos sitios Web. Los materiales que se encuentran en estos sitios web no forman parte de los materiales de este producto IBM y su utilización es por cuenta y riesgo del usuario.

IBM puede utilizar o distribuir la información que proporcione el usuario de la manera que IBM crea más conveniente sin incurrir en ninguna obligación con el usuario.

Los licenciatarios de este programa que deseen obtener información acerca del mismo con el fin de: (i) intercambiar la información entre programas creados independientemente y otros programas (incluyendo éste) y (ii) utilizar mutuamente la información que se ha intercambiado, deben ponerse en contacto con:

IBM Software Group
A la atención de: Licensing
200 W. Madison St.
Chicago, IL; 60606
EE.UU.

Esta información estará disponible, bajo las condiciones adecuadas, incluyendo en algunos casos el pago de una cuota.

IBM proporciona el programa bajo licencia descrito en este documento y todo el material bajo licencia disponible para el mismo bajo los términos del Acuerdo de cliente de IBM, el Acuerdo de licencia de programa internacional de IBM o cualquier acuerdo equivalente entre las dos partes.

Cualquier dato de rendimiento mencionado aquí ha sido determinado en un entorno controlado. Por lo tanto, los resultados obtenidos en otros entornos operativos pueden variar de forma significativa. Es posible que algunas mediciones se hayan realizado en sistemas en desarrollo y no existe ninguna garantía de que estas mediciones sean las mismas en los sistemas comerciales. Además, es posible que algunas mediciones hayan sido estimadas a través de extrapolación. Los resultados reales pueden variar. Los usuarios de este documento deben consultar los datos que corresponden a su entorno específico.

Se ha obtenido información acerca de productos que no son de IBM de los proveedores de esos productos, de sus publicaciones anunciadas o de otros orígenes disponibles públicamente. IBM no ha probado estos productos y no puede confirmar la precisión del rendimiento, la compatibilidad o cualquier otra reclamación relacionada con productos no IBM. Las preguntas acerca de las aptitudes de productos que no sean de IBM deben dirigirse a los proveedores de dichos productos.

Todas las declaraciones sobre el futuro del rumbo y la intención de IBM están sujetas a cambio o retirada sin previo aviso y representan únicamente metas y objetivos.

Esta información contiene ejemplos de datos e informes utilizados en operaciones comerciales diarias. Para ilustrarlos lo máximo posible, los ejemplos incluyen los nombres de las personas, empresas, marcas y productos. Todos esos nombres son ficticios y cualquier parecido con los nombres y direcciones utilizados por una empresa real es pura coincidencia.

Cada copia o cada parte de estos programas de ejemplo, o trabajos derivados, debe incluir un aviso de copyright como se indica a continuación:

Esta información contiene ejemplos de datos e informes utilizados en operaciones comerciales diarias. Para ilustrarlos lo máximo posible, los ejemplos incluyen los nombres de las personas, empresas, marcas y productos. Todos esos nombres son ficticios y cualquier parecido con los nombres y direcciones utilizados por una empresa real es pura coincidencia.

Cada copia o cada parte de estos programas de ejemplo, o trabajos derivados, debe incluir un aviso de copyright como se indica a continuación:

© nombre de la compañía) (año). Algunas partes de este código proceden de los programas de ejemplo de IBM Corp. Sample Programs.

© Copyright IBM Corp. _especifique el año o años_. Todos los derechos reservados.

Marcas comerciales

IBM, el logotipo de IBM e ibm.com son marcas registradas de International Business Machines Corp., registradas en muchas jurisdicciones en todo el mundo. Otros nombres de productos y servicios podrían ser marcas registradas de IBM u otras compañías. Hay disponible una lista actual de marcas registradas de IBM en la web, en “Copyright and trademark information (Información de copyright y marcas registradas)” en www.ibm.com/legal/copytrade.shtml.

Adobe, el logotipo Adobe, PostScript y el logotipo PostScript son marcas registradas o marcas comerciales de Adobe Systems Incorporated en Estados Unidos y/o otros países.

Intel, el logotipo de Intel, Intel Inside, el logotipo de Intel Inside, Intel Centrino, el logotipo de Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium y Pentium son marcas comerciales o marcas registradas de Intel Corporation o sus filiales en Estados Unidos y otros países.

Linux es una marca registrada de Linus Torvalds en Estados Unidos, otros países o ambos.

Microsoft, Windows, Windows NT, y el logotipo de Windows son marcas comerciales de Microsoft Corporation en Estados Unidos, otros países o ambos.

UNIX es una marca registrada de The Open Group en Estados Unidos y otros países.

Java y todas las marcas registradas y logotipos basados en Java son marcas registradas de Oracle y/o sus filiales.

Índice

A

Access (Microsoft) 9
archivos de bases de datos
lectura 9
archivos de datos de texto
lectura 12
archivos de ejemplo
ubicación 81
archivos de Excel
lectura 8
archivos de hoja de cálculo
lectura 8
lectura de nombres de variables 8
archivos de sintaxis
apertura 61
Asistente de fecha y hora 68
Asistente para bases de datos 9

C

cálculo de nuevas variables 65
capas
creación en tablas pivote 38
casos
ordenación 73, 75
selección 76
crear etiquetas de variable 17
Cuadro de diálogo Asistente 12

D

datos categóricos 23
medidas de resumen 23
datos continuos 23
datos cualitativos 23
datos cuantitativos 23
datos de escala 23
datos de intervalo 23
datos de razón 23
datos de serie de caracteres
entrar datos 16
datos nominales 23
datos numéricos 15
datos ordinales 23

E

edición de tablas pivote 39
Editor de datos
introducción de datos no
numéricos 16
introducción de datos numéricos 15
entrada de datos 15, 16
entrar datos
no numéricos 16
numéricas 15
etiquetas de valores
asignación 19

etiquetas de valores (*continuación*)
control de la presentación en el
Visor 19
variables numéricas 19
etiquetas de variable
creación 17
Excel (Microsoft)
exportación de resultados 50
exportación de resultados
a Excel 50
a Microsoft Word 50
a PowerPoint 50
HTML 57
expresiones condicionales 67

F

funciones en las expresiones 66

G

gráficos
barras 24, 29
creación de gráficos 29
histogramas 26
gráficos de barras 24

H

herramienta Ayuda de la sintaxis 60
histogramas 26
HTML
exportación de resultados 57

M

medidas de resumen
datos categóricos 23
variables de escala 25
mover
elementos en el Visor 35
elementos en tablas pivote 37

N

nivel de medición 23
nueva codificación de valores 63

O

ocultar filas y columnas en tablas
pivote 40
ordenación de casos 73

P

pegar sintaxis
desde un cuadro de diálogo 59

PowerPoint (Microsoft)
exportación de resultados 50
procesamiento de archivos
segmentados 74

R

recuentos
tablas de recuentos 23

S

selección de casos 76
sintaxis 59
subconjuntos de casos
basados en fechas y horas 79
eliminación de los casos no
seleccionados 79
expresiones condicionales 77
filtrado de los casos no
seleccionados 79
muestra aleatoria 78
selección 76
si se satisface la condición 77

T

tablas de frecuencias 23
tablas dinámicas
formatos de casilla 40
ocultación de filas y columnas 40
ocultar las separaciones decimales 40
tipos de datos de casilla 40
tablas pivote
acceso a las definiciones 36
capas 38
edición 39
formato 39
paneles de pivotado 37
transposición de filas y columnas 37
tipos de datos
para variables 18
transposición (volteado) de filas y
columnas en las tablas pivote 37

V

valores perdidos
para variables no numéricas 20
para variables numéricas 20
perdidos del sistema 19
valores perdidos del sistema 19
variables 15
etiquetas 17
tipos de datos 18
variables de escala
medidas de resumen 25
variables de fecha y hora 68

- ventanas de sintaxis
 - autocompletar 60
 - codificación de colores 60
 - edición de comandos 60
 - ejecución de comandos 59, 61
 - pegar comandos 59
 - puntos de separación 61
- Visor
 - desplazamiento de los resultados 35
 - ocultación y presentación de los resultados 35

W

- Word (Microsoft)
 - exportación de resultados 50



Impreso en España