

Introducción a SPSS

1.1. Introducción

La gran potencia de cálculo alcanzada por los ordenadores ha convertido a los mismos en poderosas herramientas al servicio de todas aquellas disciplinas que, como la estadística, requieren manejar un gran volumen de datos. Actualmente, prácticamente nadie se plantea hacer un estudio estadístico serio sin la ayuda de un buen programa de análisis estadístico.

SPSS®* es uno de los programas de análisis estadístico más utilizados, sobre todo en el ámbito de las ciencias biosanitarias.

El objetivo de esta práctica es introducir al alumno en la utilización de este programa, enseñándole a realizar las operaciones básicas más habituales. A lo largo de la práctica, los alumnos aprenderán a crear variables, introducir datos de las muestras, transformar variables, filtrar datos y fundir e importar archivos de datos.

1.2. Funciones básicas

1.2.1. Arranque

Como cualquier otra aplicación de Windows, para arrancar el programa hay que hacer click sobre la opción correspondiente del menú Inicio ▶ Programas, o bien sobre el icono del acceso directo del escritorio que aparece en la figura 1.2.

*Este libro está basado en la versión 20.0 de SPSS® para Windows en español.



Figura 1.1: Ventana de inicio de SPSS version 20.0.



Figura 1.2: Icono del acceso directo de SPSS.

Cuando el programa arranca, aparece la ventana del editor de datos (figura 1.3).

Como cualquier otra ventana de aplicación de Windows, la ventana del editor de datos tiene una barra de título, una barra de menú con las distintas funciones que puede hacer SPSS, entre ellas los análisis estadísticos de datos, una barra de botones que son atajos a las opciones más habituales de los menús, y una barra de estado en la parte inferior que nos indica lo que hace el programa en cada instante. Además, en la parte inferior aparecen dos pestañas que permiten pasar a la Vista de datos o a la Vista de variables.

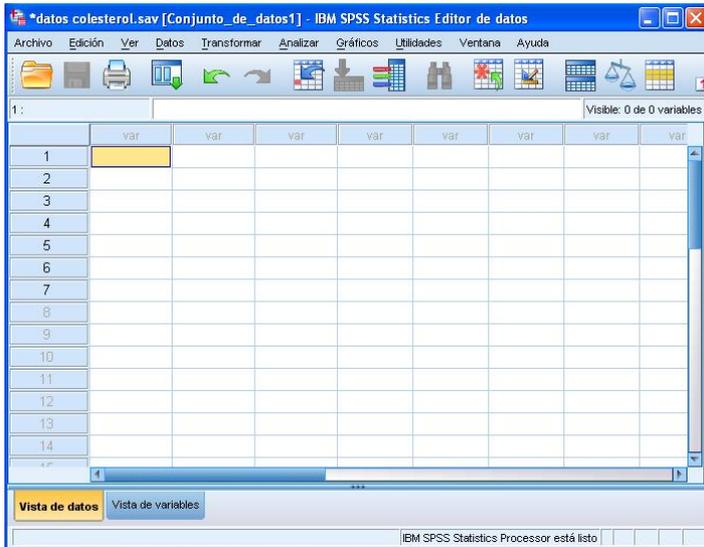


Figura 1.3: Ventana del editor de datos.

1.2.2. Introducción de datos

Para realizar cualquier análisis, la ventana del editor de datos debe contener la matriz de datos a analizar. Una vez que el usuario obtiene los datos muestrales, estos deben introducirse en esta ventana. Para ello, lo primero es definir las variables que se han considerado en el estudio. Cada variable se corresponderá con una columna de la matriz de datos.

Para definir una variable debemos pasar a la Vista de variables haciendo click sobre la correspondiente pestaña (figura 1.4).

En esta otra ventana, debemos definir cada variable en una fila, rellenando los siguientes campos:

Nombre El nombre de la variable puede ser cualquier cadena de caracteres que comience por una letra y que no contenga espacios en blanco ni caracteres especiales como ?, &, *, etc. Cada nombre de variable debe ser único y no se distingue entre mayúsculas y minúsculas.

Tipo Los tipos más comunes son Numérico (formato numérico estándar), Coma (con comas de separación cada tres cifras y punto para la parte decimal), Punto (con puntos de separación cada tres cifras y coma para la parte

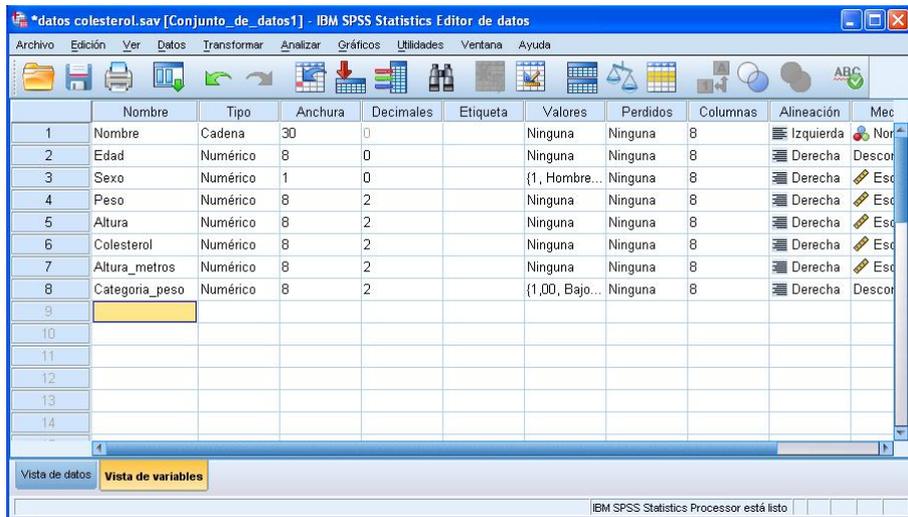


Figura 1.4: Vista de definición de variables.

decimal), Notación Científica (utiliza la E para la exponenciación), Cadena (para datos alfanuméricos) y Fecha.

Anchura Es el número máximo de caracteres que pueden tener los valores de la variable.

Decimales Para las variables numéricas es el número de cifras decimales que podrán escribirse.

Etiqueta Es una descripción de la variable. Si el nombre de la variable es suficientemente descriptivo se puede omitir.

Valores Permite asignar etiquetas a los distintos valores que puede tomar la variable. No es obligatorio pero puede ser útil en algunos casos.

Al hacer click sobre la casilla aparece un cuadro de diálogo para asignar etiquetas a valores. Para ello basta con escribir un valor en el cuadro de texto Valor y la correspondiente etiqueta en el cuadro de texto Etiqueta. Después hay que hacer click sobre el botón Añadir y repetir los mismos pasos para todos los valores de la variable. Para finalizar hay que hacer click en el botón Aceptar.

Perdidos Permite definir qué valores se utilizarán para representar los datos perdidos por el usuario. Es útil para distinguir datos que se han perdido por distintas causas. Por ejemplo, puede ser interesante distinguir el dato perdido correspondiente a un entrevistado que se niega a responder, del dato perdido debido a que la pregunta no le afectaba al entrevistado. Los valores de datos especificados como perdidos por el usuario se excluyen de la mayoría de los cálculos.

Al hacer click sobre la casilla aparece un cuadro de diálogo donde deben indicarse los valores discretos que representarán valores perdidos (pueden introducirse hasta tres), o bien, el rango de valores que se representarán como valores perdidos más un valor perdido discreto opcional.

Columnas Permite especificar el ancho de la columna en la que se introducirán los datos correspondientes a la variable.

Alineación Permite especificar la alineación de los datos correspondientes a la variable. Puede ser Izquierda, Derecha o Centrado.

Medida Permite especificar el tipo de escala utilizada para medir la variable. Puede ser Escala cuando la variable es numérica y la escala es de intervalo, Ordinal cuando los valores de la variable representan categorías con un cierto orden o Nominal cuando los valores representan categorías sin orden.

Rol Permite especificar la función que una variable tiene en el análisis. Puede ser entrada, cuando se trata de una variable independiente, objetivo, cuando es una variable dependiente, ambos, cuando la variable puede ser dependiente e independiente, ninguna si la variable no tiene ninguna función asignada, partición, cuando la variable se utilizará para dividir los datos en muestras separadas y dividir, cuando se trata de una variable introducida para asegurar la compatibilidad en SPSS.

Una vez definidas las variables se procede a introducir los datos de la muestra. Para ello hay que volver a la Ventana de datos haciendo click en la correspondiente pestaña. Ahora aparecerán en las cabeceras de columna los nombres de las variables definidas. Cada individuo de la muestra se corresponde con una fila de la matriz de datos. Para introducir el valor de una variable en un individuo determinado, nos situamos en la celda de la fila de dicho individuo y de la columna de la variable, bien haciendo click sobre la misma, o bien desplazándonos por la matriz de datos con las flechas de movimiento del cursor del teclado, y se teclea el valor seguido de la tecla Intro (figura 1.5).

	Nombre	Edad	Sexo	Peso	Altura
1	José Luis	18	1	85,00	179,00
2	Rosa Díez	32	2	65,00	173,00
3	Javier García	24	1	71,00	181,00
4	Carmen López	35	2	65,00	170,00
5	Cristobal Campos	44	1	70,00	178,00
6	Marisa López	46	2	51,00	158,00
7	Antonio Rúaiz	68	1	66,00	174,00
8					
9					
10					
11					
12					

Figura 1.5: Introducción de datos en la matriz de datos. Cada columna corresponde a una variable y cada fila a un individuo de la muestra.

1.2.3. Guardar datos

Una vez introducidos los datos, conviene guardarlos en un fichero para no tener que volver a introducirlos en futuras sesiones. Para ello, se selecciona el menú Archivo ▶ Guardar. Si el fichero ya existe, se actualizará su información, y si no, aparecerá un cuadro de diálogo en el que hay que introducir el nombre que queremos darle al fichero y la carpeta donde lo queremos ubicar. Los ficheros de datos de SPSS tienen por defecto extensión *.sav. Cuando los datos estén guardados en un fichero, el nombre del fichero aparecerá en el título de la ventana de datos (figura 1.5).

1.2.4. Recuperar datos

Si los datos con los que se pretende trabajar ya están guardados en un fichero, entonces tendremos que abrir dicho fichero. Para ello, se selecciona el menú Archivo ▶ Abrir ▶ Datos y se selecciona el fichero que se desea abrir. Automáticamente, los datos aparecerán en la vista de datos.

1.2.5. Modificación de datos

En ocasiones es necesario modificar los datos de la matriz de datos para corregir errores, añadir nuevos datos o eliminarlos. Para corregir un valor basta con seleccionar la celda que contiene el valor y teclear el nuevo. Otras operaciones habituales son:

- Insertar una variable nueva entre otras ya existentes. En la vista de variables se selecciona la fila que contiene la variable por encima de la cual queremos insertar la nueva, y se selecciona el menú Edición ▶ Insertar variable.
- Eliminar una variable. En la vista de variables se selecciona la fila que contiene la variable a eliminar y se pulsa la tecla Supr.
- Insertar un individuo entre otros ya existentes. En la vista de datos se selecciona la fila que contiene los datos del individuo por encima del cual queremos insertar el nuevo, y se selecciona el menú Edición ▶ Insertar caso.
- Eliminar un individuo. En la vista de datos se selecciona la fila que contiene los datos del individuo a eliminar y se se presiona la tecla Supr.

Cada vez que realicemos modificaciones en la matriz de datos, conviene volver a guardar los datos para que se actualice el fichero que los contiene.

¡Importante! Cuando por equivocación realicemos una operación no deseada, podemos deshacerla mediante el menú Edición ▶ Deshacer.

1.2.6. Transformación y generación de datos

En muchos análisis estadísticos se suelen transformar los datos de las variables originales en otros más convenientes para el análisis que se vaya a efectuar. Para generar una nueva variable mediante una transformación de otra ya existente o bien mediante funciones ya predefinidas se selecciona el menú Transformar ▶ Calcular Variable... Entonces aparece la ventana de transformación de variables tal y como se muestra en la figura 1.6.

En esta ventana se debe introducir el nombre de la nueva variable en el cuadro Variable de destino, y la expresión cuyo resultado será el contenido de la nueva variable en el cuadro Expresión numérica. Para ello aparecen toda una serie de operadores y funciones para realizar la transformación, así como la

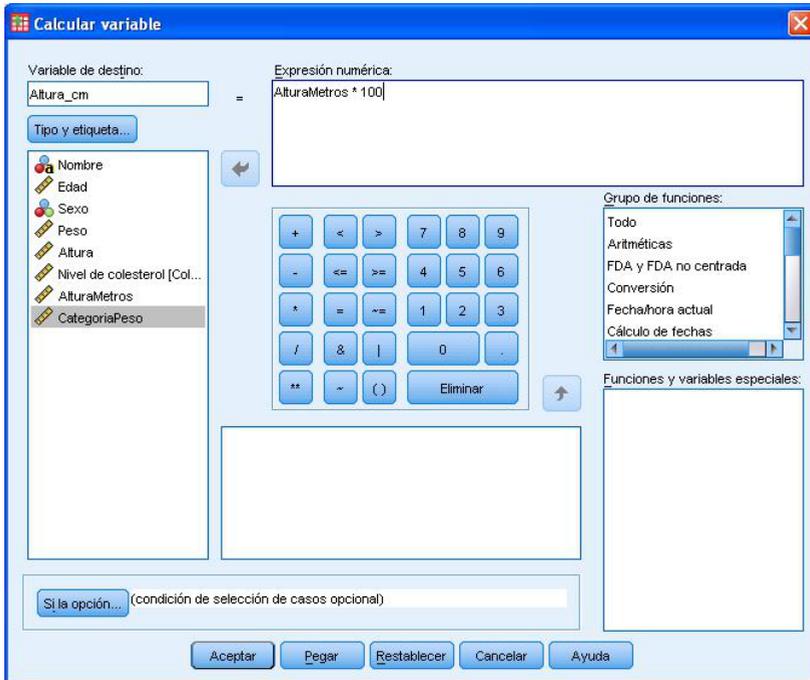


Figura 1.6: Ventana de transformación de variables. A la izquierda aparecen las variables ya definidas, a la derecha las funciones predefinidas que pueden utilizarse, y en el centro los operadores aritméticos y relacionales más comunes.

lista de variables ya definidas que pueden utilizarse como argumentos de las distintas funciones de transformación.

Los operadores más habituales para construir expresiones son los aritméticos +, -, *, /, ** (potenciación), los relacionales =, <, >, ~=, <=, >= y los lógicos & (Y), | (O) y ~ (negación). Y algunas de las funciones más habituales son: ABS (valor absoluto), SQRT (raíz cuadrada), EXP (exponencial), LN (logaritmo neperiano), SIN (seno), COS (coseno), TAN (tangente), SUM (suma), MEAN (media aritmética), SD (desviación estándar), RND (redondeo al entero más cercano), TRUNC (parte entera de un número).

Haciendo click en el botón Si la opción... se pueden establecer condiciones de aplicación de la transformación. Para establecer una condición debemos activar la opción Incluir si el caso satisface la condición y después introducir

una condición lógica como por ejemplo $Sexo=1$. De este modo, la transformación sólo se aplicará a los individuos que cumplan dicha condición. Una vez definida la expresión hay que hacer click sobre el botón Aceptar y automáticamente aparecerá en la vista de datos una nueva columna con los datos transformados de la nueva variable.

1.2.7. Recodificación de datos

Otra forma de transformar una variable es crear otra cuyos valores sean una recodificación de los de la primera, por ejemplo agrupando en intervalos. Esta recodificación podemos hacerla tanto en la misma variable como en variables diferentes. Para recodificar en variables diferentes se selecciona el menú Transformar ▶ Recodificar en distintas variables. Automáticamente aparece la ventana de recodificación de variables tal y como se ve en la figura 1.7.

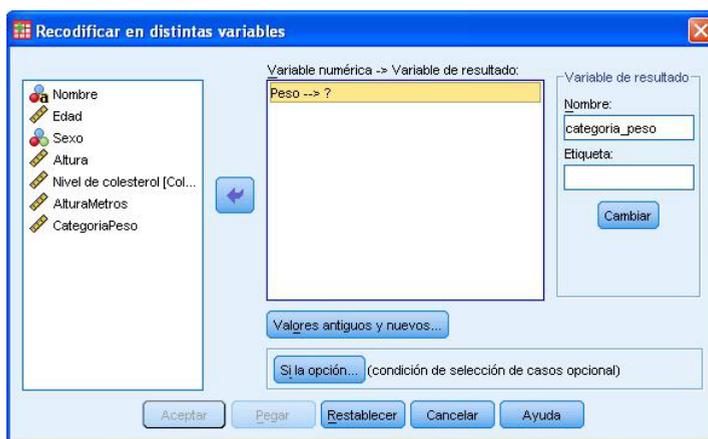


Figura 1.7: Ventana de recodificación de variables. A la izquierda aparecen las variables ya definidas, a la derecha deben especificarse las reglas de recodificación.

Para recodificar una variable en otra nueva, primero debemos seleccionar la variable que queremos recodificar y hacer click sobre el botón con una flecha que aparece al lado. Después hay que escribir el nombre de la nueva variable en el cuadro Nombre y hacer click sobre el botón Cambiar. A continuación hay que establecer las reglas de recodificación. Para ello hay que hacer

click en el botón Valores antiguos y nuevos para que aparezca la ventana de definición de reglas (figura 1.8). Las reglas pueden establecer la conversión del valor de la variable original que introduzcamos en el cuadro Valor antiguo en el valor de la variable nueva que introduzcamos en el cuadro Valor nuevo, o bien la conversión de todo un intervalo de valores de la variable original en un valor de la variable nueva. Una vez definidos dichos valores hay que hacer click sobre el botón Continuar, y después sobre Aceptar.

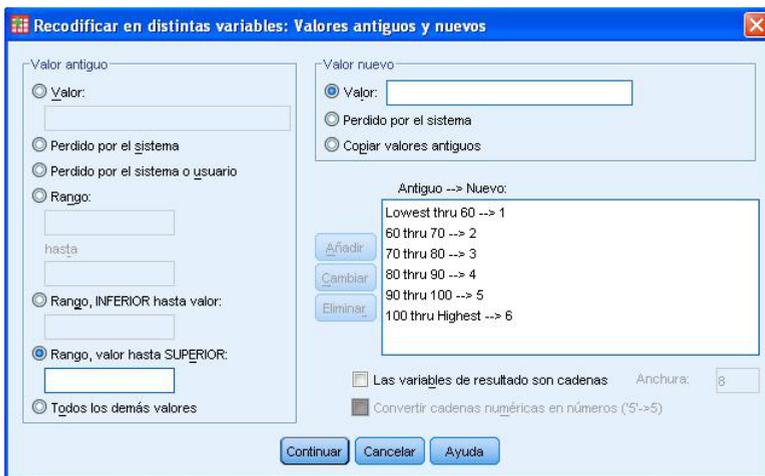


Figura 1.8: Ventana de definición de reglas.

1.2.8. Impresión

Para imprimir se utiliza el menú Archivo ▶ Imprimir. Al instante aparece un cuadro de diálogo para la impresión donde debemos indicar si queremos imprimir todo o bien la selección que hayamos hecho. Tras esto se hace click sobre el botón Aceptar y la información se envía a la impresora.

Antes de imprimir conviene hacer una previsualización de lo que se va a enviar a la impresora para estar seguros de que es eso lo que se quiere. Para ello se utiliza el menú Archivo ▶ Presentación preliminar. Entonces aparece un visor donde se ve la página, tal y como se enviará a la impresora. Si todo parece correcto se puede hacer click sobre el botón Imprimir y aparecerá el cuadro de diálogo de impresión desde el que se puede enviar a la impresora definitivamente.

1.2.9. Salir del programa

Para terminar una sesión de trabajo se utiliza el menú Archivo ▶ Salir, o bien se hace click sobre el aspa para cerrar la ventana del programa. Si quedan datos o resultados que no se han guardado, el programa nos preguntará antes de salir si deseamos guardarlos.

1.2.10. Ayuda

En esta práctica sólo hemos descrito las operaciones básicas en una sesión de trabajo. Pero quedan por describir todos los análisis estadísticos que pueden realizarse con los menús de la barra de menús. Aunque muchos de estos menús se explicarán en las siguientes prácticas, el programa dispone del menú de ayuda Ayuda en el que podemos encontrar una descripción de todos estos menús y al que podemos recurrir cada vez que tengamos dudas.

1.3. Ejercicios resueltos

1. Introducir en la matriz de datos los datos de la siguiente muestra y guardarlos en un fichero con el nombre `datos_colesterol.sav`.

Nombre	Sexo	Peso	Altura	Colesterol
José Luis Martínez Izquierdo	H	85	179	182
Rosa Díaz Díaz	M	65	173	232
Javier García Sánchez	H	71	181	191
Carmen López Pinzón	M	65	170	200
Marisa López Collado	M	51	158	148
Antonio Ruiz Cruz	H	66	174	249
Antonio Fernández Ocaña	H	62	172	276
Pilar Martín González	M	60	166	213
Pedro Gálvez Tenorio	H	90	194	241
Santiago Reillo Manzano	H	75	185	280
Macarena Álvarez Luna	M	55	162	262
José María de la Guía Sanz	H	78	187	198
Miguel Angel Cuadrado Gutiérrez	H	109	198	210
Carolina Rubio Moreno	M	61	177	194

i

- a) En la ventana de Vista de variables, crear las variables Nombre, Sexo, Peso, Altura y Colesterol e introducir los datos anteriores, siguiendo las indicaciones del apartado 1.2.2.
- b) Una vez introducidos los datos, se guardan en un fichero de nombre datos_colesterol. siguiendo lo indicado en el apartado 1.2.3.

2. Sobre la matriz de datos del ejercicio anterior realizar las siguientes operaciones:

- a) Insertar detrás de la variable Nombre una nueva variable Edad con las edades de todos los individuos de la muestra.

Nombre	Edad
José Luis Martínez Izquierdo	18
Rosa Díaz Díaz	32
Javier García Sánchez	24
Carmen López Pinzón	35
Marisa López Collado	46
Antonio Ruiz Cruz	68
Antonio Fernández Ocaña	51
Pilar Martín González	22
Pedro Gálvez Tenorio	35
Santiago Reillo Manzano	46
Macarena Álvarez Luna	53
José María de la Guía Sanz	58
Miguel Angel Cuadrado Gutiérrez	27
Carolina Rubio Moreno	20

i

- 1) En la Vista de variables seleccionar la fila correspondiente a la variable Sexo haciendo click con el ratón sobre la cabecera de la misma y a continuación seleccionar el menú Edición ▶ Insertar variable, con lo que aparece una nueva fila entre las variables Nombre y Sexo.
- 2) En la nueva fila definir la variable Edad.
- 3) En la Vista de datos rellenar los datos de la columna correspondiente a la Edad.