



Fichas para investigadores

¿Cómo optimizar una matriz de datos? Utilidades básicas del SPSS

Autora: Vilà Baños, Ruth. Departamento de Métodos de Investigación y Diagnóstico en Educación. Universidad de Barcelona.

Introducción

Una vez se dispone de los datos introducidos en el paquete estadístico del SPSS (Statistical Package of Social Science), en una matriz de datos, es el momento de plantearse optimizar esa matriz para poder extraer el máximo rendimiento a los datos, según el tipo de análisis que se pretende realizar. Para ello, el propio SPSS tiene una serie de utilidades que pueden ser de gran utilidad. Estas utilidades básicas pueden diferenciarse según su funcionalidad entre: utilidades para la edición de datos, utilidades para la modificación de variables, y las opciones de ayuda que nos brinda. A continuación se presentan algunas de estas utilidades.

1. Utilidades para la edición de datos

Por edición de datos hacemos referencia a diferentes opciones como: la selección, cambio, cortar, copiar y pegar datos, insertar o borrar variables o casos, buscar valores de la variable, etc. Opciones como cortar, copiar y pegar datos son relativamente usuales en muchos softwares y su procedimiento es similar también en el SPSS. De este modo, entre las opciones enumeradas, en este apartado nos centraremos en aquellas que consideramos excepcionales para la optimización de la matriz de datos en este paquete estadístico.

Una de las primeras utilidades es la selección de casos. Esta opción es muy útil para efectuar análisis focalizados en casos concretos, por ejemplo, imaginemos que necesitamos analizar los resultados obtenidos en un cuestionario sólo en las respuestas dadas por las chicas participantes en la muestra. En este ejemplo, tal como se indica a la figura 1, primero se debe escoger la opción "seleccionar casos" del menú DATOS. Esta opción abre un nuevo cuadro de diálogo en el que haciendo clic en la opción "Si se satisface la condición" abrirá un nuevo cuadro que nos permitirá especificar los casos que deseamos seleccionar (en el ejemplo de las chicas, el valor que corresponda a femenino de la variable sexo).

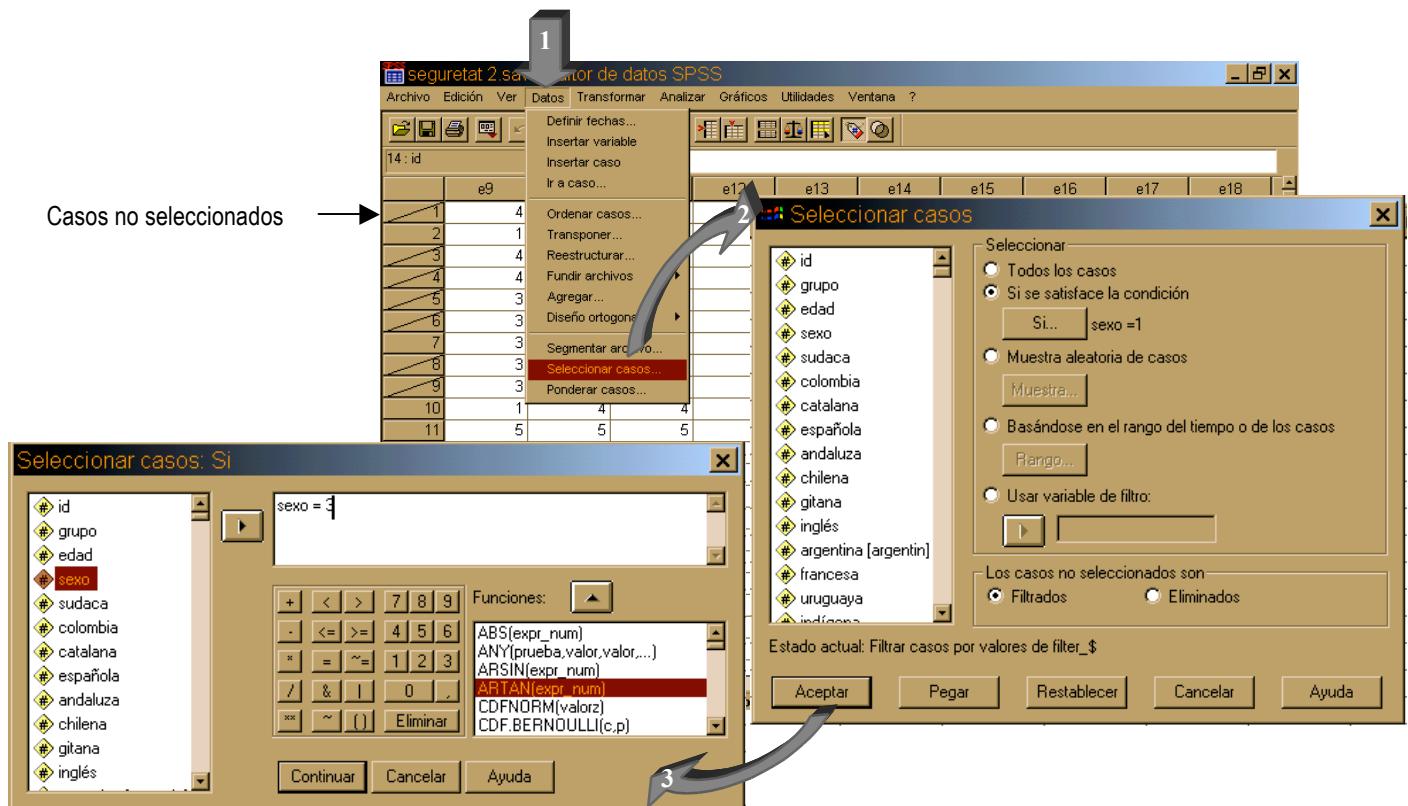


Fig. 1. Procedimiento para seleccionar casos

A partir de esta selección, es posible visualizar los casos seleccionados en la propia vista de datos, ya que aquellos casos que no son seleccionados aparecen con una línea de tachado en la primera columna de numeración de casos. Cualquier análisis que solicitemos a continuación (gráficos, tablas, estadísticos descriptivos, contrastes, correlaciones, etc.) se efectuará únicamente considerando los casos seleccionados, hasta que deshagamos esta opción mediante el mismo procedimiento pero marcando la opción todos los casos.

Otra opción interesante para la edición de datos es la búsqueda de casos. Se utiliza cuando los casos que necesitamos buscar son pocos. Por ejemplo, imaginemos que una vez introducidos los datos, depuramos la matriz en búsqueda de posibles errores, y en algunas tablas de frecuencias existe el valor 3 en sexo por error. Para poder localizar este valor, la función búsqueda de casos puede ser de gran utilidad.

En el menú edición, una vez seleccionada la variable (haciendo clic sobre su nombre en la vista de datos) en la que queremos efectuar la búsqueda, optamos por la opción buscar, y se abrirá un cuadro de diálogo en el que especificaremos el valor que deseamos buscar (figura 2).

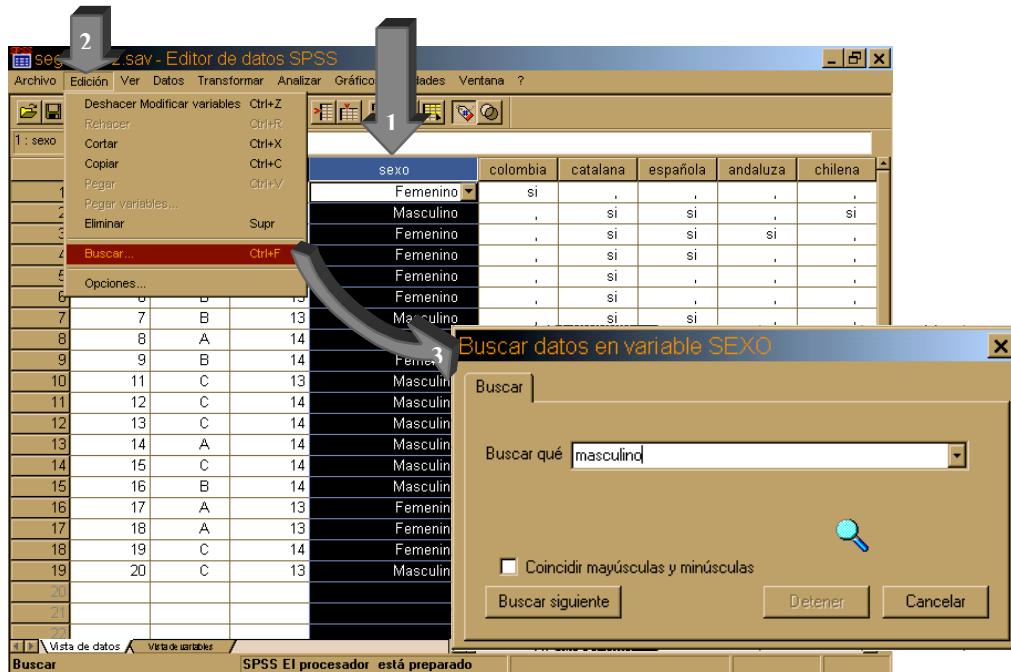


Fig. 2. Procedimiento para buscar sujetos o valores

Por último hacemos referencia a la utilidad de insertar variables y casos en una matriz ya elaborada. Estas opciones son de gran utilidad una vez ya tenemos la matriz y necesitamos añadir nuevos datos; por ejemplo, añadir una variable que identifique los cuestionarios, y por tanto, los casos de la matriz.

La figura 3 nos indica cómo insertar una nueva variable, teniendo en cuenta que ésta se generará a la izquierda de donde estaba situado el cursor en el momento de pedir la función. Una vez insertada se deberá definir la variable e introducir los valores, como se hizo al crear la matriz.



Butlletí LaRecerca

ISSN: 1886-1946 / Dipòsit legal: B.20973-2006

Ficha 11. Mayo, 2008

Vilà, R.

¿Cómo optimizar una matriz de datos? Utilidades básicas del SPSS

The screenshot shows the SPSS Data Editor window titled "seguretat 2.sav - Editor de datos SPSS". The menu bar includes Archivo, Edición, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, and Ventana. The "Datos" menu is currently active. A context menu is open over the data grid, with "Insertar variable" highlighted in red. The data grid contains columns labeled e12 through e17 and rows numbered 1 to 21. The status bar at the bottom displays "Insertar variable" and "SPSS El procesador está preparado".

Fig. 3. Procedimiento para insertar nuevas variables

Respecto a insertar nuevos casos, la figura 4 nos indica como hacer-lo, considerando que éste se insertará en la misma fila donde está situado el cursor en el momento de ejecutar la función.

The screenshot shows the SPSS Data Editor window titled "seguretat 2.sav - Editor de datos SPSS". The menu bar includes Archivo, Edición, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, and Ventana. The "Datos" menu is currently active. A context menu is open over the data grid, with "Insertar caso" highlighted in red. The data grid contains columns labeled e1 through e10 and rows numbered 1 to 21. The status bar at the bottom displays "Insertar casos" and "SPSS El procesador está preparado".

Fig. 4. Procedimiento para insertar nuevos casos



A continuación, profundizamos en otras utilidades del paquete estadístico del SPSS para la optimización de la matriz de datos y algunos de sus análisis, en esta ocasión, nos centramos en aquellas opciones que pueden ser de gran ayuda para la modificación de variables existentes y creación de nuevas.

2. Utilidades para la modificación de variables

Existen dos opciones básicas de gran utilidad para modificar las variables existentes y conseguir que se adecuen al análisis que más responda a nuestros intereses o propósitos: calcular nuevas variables o recodificarlas (cambiándoles los códigos de las categorías o alternativas de respuesta).

Efectivamente, es muy posible que necesitemos crear nuevas variables, fruto del cálculo de alguna o algunas de las variables que ya tenemos creadas e introducidas en la matriz de datos. Un ejemplo claro es una matriz que conste de las puntuaciones de los ítems de un test de rendimiento o de una escala de actitud. Evidentemente, necesitaremos realizar cálculos diversos como la corrección de las puntuaciones en algunos ítems (si la escala de actitudes tiene juicios negativos y positivos, por ejemplo), la puntuación total obtenida en el instrumento, e incluso, diferentes puntuaciones según las dimensiones del constructo medido.

Si necesitamos el total que ha obtenido cada sujeto en este instrumento, la operación que debemos efectuar es calcular una nueva variable. ¿Cómo lo hacemos? En la opción transformar del menú se encuentra la función calcular, si se abre, haciendo clic en ella, el desplegable que vemos en la figura 5, nos ofrece dos opciones:

- a) Utilizar símbolos (suma, resta, multiplicación etc.) que tenemos en el teclado y que también aparecen en el propio desplegable. Si optamos por esta alternativa debemos construir el cálculo en lenguaje matemático seleccionando las variables en el cuadro de la izquierda y los conectores del teclado o del propio desplegable (en el ejemplo anterior: item1+item2+item3+item4...etc.).
- b) Utilizar las funciones que el propio menú desplegable nos ofrece, situadas en el recuadro de la parte derecha. Si optamos por esta alternativa, debemos ser cuidadosos con la sintaxis que el propio sistema nos sugiere, en cuanto seleccionamos la función y hacemos clic en la flecha correspondiente. En el ejemplo anterior: SUM (item1,item2,item3,item4,..etc).

En los dos casos seleccionaremos las variables entre las que aparecen en el recuadro de la izquierda, siendo éstas todas las que se han definido previamente. En el ejemplo todas las variables que representan los ítems del test, y las iremos sumando entre ellas para conseguir la puntuación total del instrumento. No obstante, también podríamos haber utilizado la función sumatorio entre las funciones establecidas.

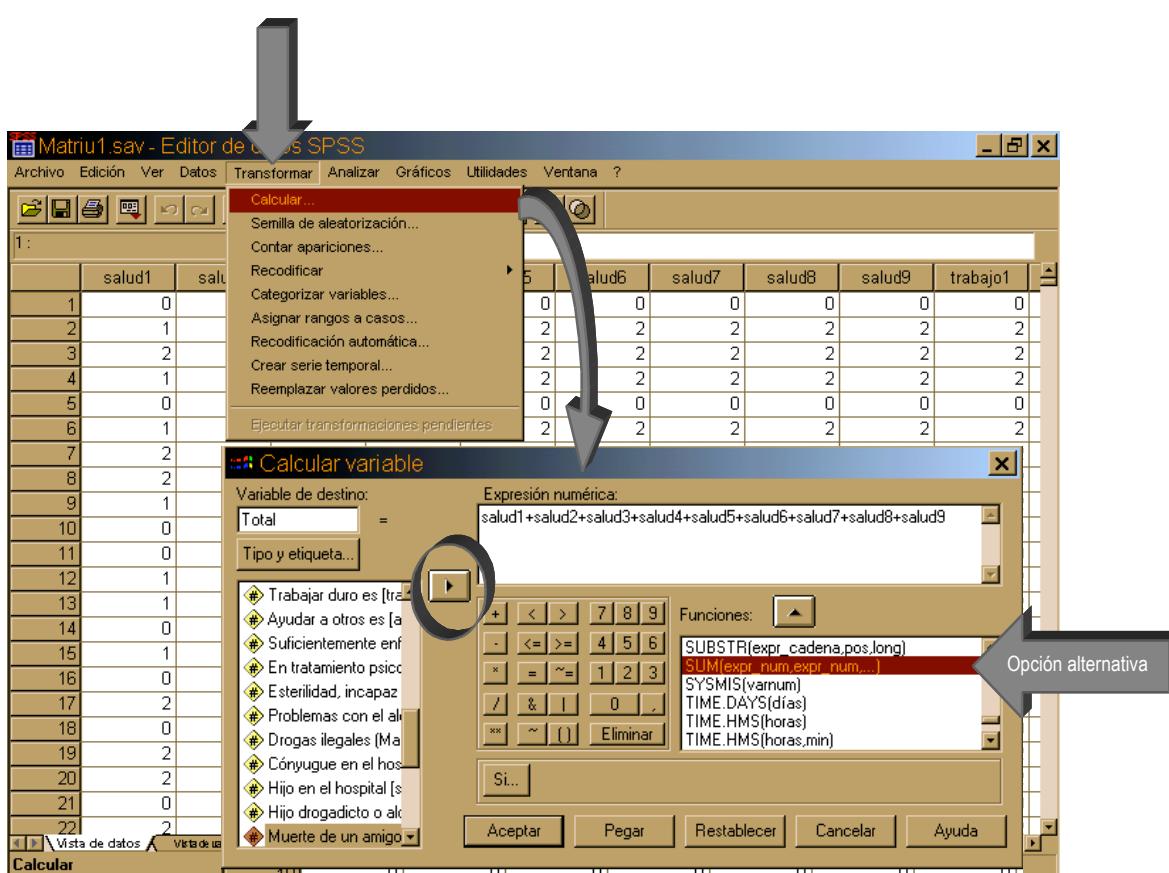


Fig. 5. Procedimiento para calcular nuevas variables

Por otra parte, es muy probable que también necesitemos, en otras ocasiones, modificar variables cambiando los valores o categorías, realizando lo que se denomina recodificar. Imaginemos que disponemos de información sobre el rendimiento académico de un grupo de alumnos en puntuaciones de 0-10, pero para nuestro estudio necesitaríamos esta información en las categorías: suspendido, aprobado, notable y excelente. Para transformar los valores del rendimiento académico debemos realizar la recodificación de la variable.

¿Cómo lo hacemos? Para recodificar las variables hemos de acceder a la opción transformar del menú y después de hacer clic en la función recodificar aparecen dos opciones:

- Recodificar en las mismas variables: Esta opción implica crear la nueva variable encima de la variable original que deseamos recodificar, eliminando el contenido anterior. En el ejemplo anterior, eliminaríamos las puntuaciones de 0-10 por las categorías propuestas.



- b) Recodificar en distintas variables: Esta opción implica crear la variable en una nueva columna, como una nueva variable sin eliminar la originaria. Si decidimos crear una nueva variable a parte, hemos de dar un nombre a la nueva variable. Por esta finalidad es importante que activemos la opción cambiar (tal como se exemplifica en el punto 3 de la figura 6).

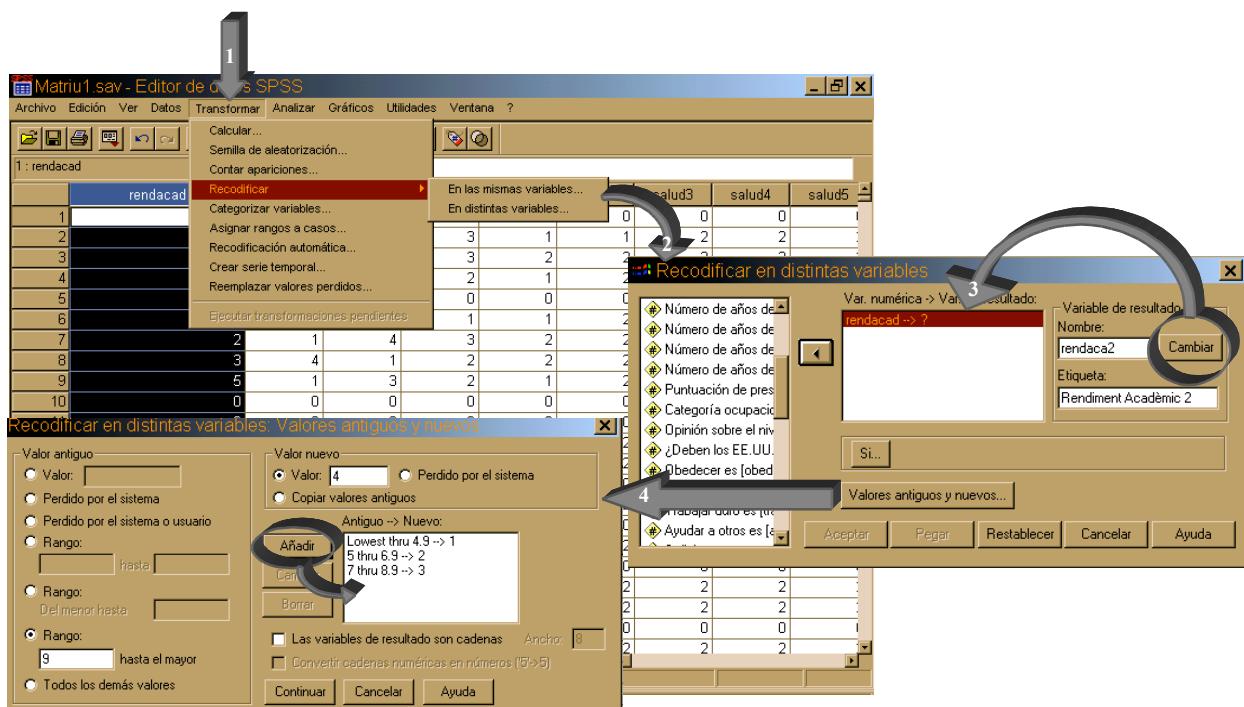


Fig. 6. Procedimiento para recodificar variables

Seguidamente debemos de indicar cuáles son los valores de cada código que necesitamos recodificar y en qué valores se transformarán. Es decir, deberemos indicar que un suspendido es aquella persona que ha obtenido una puntuación que oscila entre 0 y 4'9, etc. Para ello accedemos a la opción valores antiguos y nuevos (punto 4 de la figura 6). La ventana que se abre está dividida en dos partes verticales:

- La parte de la izquierda define los valores antiguos, es decir, los intervalos de las puntuaciones que en el ejemplo anterior son: 0-4'9 (para el suspenso), 5-6'9 (para el aprobado), 7-8'9 (para el notable), 9-10 (para el excelente).
- La parte de la derecha define los nuevos valores, que en el ejemplo anterior corresponderán a: 1 (para el suspenso), 2 (para el aprobado), 3 (para el notable), 4 (para el excelente). Como podemos comprobar introducimos valores numéricos para designar las categorías, pudiendo definir la variable y poner las



etiquetas correspondientes (suspenso, aprobado,...) en otro momento, desde la vista de variables.

Para definir estos intervalos de puntuaciones podemos utilizar la opción de rango de valores, o bien otras alternativas: por ejemplo el suspenso puede ser definido a partir del rango del menor hasta..., indicando únicamente el 4'9 como puntuación máxima; o bien, el excelente puede ser identificado a través de rango de... hasta el mayor, indicando en este caso el 9 como puntuación mínima. En todo caso, cada vez que indiquemos valores antiguos y nuevos de una categoría debemos hacer clic en añadir para que se mantengan los cambios. Una vez añadidas todas las categorías a la pequeña pantalla que se encuentra en la parte derecha de la ventana podremos aceptar la operación y veremos cómo se crea una nueva variable en la última columna de la matriz.

3. Opciones de ayuda

SPSS dispone de diferentes opciones de ayuda que pueden ser muy útiles para su uso. Esta ayuda, puede consultarse desde distintos lugares y para distintas finalidades, que intentaremos sintetizar a continuación.

En cada ventana y cuadros de diálogo aparecen botones de ayuda, que permiten el acceso a los diferentes contenidos, índices y opciones de búsqueda de términos específicos.

Por otra parte, también tenemos la posibilidad de obtener ayuda en los diferentes cuadros de diálogo, utilizando el botón derecho del ratón sobre el concepto que no se comprende. Por ejemplo, si no recordamos el concepto de la mediana, podemos situar el cursor sobre dicho término y con un clic del botón derecho del ratón (figura 7) nos aparece su definición.

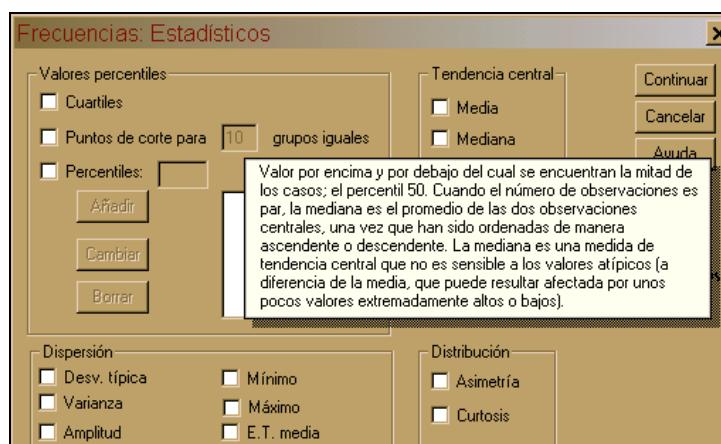


Fig. 7. Ejemplo de desplegable de ayuda que se activa con el botón derecho del ratón



Siguiendo el mismo procedimiento podremos obtener información de ayuda sobre los términos de una tabla. Para ello, la activaremos haciendo clic sobre ella y después con el botón derecho del ratón seleccionar la opción ¿Qué es esto? del menú que aparece para consultar las definiciones de los términos. Procedimiento que se resume en la figura 8.

También disponemos del asesor de resultados. Éste aparece mediante el mismo procedimiento anterior, pero seleccionando la opción asesor de resultados en el menú que se despliega (última opción tal como se representa en la figura 7). De esta manera, se obtiene información sobre la interpretación de los resultados.

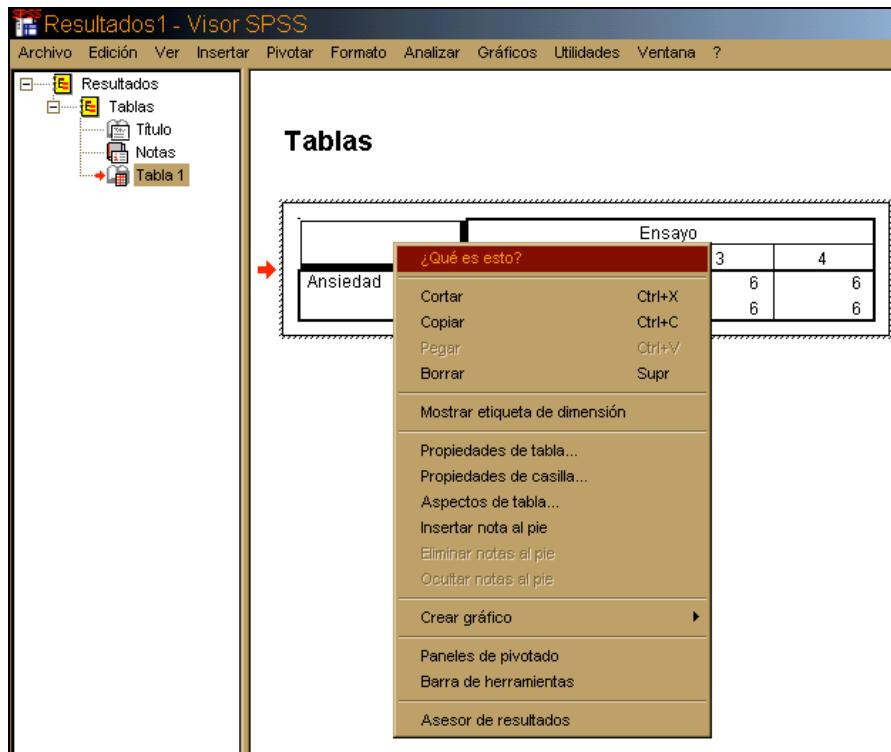


Fig. 8. Opción de ayuda que se activa con el botón derecho del ratón sobre una tabla

Otra opción de ayuda útil es el asesor estadístico. Éste aparece en el menú de ayuda y proporciona un método de asistencia para encontrar el procedimiento de análisis con el SPSS más adecuado según el tipo de resultados que deseamos obtener. Si lo activamos, el asistente planteará una serie de interrogantes sencillos a las que deberemos dar respuesta, referidos a los datos que tenemos y a los objetivos del análisis, y finalmente se propondrá el procedimiento más adecuado a seguir. En la figura 9 se resume cómo activar esta útil herramienta.

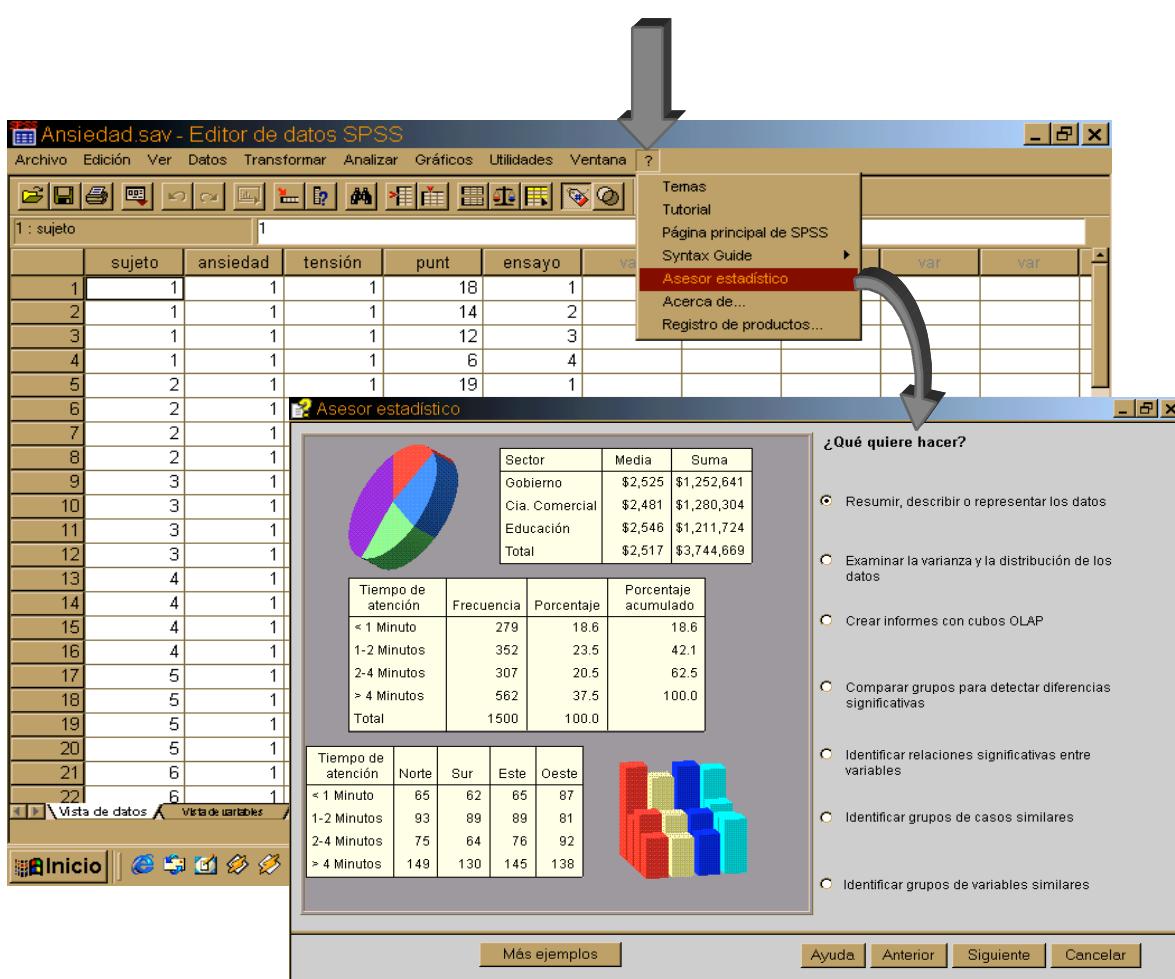


Fig. 9. Opción de ayuda del asesor estadístico

Finalmente, otra interesante aplicación es el tutorial. Se trata de una opción que también aparece en el menú general de ayuda, tal como se aprecia en la figura 10. Éste nos ofrece una explicación paso a paso de diferentes temáticas de funcionamiento del SPSS, pudiendo buscar y/o elegir estas temáticas en un índice de contenidos de ayuda.

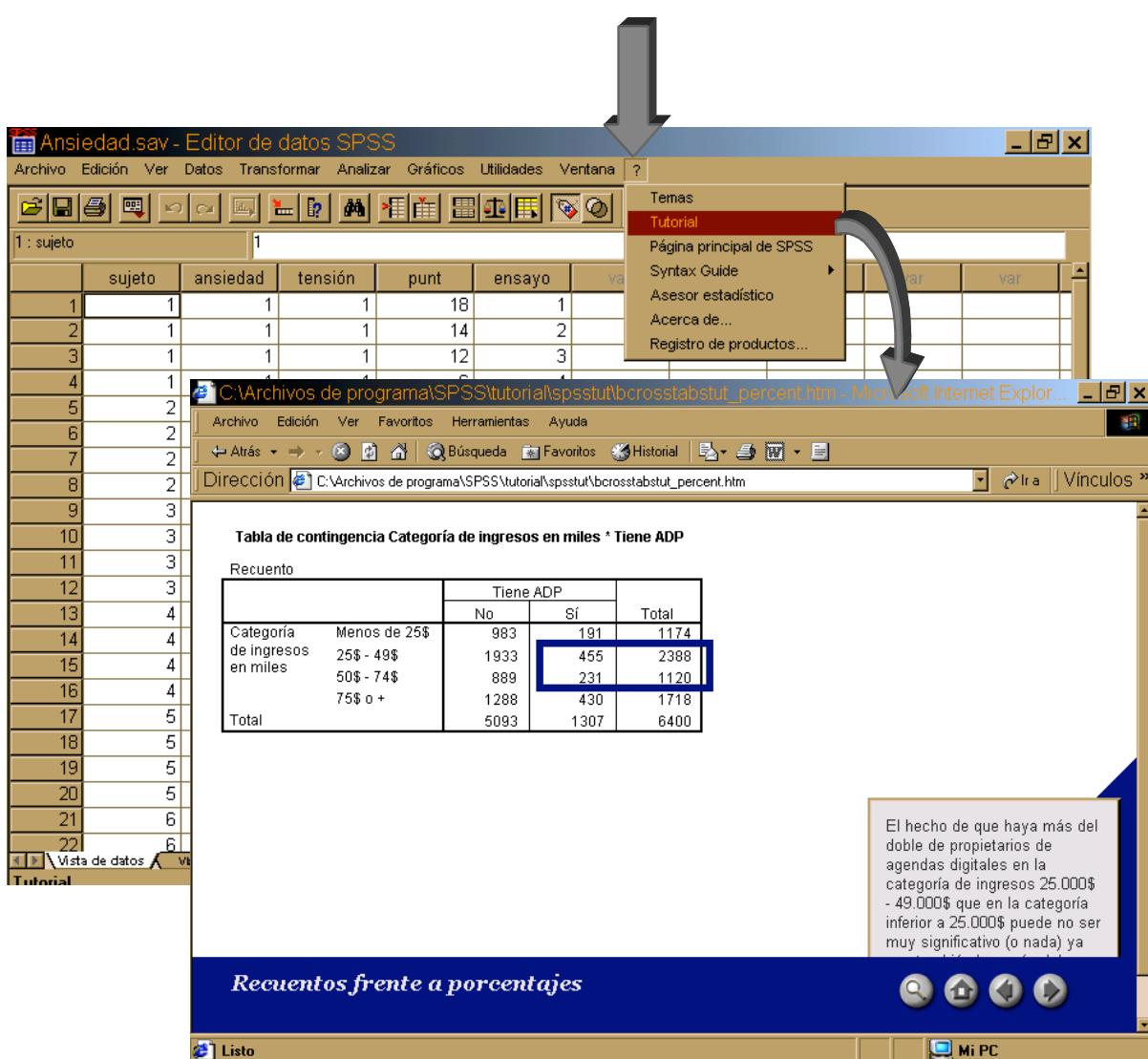


Fig. 10. Opción de ayuda tutorial



Bibliografía recomendada

- Álvarez, M. (2000). *Análisis estadístico con SPSS: procedimientos básicos*. Bilbao: Deusto.
- Bisquerra, R. (1987). *Introducción a la estadística aplicada a la investigación educativa. un enfoque informático con los paquetes BMDP y SPSS-X*. Barcelona: PPU.
- Camacho, J. (2000). *Estadística con SPSS*. Madrid: Ra-Ma.
- Filgueira, E. (2001). *Análisis de datos con SPSSWIN*. Madrid: Alianza editorial.
- Gondar, J. E. (2002). *Técnicas estadísticas con SPSS*. Madrid: Data mining institute.
- Lizasoain, L., i Joaristi, L. (1998). *SPSS para Windows*. Madrid: Paraninfo.
- Norusis, M. J. (2000). *SPSS 10.0: guide to data analysis*. New Jersey: Prentice Hall.
- Pardo, A., i Ruiz, M. A. (2002). *SPSS 11. Guía para el análisis de datos*. Madrid: McGraw-Hill.
- Visauta, B. (1997). *Análisis estadístico con SPSS para windows*. (Vol. 1. Estadística básica). Madrid: McGraw-Hill.
- Visauta, B., i Batallé, P. (1991). *Métodos estadísticos aplicados. Estadística descriptiva*. Barcelona: PPU.
- Vilà, R., i Bisquerra, R. (2004). El análisis cuantitativo de los datos. En: R. Bisquerra (ed.), *Metodología de la investigación educativa*. Madrid: La Muralla.