

ANÁLISIS DE DATOS EN PSICOLOGÍA I: CUADERNO DE PRÁCTICAS DE SPSS

Carmen Ximénez Gómez

Con la colaboración de los profesores de la asignatura:

**Juan Botella Ausina
Ludgerio Espinosa Gil
Manuel Suero Suñe**

**Asignatura Análisis de datos I
Universidad Autónoma de Madrid.
Curso 2005-2006**

El objetivo de estas prácticas es familiarizar a los alumnos con las herramientas informáticas que harán posible la aplicación de las técnicas de análisis de tipo descriptivo que han aprendido en la asignatura *Análisis de datos en Psicología I*.

Como se ha podido comprobar en la resolución de problemas en clase, los cálculos a mano e incluso con calculadora son muy costosos cuando se manejan amplias muestras de sujetos. De ahí la necesidad de utilizar computadoras que los agilicen y nos permitan la interpretación de los resultados, que es lo que más nos interesa. Existen diversos programas informáticos que pueden emplearse para analizar datos. Por ejemplo, el SPSS, el SAS, el STATGRAPHICS, el SYSTAT o incluso una hoja de cálculo de EXCEL. Se ha seleccionado el programa SPSS porque es uno de los más conocidos y empleados en la práctica profesional y porque se encuentra disponible en las aulas de informática de la Facultad.

En un primer momento se pretende que el alumno se familiarice con la estructura y las principales funciones que el SPSS ofrece. En el tema 1 trataremos de cubrir este objetivo. No obstante, el verdadero propósito de estas prácticas es que el alumno aprenda a realizar los principales análisis estudiados en clase mediante el ordenador. En el tema 2 se resumen los análisis descriptivos con una variable (distribuciones de frecuencias, índices de tendencia central, de variabilidad, etc.) y en el tema 3 los de dos variables (índices de asociación lineal, combinación lineal de variables y análisis de regresión lineal). Por último, se pretende que el alumno sepa cómo interpretar los resultados obtenidos a partir del SPSS. Para ello, no es suficiente conocer cómo funciona el programa sino que se precisa una comprensión de los contenidos explicados en clase.

Las prácticas consisten en analizar los datos reales correspondientes a 200 candidatos a un puesto de trabajo que respondieron un cuestionario (las 16 variables del cuestionario se describen en el anexo II). A partir de estos datos, realizaremos diferentes análisis descriptivos y extraeremos conclusiones sobre la muestra. Los datos se incluyen en el anexo III y se proporcionarán en soporte informático (fichero **practica.sav** que el alumno habrá de descargar de la página web de la asignatura y grabarlo en un disquete 3 ½, que traerá a cada clase). En las clases prácticas sólo realizaremos un tipo concreto de análisis, a modo de ejemplo, pero el alumno debe acudir al aula de informática por su cuenta y experimentar nuevos análisis con los datos. Con este fin se ha incluido un ejercicio práctico integrado en el anexo IV, que se resuelve con otro fichero diferente (**ejercicios.sav** que también se puede descargar de la página web de la asignatura).

Para realizar las prácticas es necesario un conocimiento elemental del manejo del ordenador: copiar archivos, disquetes, entorno Windows, entrar y salir de los programas, etc. Debido al amplio tamaño de los grupos y las limitaciones de espacio en las aulas de informática, para impartir las clases prácticas cada grupo de alumnos se desdoblará en dos subgrupos (por ejemplo el grupo 12 se desdoblará en el subgrupo 12A y el 12B). Es importante que cada alumno sepa a qué subgrupo pertenece y dónde y cuándo recibirá cada clase práctica. En el tablón de anuncios de cada clase se incluirá dicha información. Asimismo, los alumnos pueden conectarse a la página http://www.uam.es/personal_pdi/psicologia/carmenx/practicas.html donde encontrarán dicho listado y toda la información relativa a las prácticas (ficheros de datos, ejercicios, etc.).

Las prácticas influirán sobre la calificación final de la asignatura. La parte práctica del examen de *Análisis de datos en Psicología I* (calificada sobre 7 puntos) se puntuará junto a una práctica en la que el alumno/a podrá obtener de 0 a 3 puntos.

El aprendizaje del SPSS no termina en este curso sino que tendrá continuidad en otras asignaturas de segundo curso: *Análisis de datos en Psicología II* e *Introducción a la Psicometría*. Asimismo, muchas de las prácticas de otras asignaturas requieren el manejo de datos con el SPSS. Por tanto, todo lo aprendido será necesario para una comprensión adecuada de éstas.

ÍNDICE

Tema 1. Introducción al SPSS

1.1. Estructura del SPSS	1
El editor de datos	1
El visor de resultados	2
1.2. Barra de menús y de herramientas	2
Menús generales	2
Menú <i>Analizar</i>	3
Menú <i>Ayuda</i>	3
1.3. Cómo manejar archivos de datos	4
Elaborar archivos nuevos	4
Definir variable	4
Guardar	6
Abrir archivos de datos existentes	6

Tema 2. Análisis descriptivos con una variable

2.1. Distribuciones de frecuencias	8
Tablas de frecuencias	8
Representaciones gráficas	10
2.2. Medidas de posición y de tendencia central	13
Los centiles, los deciles y los cuartiles	13
Media, mediana y moda	13
2.3. Medidas de variación	15
Varianza y desviación típica	15
Diagramas de cajas	16
2.4. Puntuaciones típicas y escalas derivadas	16
2.5. Asimetría y curtosis	19
Ejercicios	20

Tema 3. Análisis descriptivos con más de una variable

3.1. Índices de asociación lineal	21
La matriz de varianzas-covarianzas	21
La matriz de correlaciones	22
3.2. Combinación lineal de variables	23
3.3. Análisis de regresión lineal simple	25
Diagrama de dispersión	25
Ecuación de regresión	26
Bondad de ajuste del modelo	30
3.4. Análisis de regresión lineal múltiple	34
3.5. Distribuciones de frecuencias y representaciones gráficas	36
Ejercicios	40

ANEXOS:

ANEXO I: Notación científica	41
ANEXO II: Cuestionario	42
ANEXO III: Datos para las prácticas	43
ANEXO IV: Ejercicio práctico	49

Tema 1. Introducción al SPSS

El objetivo de este primer tema es que el alumno aprenda a manejarse en el programa SPSS dentro del entorno Windows. Para ello, primero hablaremos de su estructura y de sus principales ventanas (el editor de datos y el visor de resultados). Más tarde, de las barras de menús y de herramientas y de las funciones que realizan. Por último, veremos cómo se manejan los archivos de datos.

1.1. Estructura del SPSS

Antes de entrar en aspectos técnicos, conviene que el alumno tenga claro que el SPSS realiza tres funciones esenciales: (1) leer datos de n sujetos en J variables, (2) especificar el tipo de análisis que se quiere realizar con esos datos (estadísticos de tendencia central, de variabilidad, etc.) y (3) mostrar los resultados obtenidos tras los análisis. Lo único que no hace el SPSS es darnos la interpretación de los resultados. Ésa es tarea del investigador. Como veremos a continuación cada una de estas funciones aparece en una ventana diferente.

El primer paso para acceder al SPSS es abrir el programa. Para ello hay que entrar en Windows y pulsar dos veces con el botón izquierdo del ratón sobre el siguiente icono:



El editor de datos

La primera pantalla que aparece al entrar en el SPSS es la ventana del *editor de datos*, la cual ofrece el siguiente aspecto:

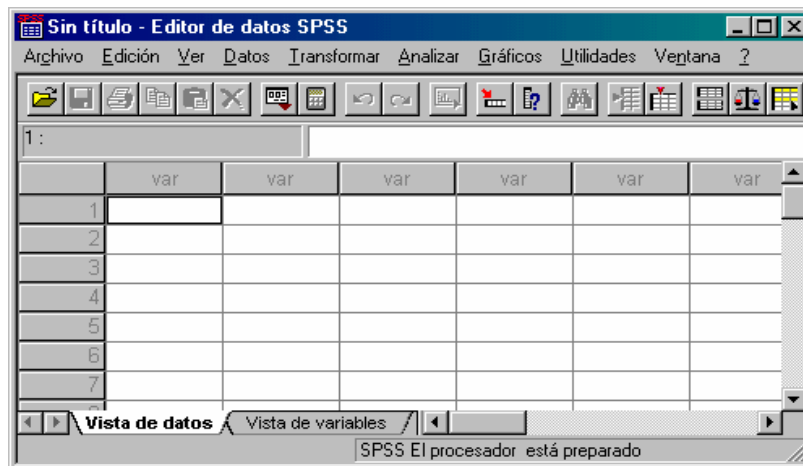



Figura 1.1.

Esta ventana es donde deben encontrarse los datos empleados en el análisis. Pueden darse varias situaciones, aunque nosotros solo veamos dos:

1. Los datos fueron introducidos previamente; probablemente en una sesión anterior de trabajo y sólo es necesario recuperar el archivo. Para ello se selecciona el menú Archivo o se pulsa en el icono .
2. Los datos no fueron introducidos previamente y queremos hacerlo. La forma de introducir los datos es muy similar a las tablas de datos que hemos empleado en clase. Como se observa en la figura 1.1., la ventana *editor de datos* contiene una tabla donde las filas se refieren a los n sujetos u observaciones de la muestra y las columnas a las J variables X_1, X_2, \dots, X_J . El resultado final es una matriz de n sujetos $\times J$ variables.

* Si el icono no está disponible en el Escritorio, pulsar en 'Inicio -> Programas -> SPSS' o buscar su ubicación con "el Explorador de Windows".

El Visor de resultados

Una vez se solicite un análisis con los datos que se encuentran en *el editor de datos*, los resultados obtenidos se muestran en la ventana *visor de resultados*, cuyo aspecto es el siguiente:



Figura 1.2.

En la parte derecha se muestran los resultados obtenidos (tablas, gráficos, etc.) y en la parte izquierda (donde dice Resultados) aparece una lista con los títulos de los contenidos mostrados en la ventana *Visor de resultados*.^{*} Si se desea pasar de la ventana de resultados a la de datos se selecciona con el ratón la barra de botones de Windows que aparece señalada en la figura 1.2. en la parte inferior de la pantalla.

1.2. Barra de menús y de herramientas

Como se observa en las figuras 1.1. y 1.2., en cada una de las ventanas descritas aparecen dos barras en la parte superior. Por ejemplo, en la ventana *Editor de datos* su aspecto es el siguiente:^{**}



La barra de la parte superior consta de una serie de menús (Archivo, Edición, Ver, Datos, etc.). Si se selecciona con el ratón cada una de ellas aparece un menú desplegable donde se ofrecen otros sub-menús, cada uno de los cuales tiene a su vez un ‘cuadro de diálogo’. Algunos de estos sub-menús (los más utilizados) pueden definirse en la barra inferior. En lugar de menús, ésta consta de una serie de botones cada uno de ellos con un icono dibujado (p.e. : abrir archivo; : guardar archivo; etc.) y se denomina barra de herramientas.

Menús generales

Entre los menús de ‘la barra de menús’ se encuentran los que son comunes a todos los programas de tipo Windows: Archivar, Editar, Ver, Ventana, etc. En el SPSS, los que ofrecen mayores diferencias son los menús ‘Datos’, ‘Transformar’ y ‘Analizar’, los cuales permiten realizar operaciones con los datos. Dichos menús serán revisados en detalle en los temas 2 y 3.

* Para ver ejemplo ir a páginas 9 y 10.

** Cada ventana tiene su propia barra de menús y de herramientas, aunque son bastante similares.

Menú *Analizar*

Este es el menú más importante de todos. Si lo seleccionamos, aparece lo siguiente:

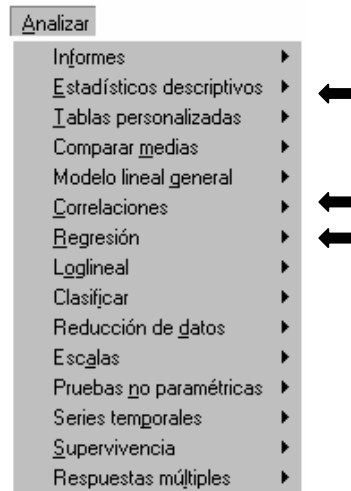


Figura 1.3.

Decimos que es el menú más importante porque desde aquí es donde indicaremos al SPSS el tipo de análisis que queremos realizar con los datos. Cada fila (terminada en una punta de flecha: ▶) ofrece a su vez otros menús, cada uno con su correspondiente *cuadro de diálogo* que es el lugar donde se especifica el número de variables que entran en el análisis y otra serie de detalles que veremos más adelante.

En estas prácticas sólo veremos los menús que se corresponden con las técnicas estudiadas en la asignatura *Análisis de datos en Psicología I*. En concreto, los menús ‘Estadísticos descriptivos’, ‘Correlaciones’ y ‘Regresión’. En las asignaturas troncales de segundo curso: *Análisis de datos en Psicología II* e *Introducción a la Psicometría*, se verán muchos otros de los restantes menús. Y en la asignatura optativa *Proceso de datos* se profundiza mucho más en el manejo del programa.

Menú *Ayuda*

Por último, el menú que aparece con la expresión ? se refiere a la ayuda que ofrece el programa y que será muy útil en todo momento para el usuario. Puede accederse a él desde la ventana *Editor de datos* y *Visor de resultados* y en ambos casos ofrece el siguiente aspecto:

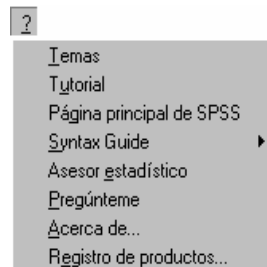


Figura 1.4.

Desde ‘Temas’ puede escribirse un tipo de análisis (por ejemplo ‘Frecuencias’) y el SPSS muestra una descripción de sus opciones. También es posible obtener ayuda pulsando el botón derecho del ratón. Por ejemplo, si nos encontramos en el cuadro de diálogo en el que se calculan los estadísticos descriptivos y se quiere saber la definición de ‘curtosis’, basta con seleccionarlo y pulsar el botón derecho del ratón.

1.3. Cómo manejar archivos de datos

En este apartado aprenderemos a manejar la ventana *Editor datos* y a leer y almacenar en un archivo los datos necesarios para realizar un análisis. El objetivo que se persigue es que el alumno aprenda estas operaciones para finalmente saber manejar los datos de la práctica.

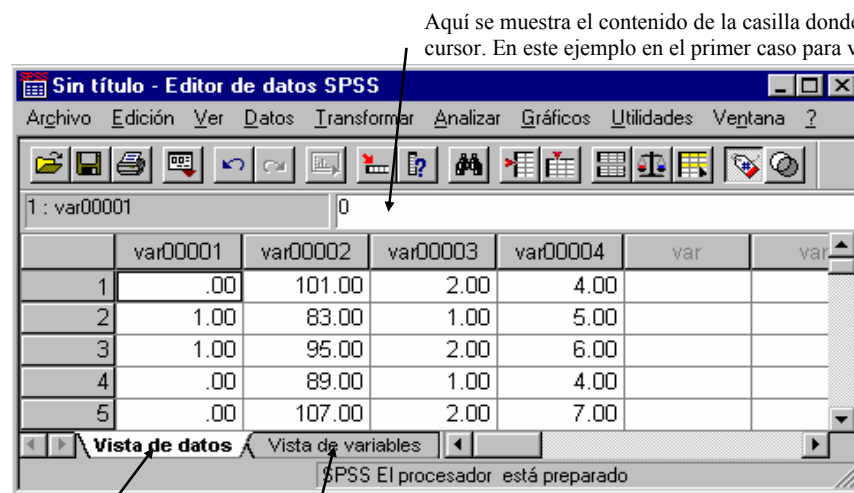
Los datos de la práctica se corresponden con un cuestionario que fue administrado por un equipo de investigación en psicología diferencial a 200 candidatos a un puesto de trabajo en una empresa. Dicho cuestionario consta de 16 preguntas. Cada una de ellas constituye una variable medida en diferentes niveles. Por ejemplo, la primera variable, *sexo*, está medida a nivel nominal y la cuarta, *estatura*, a nivel de razón (ver anexo II para una descripción detallada de cada variable). Aquí trabajaremos con las respuestas correspondientes a una muestra de 200 sujetos. Por tanto, la interpretación de los análisis realizados se referirá tan solo a esta muestra de sujetos. Los datos empleados aparecen en el anexo III. Puesto que constan de 200 casos y su grabación sería muy costosa, dichos datos serán facilitados en un archivo (*practica.sav*) que se puede descargar desde la dirección: http://www.uam.es/personal_pdi/psicologia/carmenx/practicas.html. Es tarea del alumno descargar este archivo, grabarlo en un disquete de 3 ½ y traerlo a cada una de las clases prácticas.

Elaborar archivos nuevos

Cuando recogemos datos de n sujetos en J variables, lo primero que hay que hacer es introducirlos en el SPSS. Para ello hay que situarse en la ventana *Editor de datos* de la figura 1.1. Esta ventana contiene una matriz de n sujetos \times J variables. Veamos cómo se introducen datos a partir de un ejemplo práctico. Supongamos que tenemos los siguientes datos:

Sujeto	Sexo	Inteligencia	Nivel cultural	Estrés
1	0	101	2	4
2	1	83	1	5
3	1	95	2	6
4	0	89	1	4
5	0	107	2	7

Los datos corresponden a 5 sujetos medidos en cuatro variables: *Sexo* (0: mujer y 1: varón), nivel cultural (1: bajo y 2: alto), *inteligencia* y *estrés*. Por tanto, tenemos que elaborar una matriz de 5 filas por cuatro columnas. Los datos se introducen colocando un sujeto en cada fila y una variable en cada columna. Para ello, simplemente se escriben las puntuaciones correspondientes a cada caso. Para moverse de una celda a otra se pulsan las teclas marcadas con flechas. En nuestro ejemplo, los datos de los 5 primeros sujetos se colocan del siguiente modo:



Pulsar para ver
los datos

Pulsar para ver la
definición de las variables

Una vez introducidos los datos es necesario dar nombre a las variables. Como se observa en la figura superior, el SPSS ha dado por defecto los nombres *var00001*, *var00002*, *var00003* y *var00004*. Asimismo, ha escrito los valores numéricos con dos decimales.

Si se desea cambiar el nombre de una variable y definir sus características (por ejemplo nivel de medida, etiquetas de valores, formato de columna, definición de valores perdidos, etc.) hay que pulsar con el ratón en la solapa “Vista de variables”. A continuación aparece una ventana que contiene tantas filas como variables haya definidas en el editor de datos:

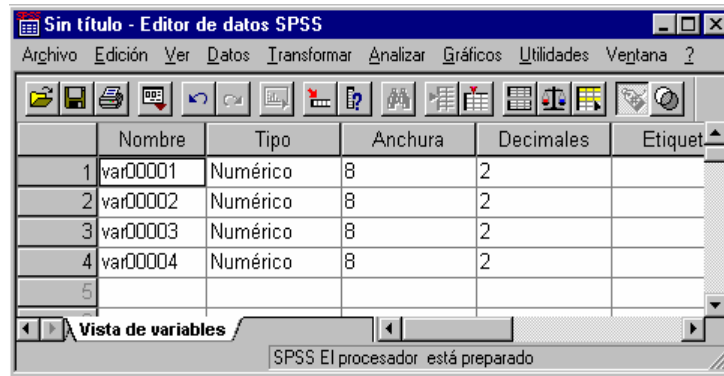
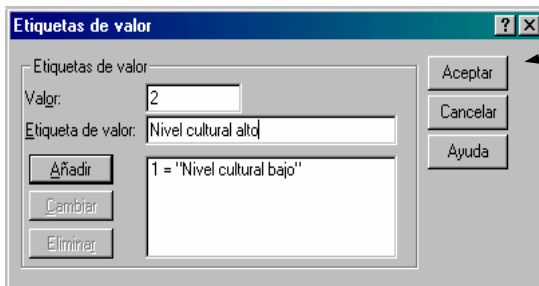


Figura 1.5.

En cada fila se da la posibilidad de: modificar el **nombre** de la variable (teniendo en cuenta que no puede tener más de 8 caracteres de largo) además de otras propiedades como el **tipo** (numérico, cadena, etc.), **la anchura**, el número de **decimales**, **la etiqueta**, el significado de los **valores** que admite, los valores **perdidos**, la **alineación** del texto, etc. Basta posicionar el ratón o el cursor en cada casilla y escribir la información deseada. En nuestro ejemplo, llamemos a las variables: *sexo*, *intelig*, *nc*, y *estres*. Algunas de las casillas de la ventana ‘Vista de variables’ incorporan su propio cuadro de diálogo que se activa al seleccionarlás. Por ejemplo, al pulsar en ‘Valores’ aparece lo siguiente:



Desde este cuadro de diálogo se puede dar etiqueta a los valores de una variable. Por ejemplo en *nc* el valor 1 significa nivel cultural bajo y el valor 2, nivel cultural alto. Cada vez que se introduzca la etiqueta de un valor se pulsa la tecla *Añadir*. Para terminar se pulsa en *Aceptar*.

Una vez definidas todas las características de las variables, para volver a visualizar los datos numéricos se pincha en la solapa “Vista de datos”. El resultado para nuestro ejemplo es el siguiente:

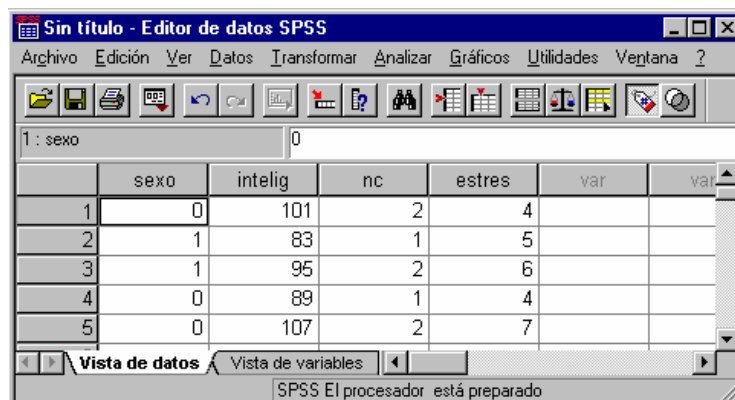


Figura 1.6.

Guardar

Para guardar la información introducida en el editor de datos se selecciona el menú *Archivo* de la barra de herramientas, se elige ‘Guardar’ y aparece el correspondiente cuadro de diálogo:

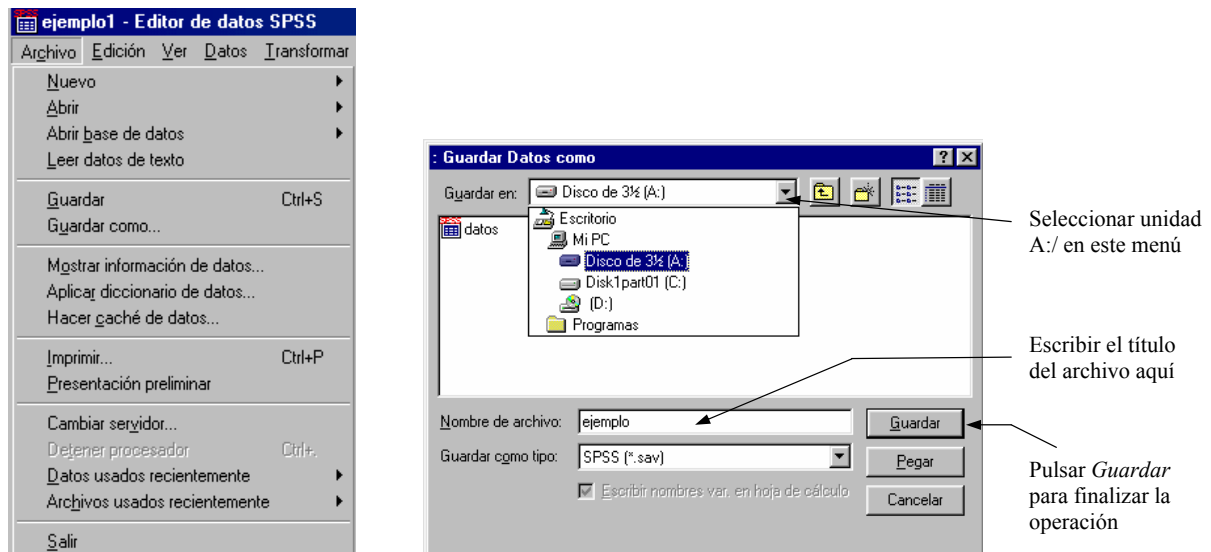


Figura 1.7.

El SPSS guarda los archivos de datos con la extensión "*sav*" por tanto basta con dar un nombre a los datos (*ejemplo*) y por defecto se grabarán en un archivo con la extensión *sav* (aquí le hemos llamado: *a:\ejemplo.sav*). También es posible guardar archivos de resultados. Para ello nos situamos en la ventana *Visor de resultados* seleccionamos el menú 'Guardar' de la figura 1.2. y aparece un cuadro de diálogo idéntico al de la figura 1.7. El procedimiento a seguir es exacto al de archivar datos con la única diferencia de que la extensión de este tipo de archivos es "*spo*". Por tanto, los archivos SPSS con la extensión "*sav*" contienen datos y los de la extensión "*spo*" resultados.

Veamos ahora cómo se procede en el caso de no tener introducidos los datos de las prácticas. Aquí trabajaremos con un fichero que contiene las respuestas de 200 candidatos a un puesto de trabajo a un cuestionario de 16 preguntas (ver anexo II para una descripción de cada variable). Por tanto, tenemos las respuestas de 200 sujetos en 16 variables. Para introducirlas en el SPSS hay que situarse en la ventana *Editor de datos* y elaborar una matriz de 200 filas x 16 columnas. En este caso la ventana 'Vista de variables' tiene 16 filas, una para cada variable considerada.* En cada fila existe la posibilidad de modificar sus características (nombre, etiqueta, decimales, valores, etc.). El resultado obtenido es igual al de la tabla de datos que se incluye en el anexo III.

Abrir archivos de datos existentes

Cuando el archivo de datos se encuentra ya introducido en el ordenador o en un disquete, bastará simplemente con abrirlo. Para ello, se selecciona el menú 'Abrir' de la figura 1.7. y la unidad donde se encuentra el archivo. En la figura 1.8. aparece el cuadro de diálogo correspondiente a 'Abrir'.

Nuestros datos se encuentran grabados en el archivo *a:\practica.sav*. Para acceder a ellos los seleccionamos y pulsamos en el botón *abrir*.

* En este caso las variables son: sexo, edad, peso, estatura, prov, idprov, rama, licen, inteli, compren, orient, extra, respon, emocio, sincer y fumar.

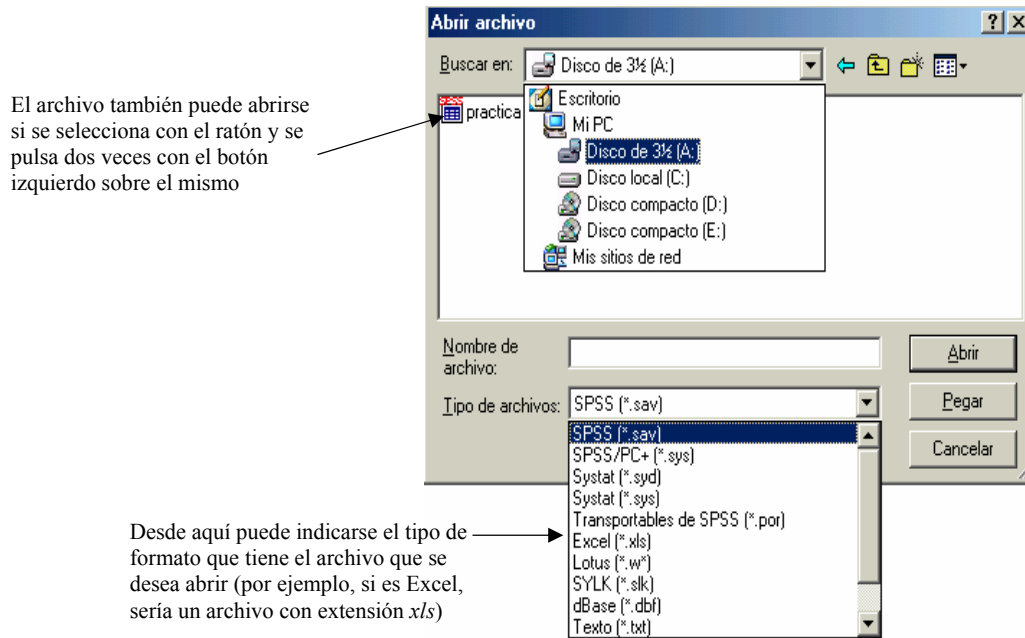


Figura 1.8.

Una vez completada esta operación aparecen los datos en el *Editor de datos*. A continuación se muestra un ejemplo para los 5 primeros sujetos en las siete primeras variables:*

	sexo	edad	peso	estatura	prov	idprov	rama
1	0	34	46.00	1.63	MADRID	46	1
2	0	34	51.00	1.72	BURGOS	22	4
3	1	25	45.00	1.53	VALLADOLID	28	1
4	1	26	64.00	1.75	PALENCIA	24	1
5	1	27	58.00	1.68	MADRID	46	2

Figura 1.9.

Siempre que se desee iniciar una sesión con el SPSS en la que se lleve a cabo uno o más análisis de datos, es necesario que el archivo donde se encuentran los datos a analizar esté abierto en el *Editor de datos* del SPSS. Como veremos más adelante, los menús donde se encuentran las diferentes opciones de análisis no están activos a no ser que el editor de datos tenga algún tipo de contenido.

Para que el SPSS realice algún tipo de análisis estadístico, no basta con que los datos estén grabados en un archivo, es necesario que dicho archivo esté abierto.

* El SPSS es capaz de leer datos no sólo de tipo *.sav (el formato de SPSS) sino también de otros formatos, por ejemplo archivos de tipo EXCEL. Para abrir un archivo de tipo EXCEL (por ejemplo, el fichero practica.xls) se selecciona la opción "Tipo de archivos" de la figura 1.8 y "Excel (*.xls)". Sin embargo los ficheros de datos de tipo *.sav sólo pueden leerse con el programa SPSS.

Tema 2. Análisis descriptivos con una variable

En este tema aprenderemos a llevar a cabo los análisis descriptivos con una variable revisados en la primera parte de la asignatura *Análisis de datos en Psicología I*. Todos los análisis se realizan desde el menú *Analizar -> Estadísticos descriptivos*:

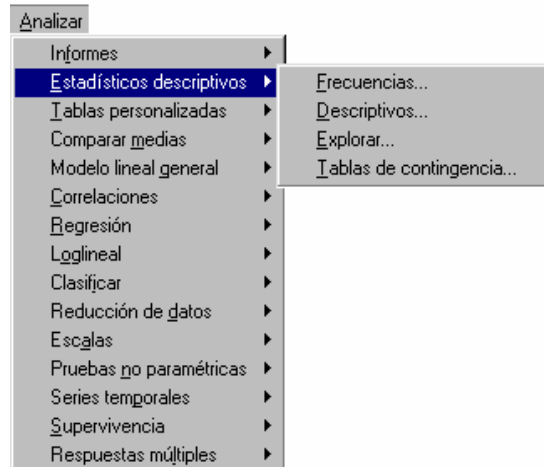


Figura 2.1.

Para realizar análisis descriptivos univariados basta con los tres primeros procedimientos ('*Frecuencias*', '*Descriptivos*' y '*Explorar*'). A continuación se presenta una explicación detallada del cuadro de diálogo correspondiente a cada una de ellos. Para ello, seguiremos el mismo orden de contenidos que el que aparece en el libro de la asignatura.

2.1. Distribuciones de frecuencias

Con el SPSS pueden confeccionarse distribuciones de frecuencias para organizar los datos correspondientes a una variable. En este apartado veremos cómo se construyen las tablas de frecuencias y cómo se realizan algunas de las representaciones gráficas vistas en clase. Ambas cuestiones se llevan a cabo desde el cuadro de diálogo correspondiente al procedimiento '*Frecuencias*' de la figura 2.1.

Tablas de frecuencias

El cuadro de diálogo '*Frecuencias*' ofrece el siguiente aspecto:

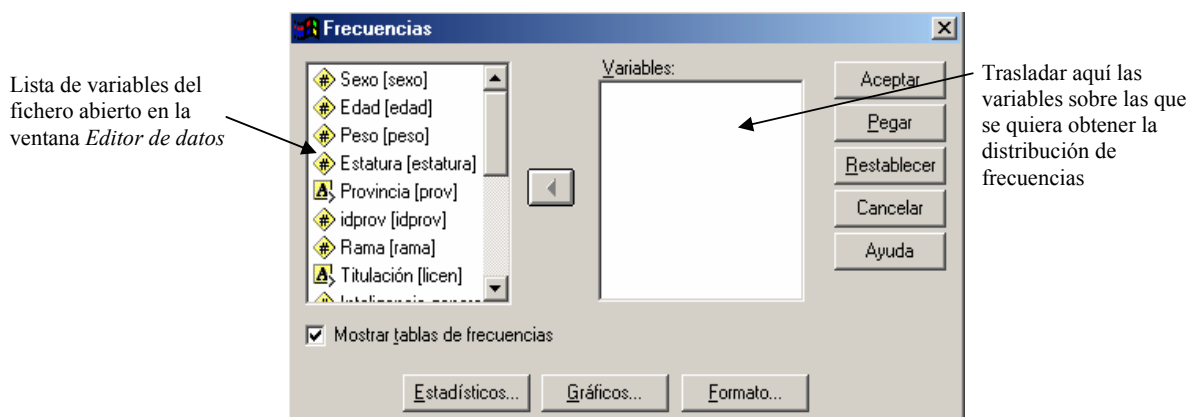



Figura 2.2.

En el cuadro izquierdo aparece el listado de variables de que consta el archivo de datos que está abierto. El cuadro derecho está vacío pues está esperando a que le indiquemos qué variables queremos analizar. Dichas variables se trasladan de un cuadro a otro mediante el botón . Por defecto, la opción ‘mostrar tablas de frecuencias’ está activada pero si no queremos que aparezca puede desactivarse. Veamos un ejemplo con la variable ‘Edad’. Para ello hay que trasladarla al cuadro de la derecha de la figura 2.2. y pulsar en el botón *Aceptar* para que se ejecute la orden. Los resultados obtenidos aparecen en el *Visor de resultados*:

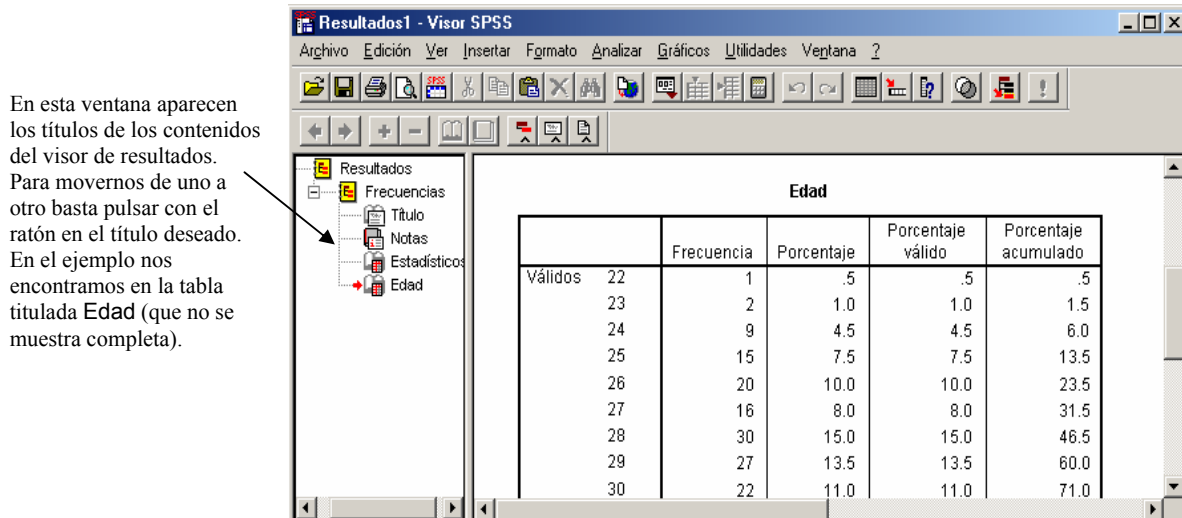


Figura 2.3.

La tabla de frecuencias tiene el mismo aspecto que las que hemos empleado en clase. En la primera columna aparecen los valores de la variable en sentido ascendente (de 22 a 46); en la segunda ('Frecuencia') sus frecuencias absolutas: n_i ; en la tercera ('Porcentaje') las frecuencias relativas en porcentaje: P_i ; y en la última ('Porcentaje acumulado') las relativas acumuladas en porcentaje: P_a .

Representaciones gráficas

Desde el cuadro de diálogo de la figura 2.2. también es posible pedir al SPSS que realice tres tipos de gráficos para observar las frecuencias de los datos:

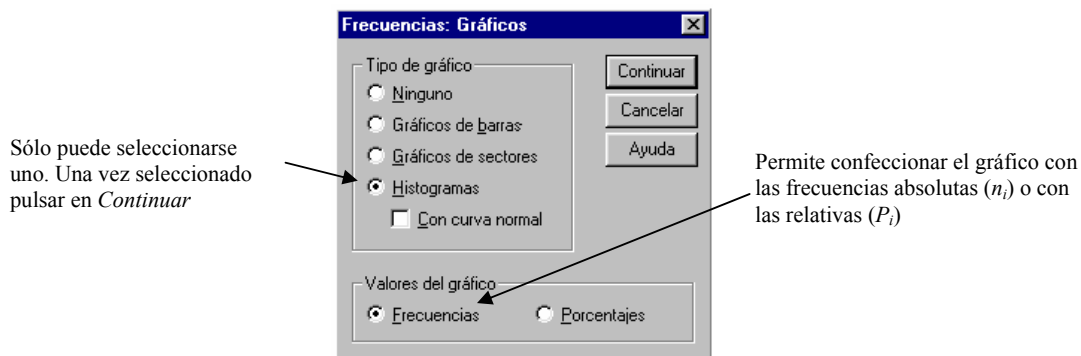


Figura 2.4.

Como se observa, se puede obtener un histograma (con el ajuste a la curva normal), un diagrama de barras y un diagrama de sectores (o pictograma). Desde este cuadro de diálogo sólo es posible seleccionar uno de los gráficos. Si se quisiera obtener los tres hay que repetir la operación tres veces. Seguimos con el ejemplo de la variable ‘Edad’. Vamos a pedir al SPSS que nos muestre los tres tipos de gráficos de la distribución de frecuencias para la *edad*.* Si pulsamos en ‘continuar’ volvemos automáticamente al menú de la figura 2.2.

* En realidad puesto que la *Edad* es una variable cuantitativa el gráfico más adecuado es el histograma pero aquí veremos todos “a modo de ejemplo”.

Para ir de un resultado a otro puede utilizarse la parte izquierda de la pantalla, donde aparece el listado de resultados obtenidos (en el ejemplo: ‘EDAD’ e ‘Histograma’). Situándose con el ratón en cada uno de ellos puede accederse automáticamente al resultado correspondiente. En la figura 2.5. se muestra el histograma para la variable edad.*

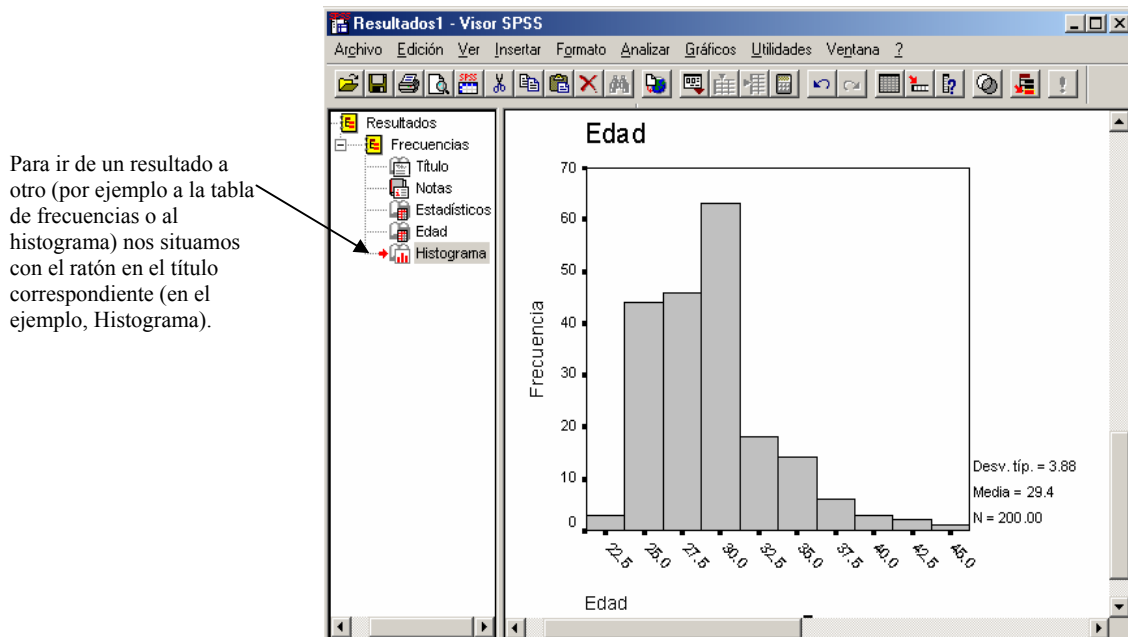


Figura 2.5.

Si se pide el diagrama de barras y de sectores los resultados son los siguientes:

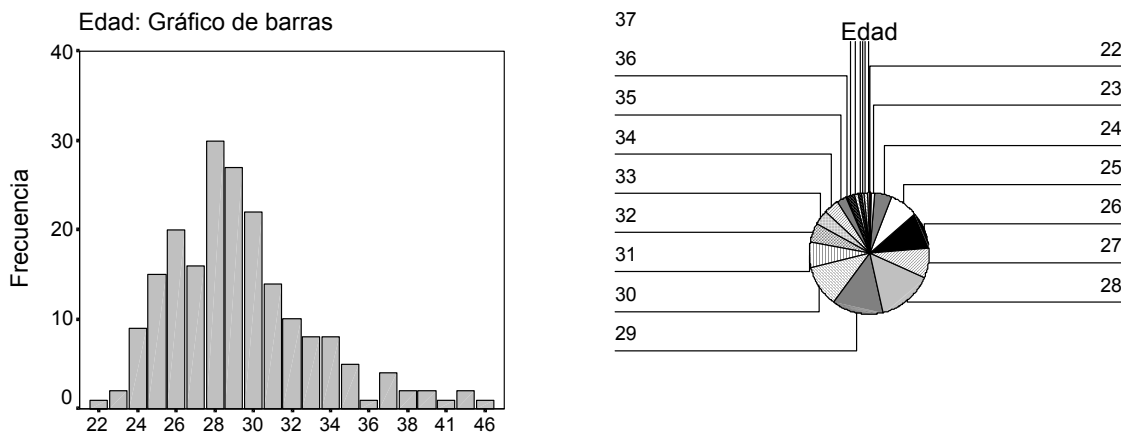


Figura 2.6.

Como se observa, no es posible elaborar el gráfico de sectores, pues al ser una variable cuantitativa, hay demasiadas categorías. En todas las gráficas la mayor parte de los sujetos tienen entre 25 y 30 años, estando el promedio en 29 años y la moda en 28. Sin embargo, el 29% de los sujetos superan los 30 años (*Ejercicio*: realizar cada uno de los gráficos que mejor describa a las variables: ‘sexo’, ‘estatura’, ‘licen’ e ‘inteli’).

Otra de las representaciones que hemos aprendido para reflejar distribuciones de frecuencias es el **diagrama de tallo y hojas**. Para confeccionarlo se utiliza el menú *Analizar -> Estadísticos descriptivos -> Explorar* (ver figura 2.1.), cuyo cuadro de diálogo es el siguiente:

* Si se desea cambiar alguna de las opciones puede hacerse pulsando dos veces con el botón izquierdo del ratón y se entra en ‘el editor de gráficos’ desde el que puede cambiarse tanto el aspecto como la disposición del gráfico.

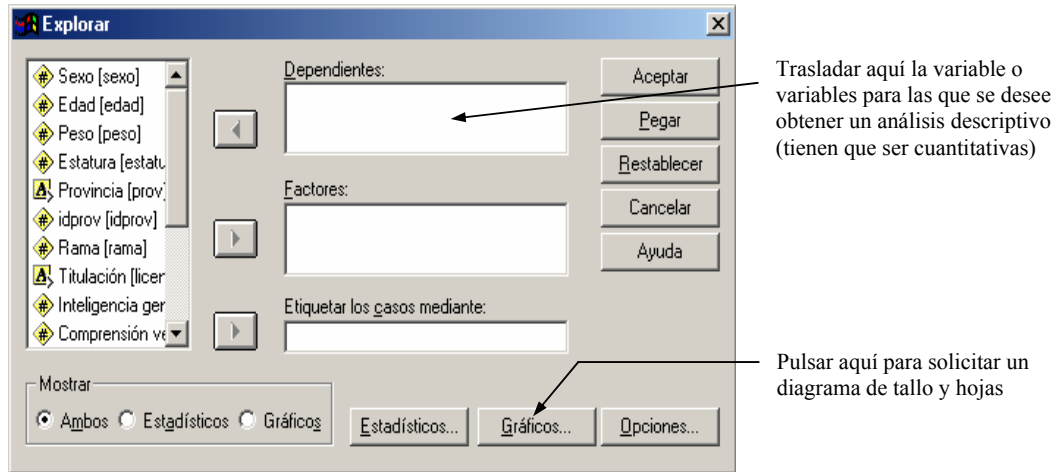


Figura 2.7a.

Este cuadro de diálogo ofrece diferentes posibilidades de análisis de datos. Por el momento nos centraremos tan sólo en la que nos interesa, que es elaborar un diagrama de tallo y hojas. Para ello, en primer lugar se seleccionan las variables para las que se desea elaborar este tipo de diagrama y se trasladan al cuadro 'Dependientes' mediante el botón . A continuación se selecciona el botón de la opción 'gráficos' y aparece el siguiente cuadro de diálogo:

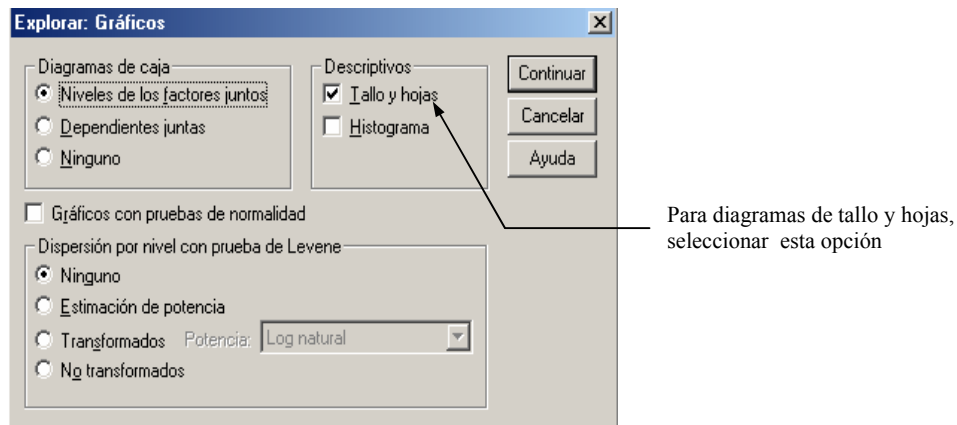


Figura 2.7b.

Veamos un ejemplo con la variable 'peso'. Los resultados obtenidos son los siguientes:

Peso	Stem-and-Leaf Plot	Tallo (en decenas)	Hoja (cada número representa 1 caso)
Frequency	Stem & Leaf		
2.00	3 . 99		
8.00	4 . 00112223		
32.00	4 . 555555555556666777777777778999999		
32.00	5 . 0000001111111222222223333334444		
38.00	5 . 555555556666666777777777888888899999		
41.00	6 . 0000000000000011111122223333333344444		
20.00	6 . 55566666666677888999		
13.00	7 . 0011333333333		
5.00	7 . 56688		
6.00	8 . 023333		
3.00	Extremes (>=85)		

Stem width: 10.00
Each leaf: 1 case(s)

El procedimiento Seleccionar casos

Supongamos que en lugar de obtener las distribuciones de frecuencias para ‘edad’ en los sujetos de toda la muestra, nos interesa obtenerla sólo para los varones. Existen varios procedimientos para seleccionar subgrupos de casos. Aquí utilizaremos el menú: **Datos -> Seleccionar Casos...**

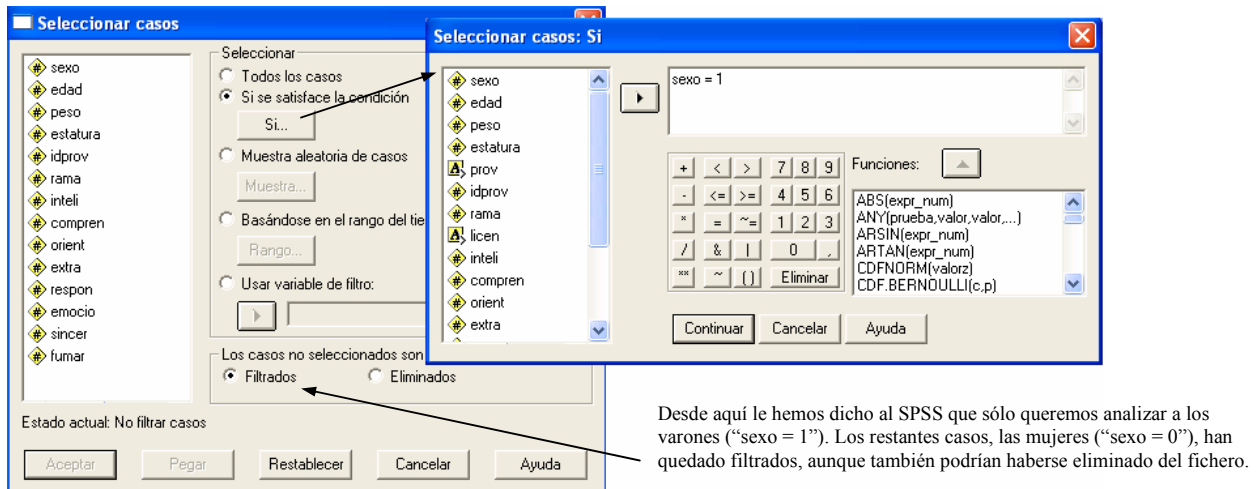


Figura 2.8.

Si activamos las opciones seleccionadas en la figura 2.8 (pulsando en “Aceptar”, como en cualquier menú), nótese que se crea automáticamente la variable `filter_$` (con etiqueta ‘sexo = 1 (FILTER)’).^{*} Mientras esta opción esté activada, todos los análisis solicitados se referirán tan sólo a los varones. Por ejemplo, si se solicita la distribución de frecuencias para ‘edad’, se obtiene lo siguiente:

Edad					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	24	5	4.2	4.2	4.2
	25	8	6.7	6.7	10.9
	26	15	12.6	12.6	23.5
	27	8	6.7	6.7	30.3
	28	22	18.5	18.5	48.7
	29	14	11.8	11.8	60.5
	30	11	9.2	9.2	69.7
	31	8	6.7	6.7	76.5
	32	7	5.9	5.9	82.4
	33	6	5.0	5.0	87.4
	34	2	1.7	1.7	89.1
	35	2	1.7	1.7	90.8
	36	1	.8	.8	91.6
	37	3	2.5	2.5	94.1
	38	2	1.7	1.7	95.8
	40	1	.8	.8	96.6
	41	1	.8	.8	97.5
	42	2	1.7	1.7	99.2
	46	1	.8	.8	100.0
Total		119	100.0	100.0	

Para *quitar el filtro* y volver a analizar todos los casos, se puede: 1) volver al menú de la figura 2.8 y activar la opción “Seleccionar todos los casos”, 2) o bien entrar en la solapa “Vista de variables” y borrar la nueva variable que se ha creado automáticamente (denominada `filter_$`).

^{*} Para utilizar más de un criterio de selección (por ejemplo, ‘varones fumadores’) se indica: (sexo = 1) & (fumar = 1).

2.2. Medidas de posición y de tendencia central

Los estadísticos de posición y tendencia central para las variables definidas en el editor de datos pueden obtenerse desde diferentes cuadros de diálogo. A continuación veremos los procedimientos más empleados para cada una de ellas.

Los centiles, los deciles y los cuartiles

Los centiles y otros cuantiles pueden obtenerse desde el procedimiento *Frecuencias* (ver figura 2.2.). En la parte inferior de este cuadro de diálogo hay un botón llamado 'Estadísticos'. Si se selecciona se obtiene el siguiente cuadro de diálogo:

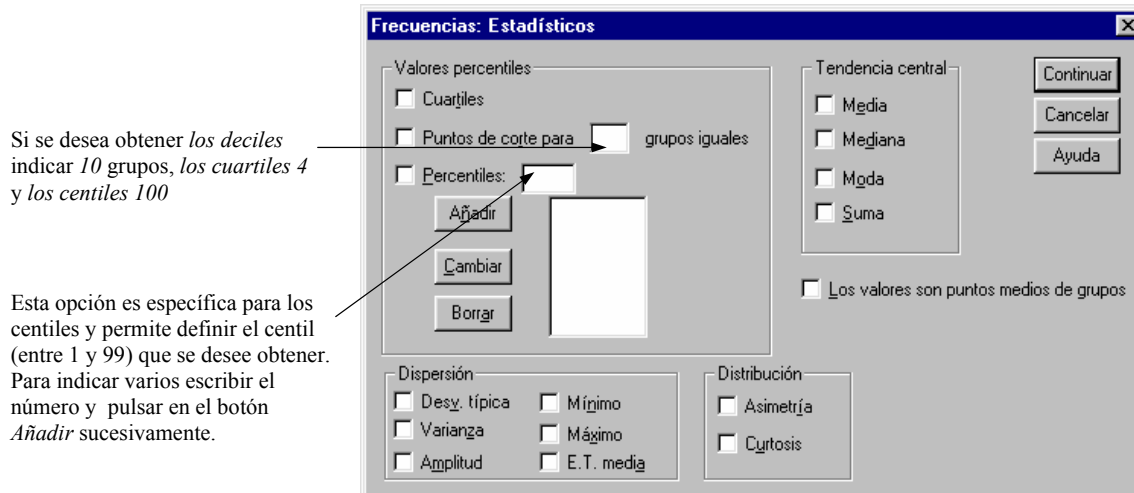


Figura 2.9.

Desde aquí puede indicarse que se obtengan diferentes tipos de cuantiles: los cuartiles o los percentiles. Se puede indicar el número exacto del centil que se desee obtener e irlo añadiendo a la lista inferior. Continuemos con nuestro ejemplo de la variable *edad*. Si le pedimos que nos dé los cuartiles y los centiles 10, 30, 60 y 90 se obtiene lo siguiente:

Estadísticos

Edad		Percentiles						
Válidos	Perdidos	10	25	30	50	60	75	90
200	0	25.00	27.00	27.00	29.00	29.60	31.00	34.00

Como se observa, hay un 75% de sujetos que tienen 31 años o menos y solo el 10% superan los 34 años de edad.

Media, mediana y moda

Estos estadísticos pueden obtenerse prácticamente desde cualquier cuadro de diálogo del menú *Analizar* de la figura 2.1. Uno de los más empleados es el del cuadro de diálogo 'Frecuencias' ya que cuando se elabora una tabla de frecuencias normalmente también se desea obtener las medidas de posición. Para ello se pulsa en el botón 'Estadísticos' del cuadro de diálogo de la figura 2.2. y aparece el cuadro de diálogo de la figura 2.9. Aquí puede seleccionarse las opciones 'media', 'mediana' y 'moda' y el SPSS calcula estos tres estadísticos para el listado de variables seleccionadas en el cuadro derecho de la figura 2.2. La media que se obtiene se corresponde con la media aritmética, la mediana con el valor central cuando los datos se ordenan en modo creciente y la

moda con el valor que obtiene la frecuencia absoluta mayor. Continuando con el ejemplo para la variable **edad** los resultados obtenidos son los siguientes:

Estadísticos

Edad		Media	Mediana	Moda
N				
Válidos	Perdidos			
200	0	29.35	29.00	28

Para conocer los estadísticos descriptivos de las variables también puede seleccionarse el procedimiento ‘*Descriptivos*’ de la figura 2.1. Este procedimiento presenta un cuadro de diálogo muy parecido al del procedimiento *Frecuencias* -> *Estadísticos*. Más específicamente:

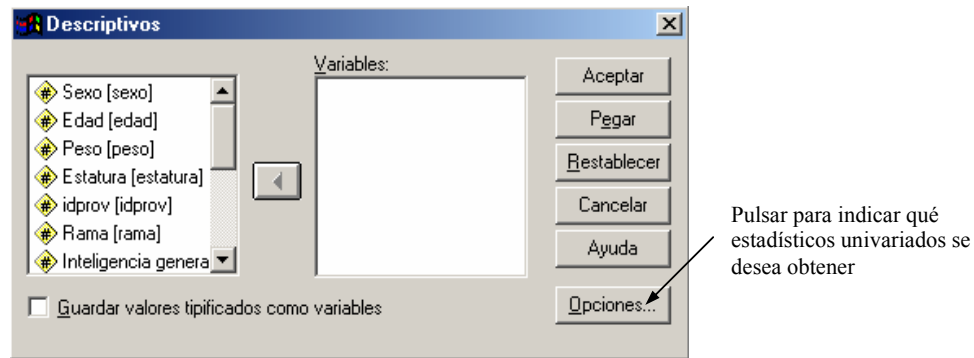


Figura 2.10.

Para obtener los estadísticos descriptivos, primero se trasladan las variables correspondientes al cuadro ‘variables’ y más tarde se pulsa en *Opciones*:

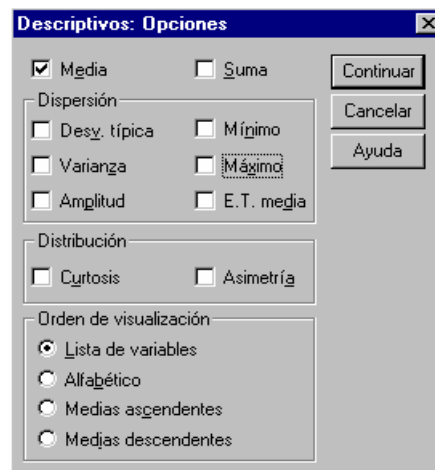


Figura 2.11.

En este caso, solamente está disponible el estadístico de tendencia central ‘media’ que, como antes, ofrece la media aritmética para las variables seleccionadas. Veamos un ejemplo para algunas variables del fichero:

Estadísticos descriptivos

	Edad	Peso	Estatura	N válido (según lista)
N	200	200	200	200
Media	29.35	58.2950	1.6756	

2.3. Medidas de variación

Las medidas o estadísticos de variación se obtienen de forma muy similar y desde los mismos menús y cuadros de diálogo que las de posición y tendencia central. Las más empleadas, la varianza y la desviación típica, aparecen por defecto en muchos menús del SPSS y en aquellas opciones que se denominan “*descriptivos*” sin especificar exactamente a qué se refieren.

Varianza y desviación típica

La varianza y desviación típica pueden obtenerse a partir del procedimiento ‘*Frecuencias*’ (ver figura 2.2.) en el botón ‘*Estadísticos*’ (ver figura 2.9.). Desde aquí puede pedirse que se calcule: la desviación típica, la varianza y la amplitud total o rango de las puntuaciones en las variables seleccionadas.

* Conviene tener en cuenta que el programa SPSS tiene la peculiaridad de que proporciona la *cuasivarianza*. Es decir, cuando se le pide que nos dé la varianza y la desviación típica, lo que da exactamente es el resultado de las fórmulas:

$$S'_x{}^2 = \frac{\sum (X_i - \bar{X})^2}{n-1} \quad \text{y} \quad S'_x = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

Otra forma de proceder para obtener medidas de variación es desde el menú *Analizar -> Estadísticos descriptivos -> Descriptivos* (ver figura 2.10.) o desde el menú *Analizar -> Explorar* (y el botón Estadísticos) de la figura 2.7. Desde ambos procedimientos se ofrecen exactamente las mismas opciones (desviación típica, varianza y amplitud total).

Veamos un ejemplo para algunas variables de las prácticas. Desde cualquiera de los cuadros de diálogo señalados se obtienen los siguientes resultados:*

Estadísticos descriptivos

	N	Rango	Desv. típ.	Varianza
Edad	200	24	3.88	15.033
Peso	200	53.00	10.3520	107.164
Estatura	200	.41	8.413E-02	7.077E-03
N válido (según lista)	200			

Diagramas de cajas

El SPSS también ofrece la posibilidad de elaborar los gráficos que expresan la dispersión de los datos en una variable. En concreto, desde el menú *Analizar -> Explorar* (y el botón Gráficos) de la figura 2.7b. puede definirse un diagrama de cajas para una variable. Este tipo de diagramas ofrecen información sobre el rango de la variable y los cuartiles.

Veamos un ejemplo para la variable ‘estatura’:

* Ver anexo I para la interpretación de resultados en notación científica (p.e. desviación típica para estatura = 8.413E-02 significa que la desviación típica de la variable estatura adopta el valor 0,0841)

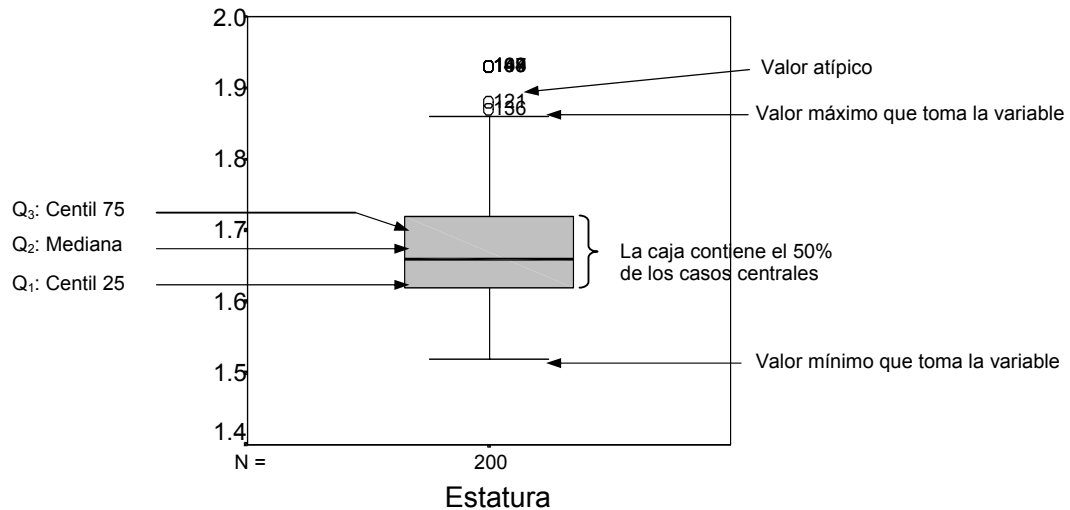


Figura 2.12.

Como se observa en la figura 2.12., la estatura oscila entre 1,52 y 1,93 metros estando la mediana en 1,66 m. y el 75% de los sujetos por debajo de los 1,71 m. Los valores superiores indican valores atípicos en los datos. En este caso hay ocho sujetos que miden más de 1,85 m.

2.4. Puntuaciones típicas y escalas derivadas

El SPSS ofrece la posibilidad de obtener las puntuaciones típicas para cada uno de los sujetos en cada una de las variables. Para solicitarlas se señala la opción 'Guardar valores tipificados como variables' del cuadro de diálogo del procedimiento *Descriptivos*:

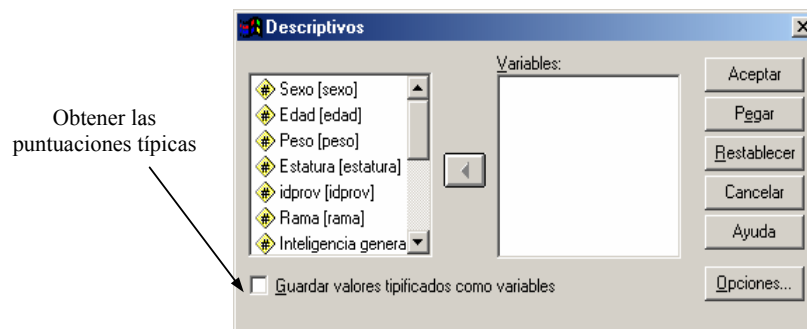


Figura 2.13.

El SPSS calcula las puntuaciones típicas a partir de la siguiente fórmula:

$$z_i = \frac{X_i - \bar{X}}{S'_X}$$

Recordemos que S'_X es la desviación típica insesgada.

Los resultados aparecen como nuevas columnas en el editor de datos (y se denominan *zvar*).

Ejercicio: Tomar como ejemplo la variable 'edad', seguir los pasos para obtener las puntuaciones típicas y una vez obtenida la nueva columna de datos (a la que el SPSS ha denominado *zedad*) comprobar que su media es 0 y su desviación típica 1 mediante el menú *Analizar -> Estadísticos descriptivos -> Descriptivos* (Opciones) de la figura 2.10.

Las escalas derivadas son aquellas construidas a partir de las puntuaciones típicas. Se definen a partir de la expresión: $T_i = a \cdot z_i + b$; donde la media de T es b y la varianza a^2 . A partir del SPSS es posible construir escalas derivadas. Para ello es imprescindible aprender a manejar el menú ‘Transformar’ de la barra de menús del *Editor de Datos de SPSS*. Como se comprobará a medida que el alumno se familiarice con el uso del SPSS este menú es bastante útil en la práctica.

Para definir la escala derivada hay que situarse en la ventana *Editor de datos* y seleccionar con el ratón el menú ‘Transformar’ cuyo aspecto es el siguiente:

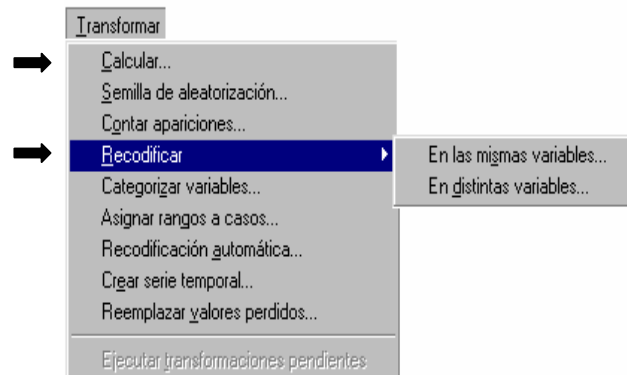


Figura 2.14.

A continuación veremos los procedimientos ‘Calcular’ y ‘Recodificar’ de este menú en detalle.

El procedimiento Calcular

Desde este procedimiento puede indicarse la operación que define la escala derivada. Por ejemplo, supóngase que queremos crear una escala derivada a partir de las puntuaciones típicas que hemos obtenido en ‘edad’ (cuya etiqueta era zedad) y la definimos como: $T_i = 10 \cdot z_i + 50$. Para indicar al SPSS la definición de T_i , hay que entrar en el cuadro de diálogo *calcular variable*:

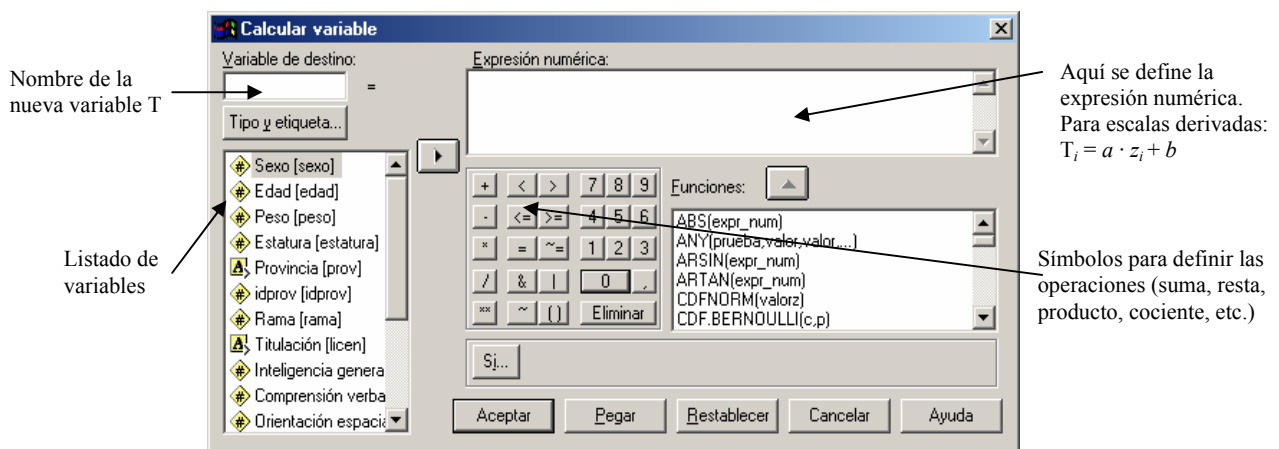


Figura 2.15.

En el recuadro ‘variable de destino’ se indica el nombre de la nueva variable (por ejemplo Tedad) y en el recuadro ‘expresión numérica’ la operación de la que resulta. Para ello, pueden seleccionarse las variables del recuadro inferior de la izquierda y trasladarse a la expresión numérica con la tecla . Los símbolos del centro indican la operación que se desea realizar (p.e. +: sumar, -: restar, *: multiplicar y /: dividir). Se recomienda poner paréntesis cada vez que se incluya una operación para una variable. En nuestro ejemplo: $Tedad = (10 * zedad) + 50$. Es decir:

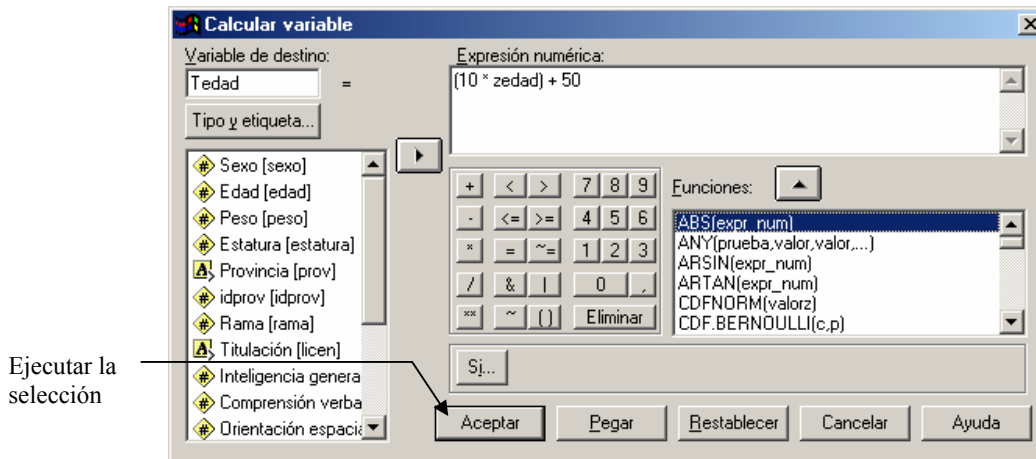


Figura 2.16.

Una vez definida la nueva variable se ejecuta la orden en el botón ‘*Aceptar*’ y el resultado aparece como una nueva columna al final del *editor de datos* del SPSS. Con esta nueva variable pueden calcularse los estadísticos descriptivos siguiendo las instrucciones dadas en los apartados anteriores y comprobar que la media es 50 y la desviación típica 10. Asimismo, también se puede calcular la puntuación típica correspondiente a T (ver figura 2.13.) y comprobar que su media es cero y su desviación típica 1.

El procedimiento Recodificar

Otra opción importante dentro del menú ‘*Transformar*’ (ver figura 2.14.) es la que permite codificar los valores originales de una variable en otros diferentes. El resultado puede reemplazar al de la variable original o puede grabarse en otra variable distinta. En caso de optar por la primera opción se perderán los valores originales; y en caso de optar por la segunda, la nueva variable aparecerá en el editor de datos con el nombre que le asignemos. El cuadro de diálogo correspondiente a esta segunda opción tiene el siguiente aspecto:

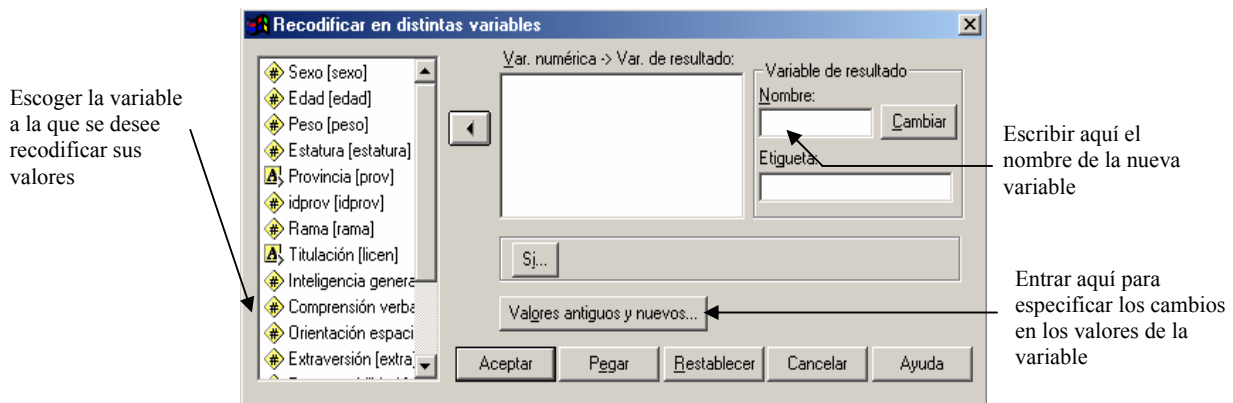


Figura 2.17.

En el recuadro “Var. numérica -> Var. de resultado” se indica cuál es la variable original que se quiere *recodificar* y cómo se llama la variable resultante. Supóngase que queremos *recodificar* los valores de la variable ‘*inteli*’ para resumirlos. Los valores de esta variable oscilan entre 1 y 30. La recodificación consiste en llamar ahora 1 a los valores entre 1 y 10 (‘nivel bajo’); 2 a los valores entre 11 y 20 (‘nivel medio’) y 3 a los valores entre 21 y 30 (‘nivel alto’). Para ello trasladamos la variable ‘*inteli*’ al recuadro derecho. En el recuadro ‘Nombre’ hay que etiquetar a la variable resultado (llamémosla por ejemplo ‘*nueva*’). Finalmente pulsamos en *cambiar*. En el recuadro tiene que aparecer *inteli -> nueva*. Para indicar los nuevos valores seleccionamos la opción

“valores antiguos y nuevos” y desde el cuadro de diálogo correspondiente indicamos cada uno de ellos. Puesto que en nuestro ejemplo los valores antiguos serán agrupados en intervalos, hay que utilizar la opción ‘Rango’. Por ejemplo, en la primera recodificación, el *valor antiguo* se define en ‘Rango’ desde 1 hasta 10; y el *valor nuevo* es 1. Una vez especificado el cambio, se pulsa en ‘Añadir’. Hay que repetir la operación para cada uno de los cambios. El cuadro final tiene que tener el siguiente aspecto:

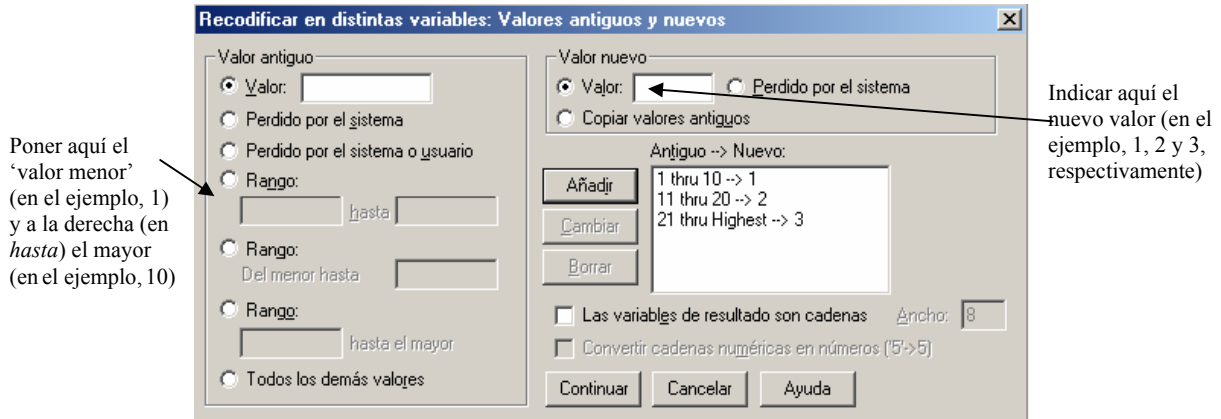


Figura 2.18.

2.5. Asimetría y curtosis

La asimetría y la curtosis pueden obtenerse a partir del menú *Analizar -> Frecuencias* (ver figura 2.2.) pulsando el botón ‘Estadísticos’ (ver figura 2.9.) y desde el menú *Analizar -> Estadísticos descriptivos -> Descriptivos* (ver figura 2.12.). Desde ambos cuadros de diálogo se ofrecen exactamente las mismas opciones: el índice de asimetría de Fisher para cada variable y el índice de curtosis. Asimismo, además del valor del estadístico tanto para la asimetría como para la curtosis, el SPSS proporciona información sobre el error típico en el cálculo de estos índices (este es un tema que no hemos estudiado en la asignatura análisis de datos I). Las fórmulas para ambos estadísticos también son insesgadas. Por ejemplo, la de la asimetría es:

$$As = \frac{n}{(n-1)(n-2)} \sum z_i^3$$

Veamos un ejemplo para algunas variables de las prácticas. Desde cualquiera de los cuadros de diálogo señalados se obtienen los siguientes resultados:

	Estadísticos descriptivos				
	N	Asimetría		Curtosis	
	Estadístico	Estadístico	Error típico	Estadístico	Error típico
Edad	200	1.205	.172	2.270	.342
Peso	200	.584	.172	.245	.342
Estatura	200	.775	.172	.659	.342
N válido (según lista)	200				

Los resultados también muestran una columna que no hemos visto en clase (el error típico asociado a cada estadístico). Los estadísticos encontrados en el ejemplo indican que casi todas las variables presentan asimetría positiva y distribuciones leptocúrticas.

* No se incluye la fórmula de la estimación insesgada de la *curtosis* porque es poco útil para la estadística descriptiva y es compleja de calcular.

EJERCICIOS:

Para afianzar lo aprendido en este tema en cuanto al manejo del SPSS para análisis descriptivos con una variable, es bueno acudir al aula de informática (por cuenta propia) y comprobar si se es capaz de responder a las siguientes preguntas, todas ellas referidas a los datos del fichero *practica.sav*:

1. ¿Cuál es la *estatura* mínima y máxima de los sujetos de la muestra? ¿Y la de los varones? ¿Y la de las mujeres?
2. ¿Qué porcentaje de sujetos mide menos de 1,65 m.?
3. ¿Cuál es el valor central de la variable *estatura*? ¿y de la variable *peso*?
4. Obtenga el valor del *peso* que es superado por el 15% de los sujetos de la muestra:
5. ¿Cuántos sujetos fuman?
6. ¿Cuántos sujetos son de provincias de Andalucía?
7. ¿En cuál de las cuatro características de personalidad (*extraversión*, *responsabilidad*, *estabilidad emocional* o *sinceridad*) han obtenido los sujetos menores puntuaciones?
8. Obtenga la representación gráfica más adecuada para las siguientes variables: *Sexo*, *estatura*, *inteli* y *rama*
9. ¿Qué porcentaje de varones obtiene en *extraversión* una puntuación de más de 40 puntos?
10. Confeccione el diagrama de tallo y hojas en la variable *estatura* para la muestra de varones, para la de mujeres y para la muestra total y trate de interpretar el resultado
11. Confeccione la gráfica que considere más adecuada para la variable "*licen*"
12. Sabiendo que para que los sujetos sean seleccionados por la empresa es necesario que obtengan en *responsabilidad* como mínimo una puntuación de 52 ¿Cuántos sujetos de la muestra total serán seleccionados?
13. ¿Quiénes obtienen mayores puntuaciones en *estabilidad emocional*, los varones o las mujeres?
14. ¿Son los sujetos de humanidades igual de homogéneos en la variable *Orientación espacial* que los de enseñanzas técnicas?
15. Obtenga los estadísticos descriptivos univariados y la representación gráfica más adecuada para la variable *sinceridad* en la muestra total, en la de varones y en la de mujeres.
16. Elabore una escala derivada con media 60 y desviación típica 10 para la variable *estabilidad emocional* y su representación gráfica.
17. Confeccione el diagrama de tallo y hojas para la variable *edad* en la muestra total, en la de varones y en la de mujeres y comente los resultados obtenidos en cada grupo.
18. Elabore un informe descriptivo sobre la variable *Edad expresada en meses*. En dicho informe tienen que aparecer los estadísticos de tendencia central, de variabilidad, las propiedades de la distribución y una representación gráfica.

Tema 3. Análisis descriptivos con más de una variable

En este tema revisaremos lo estudiado sobre *el análisis de datos con dos o más variables* y la forma de proceder con el SPSS. Se seguirá el mismo orden que en las clases teóricas.

3.1. Índices de asociación lineal

Para obtener los índices de asociación lineal con variables cuantitativas que hemos estudiado en la asignatura *Análisis de Datos en Psicología I* (la covarianza, S_{xy} y el coeficiente de correlación de Pearson, r_{xy}) se selecciona el menú *Analizar -> Correlaciones -> Bivariadas*:

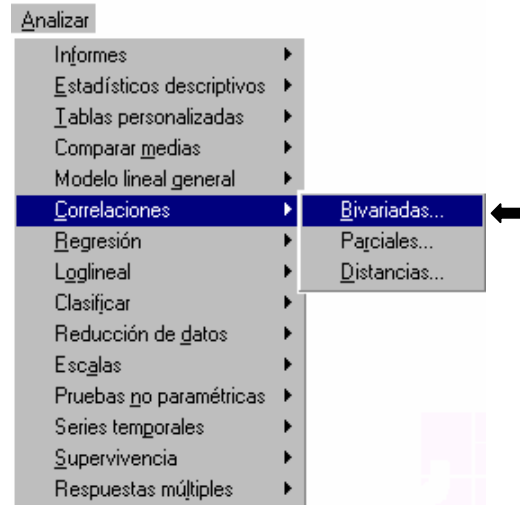


Figura 3.1.

El cuadro de diálogo para el procedimiento *Correlaciones bivariadas* tiene el siguiente aspecto:

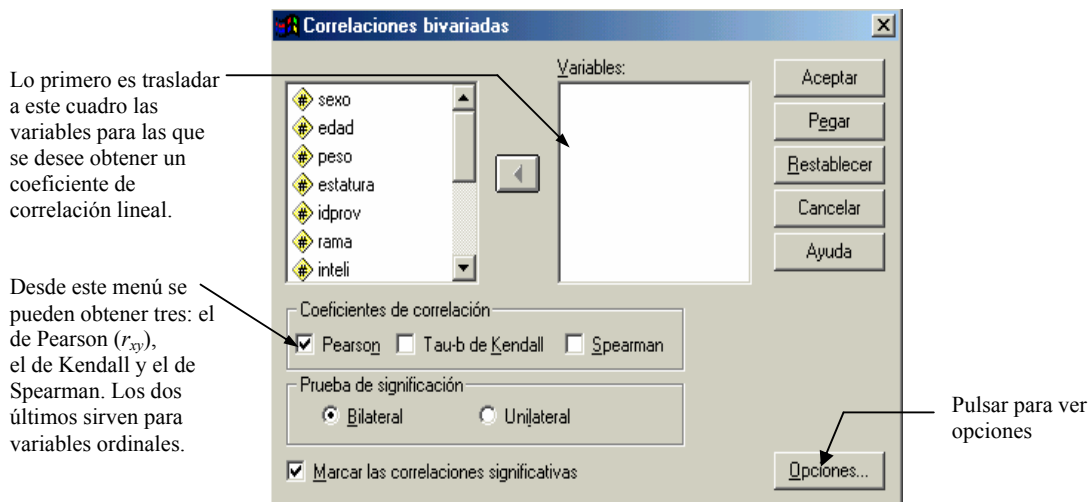


Figura 3.2.

La matriz de varianzas-covarianzas

La matriz de varianzas-covarianzas no aparece por defecto en el menú general de la figura 3.2. sino que hay que definirla desde el botón 'Opciones'. Si se pulsa este botón, aparece el cuadro de diálogo de la figura 3.3. Desde aquí puede indicarse que se muestren los estadísticos descriptivos (media y desviación típica) para cada una de las variables seleccionadas y también la matriz de varianzas-covarianzas (seleccionar la opción "Productos cruzados y covarianzas").

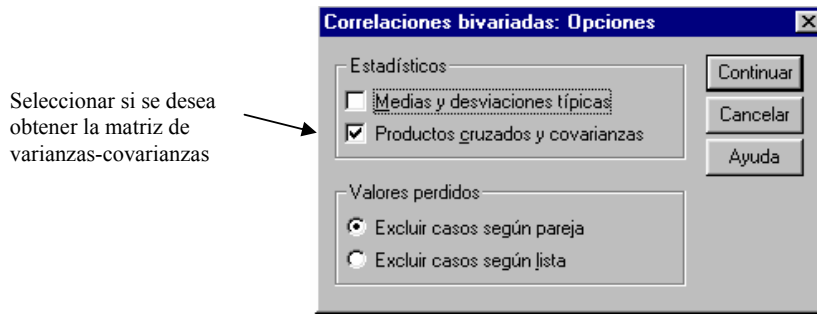


Figura 3.3.

En la covarianza pasa como en la varianza: el estadístico ofrecido es el insesgado. En este caso:

$$S'_{xy} = \frac{\sum_{i=1}^n x_i y_i}{n - 1}$$

La matriz de correlaciones

Desde el cuadro de diálogo de la figura 3.2., una vez se asignan las variables seleccionadas al cuadro 'Variables', si se pulsa en *Aceptar* aparece la matriz de correlaciones de Pearson entre todas las posibles variables seleccionadas en el visor de resultados.

Veamos cómo se resuelve en el SPSS alguno de los ejemplos que hemos calculado manualmente en clase. Por ejemplo, tomemos el ejemplo que aparece en la tabla de la página 145 del libro. En este ejemplo aparecen los datos de 15 sujetos para dos variables; *X*: inteligencia e *Y*: rendimiento. Como siempre, lo primero de todo es introducir los datos en el SPSS. En este caso tenemos una matriz de 15 sujetos x 2 variables. Una vez introducidos los datos y grabados en un archivo (en *a:\correlacion.sav*), se selecciona el menú *Analizar -> Correlaciones* y se siguen los pasos indicados en las figuras 3.4. y 3.5. Finalmente se pulsa en el botón *Aceptar* de la figura 3.4. Los resultados obtenidos son los siguientes:

Correlaciones

		X	Y
X	Correlación de Pearson	1.000	.868**
	Sig. (bilateral)	.	.000
	Suma de cuadrados y productos cruzados	118.000	55.000
	Covarianza	8.429	3.929
	N	15	15
Y	Correlación de Pearson	.868**	1.000
	Sig. (bilateral)	.000	.
	Suma de cuadrados y productos cruzados	55.000	34.000
	Covarianza	3.929	2.429
	N	15	15

Sig. (bilateral) se verá en la asignatura *Análisis de Datos II*

Estadísticos descriptivos

	Media	Desviación típica	N
X	8.00	2.90	15
Y	3.00	1.56	15

**. La correlación es significativa al nivel 0,01

Se ha señalado el coeficiente de Pearson con un círculo. Como se observa, se obtienen los mismos resultados que en el libro: $\bar{X} = 8$ e $\bar{Y} = 3$; $r_{xy} = 0,868^*$ y $S_{xy} = 3,929^*$ (ver pág. 154).

* En este cuadro también aparece la prueba de significación estadística para cada una de las correlaciones de Pearson. Sus resultados se expresan mediante asteriscos (* ó **) y el SPSS marca en amarillo las correlaciones significativas. En este curso no hemos aprendido a interpretar este dato pero se verá el próximo año en *Análisis de datos II*.

Veamos un ejemplo a partir de los datos de las prácticas. Obtengamos el coeficiente de correlación y la covarianza de las variables ‘estatura’ y ‘peso’. Los resultados obtenidos son los siguientes:

Correlaciones

		Peso	Estatura
Peso	Correlación de Pearson	1.000	.857**
	Sig. (bilateral)	.	.000
	Suma de cuadrados y productos cruzados	21325.595	148.490
	Covarianza	107.164	.746
	N	200	200
Estatura	Correlación de Pearson	.857**	1.000
	Sig. (bilateral)	.000	.
	Suma de cuadrados y productos cruzados	148.490	1.408
	Covarianza	.746	7.077E-03
	N	200	200

Los cuadrados incluyen la varianza de peso y estatura y los círculos la correlación y covarianza entre peso y estatura

** . La correlación es significativa al nivel 0,01 (bilateral).

Como se observa, el SPSS ofrece una salida que organiza los resultados en una tabla $J \times J$ (en este caso 2×2 pues hay dos variables). La salida incluye la correlación y la covarianza entre estatura y peso, peso y estatura, peso y peso y estatura y estatura. Su valor no cambia con el orden de las variables (es decir $S_{xy} = S_{yx}$; $r_{xy} = r_{yx}$). En este caso existe una relación lineal alta positiva entre estatura y peso ($r = 0,86$). Cuando se trata de las mismas variables la correlación es siempre 1 y la covarianza es la varianza (es decir $S_{xx} = S_x^2$; $r_{xx} = 1$). Por tanto, esta salida de SPSS incluye tanto la matriz de correlaciones como la matriz de covarianzas pues además de ofrecer la correlación y la covarianza entre la variables ofrece sus varianzas. Por ejemplo, la varianza de peso es 107,64.

3.2. Combinación lineal de variables

En este apartado veremos cómo se realizan transformaciones lineales con los datos de dos o más variables para obtener nuevas puntuaciones (p.e. $T_i = X_i + Y_i$; $T_i = AX_i + BY_i$; etc) y cómo se cumplen las propiedades sobre la media y la varianza de estas nuevas variables.

Lo primero es definir la nueva variable: T_i . Para ello se utiliza el menú *Transformar -> Calcular*, que vimos en el apartado 2.4. Con los datos de las prácticas, crearemos una combinación lineal de las cuatro facetas de personalidad (extra, respon, emocio y sincer). A la variable resultante la llamaremos PERSONA. El cuadro de diálogo de la figura 2.15. ha de quedar del siguiente modo:

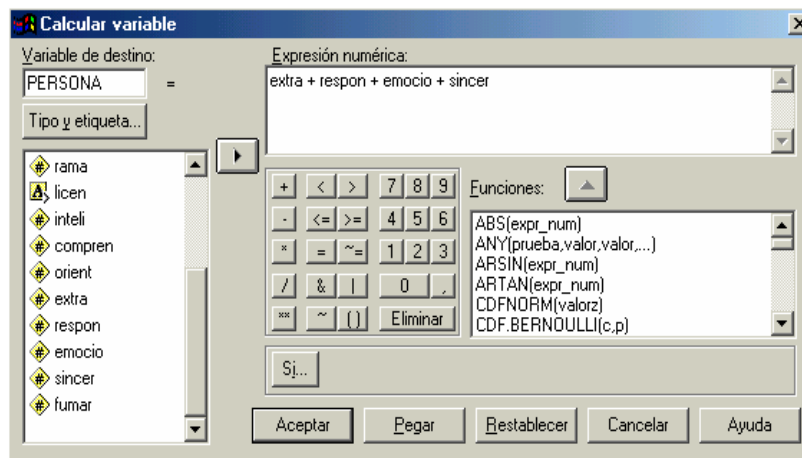


Figura 3.4.

La nueva variable es del tipo $T_i = X_i + Y_i + V_i + W_i$. Obtengamos ahora sus estadísticos descriptivos univariados desde el menú *Analizar -> Estadísticos descriptivos -> Descriptivos* y las covarianzas desde *Analizar -> Correlaciones -> Bivariadas (Opciones)*. El resultado es el siguiente:

Estadísticos descriptivos

	Extraversión	Responsabilidad	Estabilidad emocional	Sinceridad	PERSONA
N	200	200	200	200	200
Media	41.3600	46.1250	48.6250	23.1900	159.3000
Varianza	18.483	21.979	24.748	23.682	164.693

Correlaciones

		EXTRA	RESPON	EMOCIO	SINCER
EXTRA	Correlación de Pearson	1.000	.467**	.396**	.177*
	Sig. (bilateral)	.	.000	.000	.012
	Suma de cuadrados y productos cruzados	3678.080	1875.000	1685.000	735.320
	Covarianza	18.483	9.422	8.467	3.695
	N	200	200	200	200
RESPON	Correlación de Pearson	.467**	1.000	.422**	.186**
	Sig. (bilateral)	.000	.	.000	.008
	Suma de cuadrados y productos cruzados	1875.000	4373.875	1956.375	845.250
	Covarianza	9.422	21.979	9.831	4.247
	N	200	200	200	200
EMOCIO	Correlación de Pearson	.396**	.422**	1.000	.092
	Sig. (bilateral)	.000	.000	.	.193
	Suma de cuadrados y productos cruzados	1685.000	1956.375	4924.875	445.250
	Covarianza	8.467	9.831	24.748	2.237
	N	200	200	200	200
SINCER	Correlación de Pearson	.177*	.186**	.092	1.000
	Sig. (bilateral)	.012	.008	.193	.
	Suma de cuadrados y productos cruzados	735.320	845.250	445.250	4712.780
	Covarianza	3.695	4.247	2.237	23.682
	N	200	200	200	200

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

Como se observa la media en la nueva variable es 159,30; valor que se corresponde exactamente con el resultado de la suma de las medias en las cuatro variables originales ($\bar{T} = \bar{X} + \bar{Y} + \bar{V} + \bar{W}$). En cuanto a la varianza, el resultado es 164,69. Puede comprobarse que se llega al mismo resultado mediante: $S_T^2 = S_X^2 + S_Y^2 + S_V^2 + S_W^2 + 2(S_{XY} + S_{XV} + S_{XW} + S_{YV} + S_{YW} + S_{VW})$; o lo que es lo mismo, la suma del total de los elementos de la matriz de varianzas-covarianzas (aquí las correlaciones se han rodeado con círculos, las covarianzas con cuadrados y las varianzas con cuadrados en línea discontinua).

Veamos otro ejemplo. Supóngase que queremos crear una nueva variable con los datos de las prácticas que se define como el cociente entre el peso (en kilogramos) y la estatura al cuadrado (en metros). A este cociente se le denomina "Índice de masa corporal". Si el valor resultante se encuentra entre 20 y 25 quiere decir que la persona está en su peso normal. Si está entre 25 y 28 ó 30 indica que tiene sobrepeso. Si está por encima de 28 ó 30 indica que existe un problema de

obesidad y si está por debajo de 20, de excesiva delgadez. Nuestra nueva variable se definiría como $T_i = X_i / Y_i^2$. Para indicar al SPSS la definición de esta nueva variable se utiliza el procedimiento *Calcular variable*. Llamemos a la variable resultante IMC:

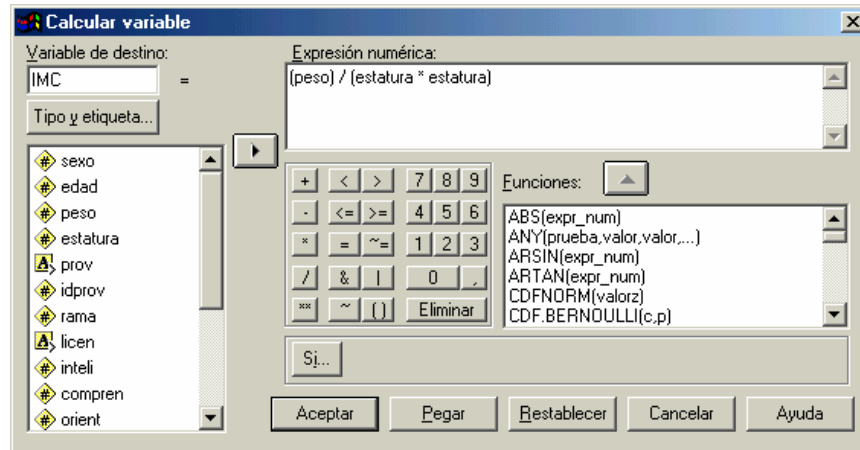
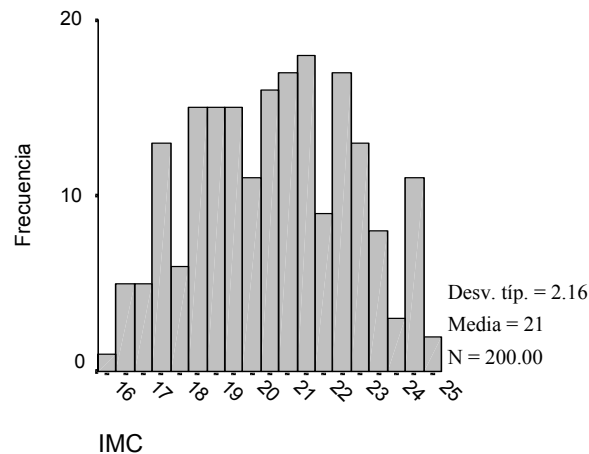


Figura 3.5.

Una forma rápida de interpretar los resultados en la nueva variable (IMC) es representando gráficamente el histograma. Para ello basta ir al menú *Analizar -> Frecuencias* y pulsar el botón *Gráficos* (ver figura 2.2.). El resultado obtenido con nuestra muestra es el siguiente:



Como se observa, la mayor parte de los sujetos tienen un peso normal (IMC entre 20 y 25), ninguno tiene tendencia a la obesidad, aunque hay un sector de sujetos considerable que puntúa por debajo de 20, por lo que está por debajo del peso normal.

3.3. Análisis de regresión lineal simple

El análisis de regresión lineal es una técnica estadística que se utiliza para estudiar la relación entre variables. En la investigación psicológica suele emplearse para pronosticar valores en una variable criterio (Y) desde las puntuaciones en una variable predictora (X). En esta asignatura solamente hemos estudiado la regresión lineal con una sola variable predictora (regresión simple) sin embargo, la regresión también puede efectuarse a partir de 2 o más variables predictoras (X_1 , X_2 , etc.). A medida que vayamos entrando en las opciones del SPSS para regresión simple, veremos de forma intuitiva cómo se procede para la regresión con más de un predictor (regresión múltiple).

Diagrama de dispersión

La primera aproximación a la relación entre dos variables (X e Y) puede hacerse a partir de un diagrama de dispersión. Por ejemplo, tomemos el ejemplo *a*) del cuadro 10.2. del libro (página

195). Como en otros ejemplos previos, tendremos que introducir los datos correspondientes a los 4 sujetos en las dos variables en un archivo nuevo del editor de datos. Una vez introducidos los datos (en el fichero *a:regresion.sav*) para realizar el diagrama de dispersión se selecciona el menú *Gráficos* y se elige el procedimiento 'Dispersión'. El cuadro de diálogo correspondiente es el siguiente:

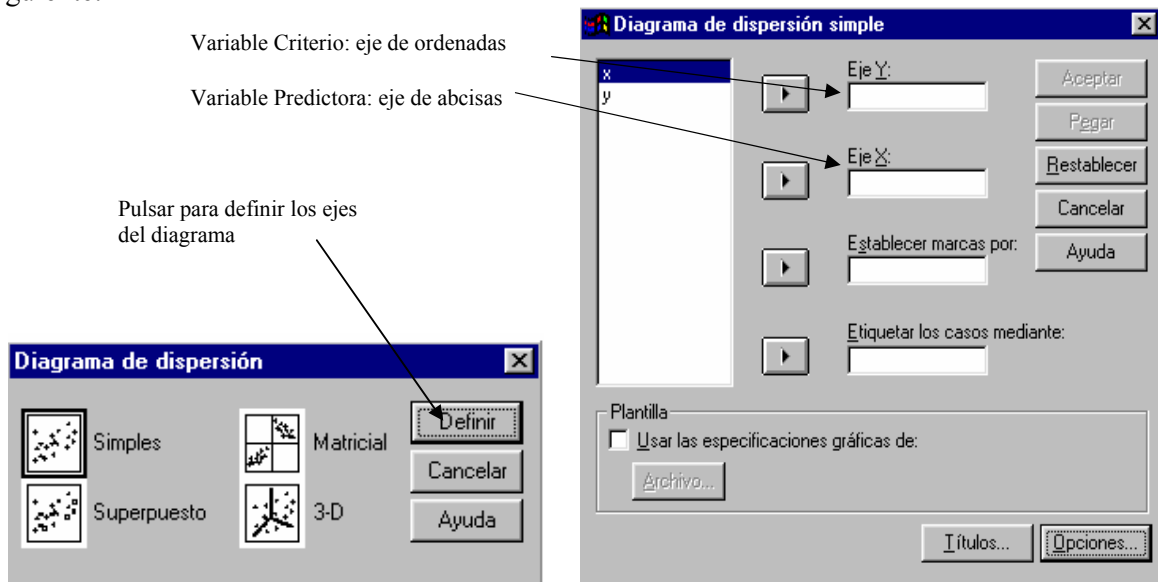
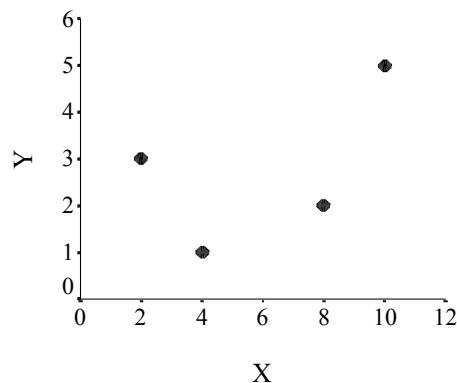


Figura 3.6.

Con esta definición se obtiene el siguiente diagrama de dispersión simple en el visor de resultados:



Como se observa, este diagrama es idéntico al de la figura 10.4.a) de la página 196 del libro. Su forma indica que los puntos en el diagrama no están perfectamente alineados pero se acercan a una hipotética línea recta.

Ecuación de regresión

Podrían trazarse diferentes rectas para realizar pronósticos de una variable a partir de la otra (por ejemplo de Y a partir de X en una regresión de Y sobre X). Las rectas de regresión tienen una fórmula muy simple:

$$Y'_i = A_{YX} + B_{YX} X_i$$

El objetivo es encontrar aquella recta que minimice la distancia entre lo encontrado (Y) y lo pronosticado (Y'). Es decir, que minimice la expresión:

$$\sum \frac{(Y_i - Y'_i)^2}{n}$$

Para ello calculamos los coeficientes del modelo mediante las fórmulas:

$$B_{YX} = \frac{n \cdot \sum X_i \cdot Y_i - \sum X_i \cdot \sum Y_i}{n \cdot \sum X_i^2 - (\sum X_i)^2} \quad A_{YX} = \bar{Y} - B_{YX} \bar{X}$$

Veamos cómo se procede en el SPSS para calcular dichos coeficientes, obtener la recta de regresión y valorar la bondad del modelo. Se selecciona el menú *Analizar* -> *Regresión* -> *Lineal*:

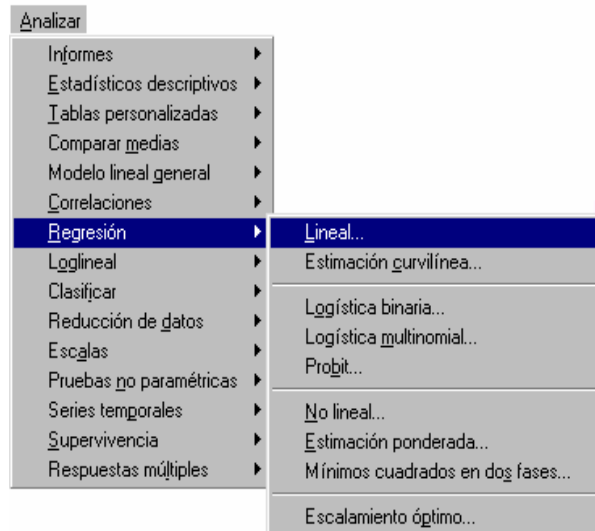



Figura 3.7.

El cuadro de diálogo correspondiente al procedimiento *Regresión lineal* aparece en la figura 3.8. Como en otros cuadros de diálogo del SPSS, lo primero es seleccionar las variables. En este caso hay que distinguir entre 'Dependiente' e 'Independiente' (o independientes si se trata de una regresión múltiple). Como siempre, las variables se trasladan mediante el botón .

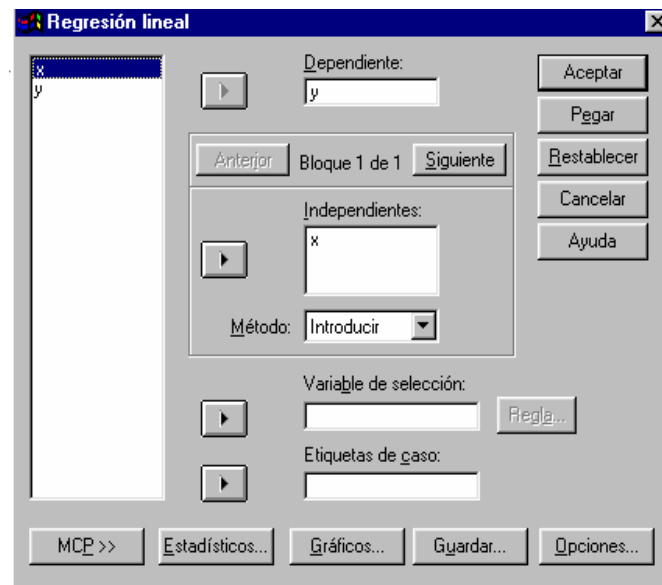


Figura 3.8.

Dentro de este menú hay otras opciones. Por el momento, la que nos interesa es la que se encuentra en el botón '*Estadísticos*', cuyo cuadro de diálogo aparece en la figura 3.9. Desde aquí podemos solicitar que se ofrezcan las estimaciones de la pendiente y el origen de la recta de regresión (según

el método de mínimos cuadrados), la matriz de covarianzas para las variables, el ajuste del modelo (coeficiente de Pearson al cuadrado o *coeficiente de determinación*) y los estadísticos descriptivos (media y varianza):

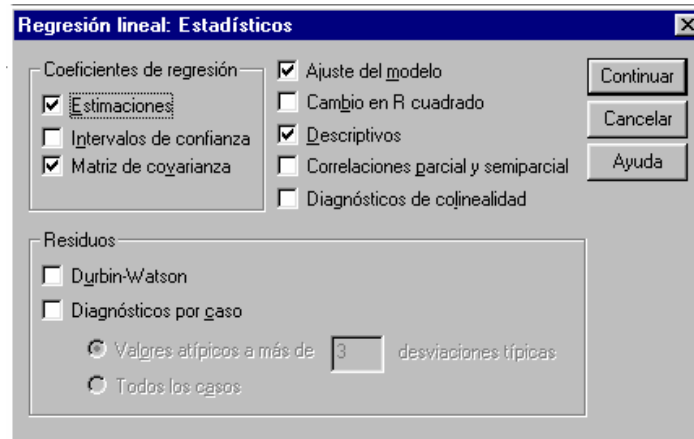


Figura 3.9.

Otra opción relevante del menú regresión lineal es la de “*Guardar*” (ver figura 3.10).^{*} Desde su correspondiente cuadro de diálogo es posible indicar que se guarden los valores pronosticados por el modelo (las Y'_i) y los residuos (las $Y_i - Y'_i$) en el editor de datos:

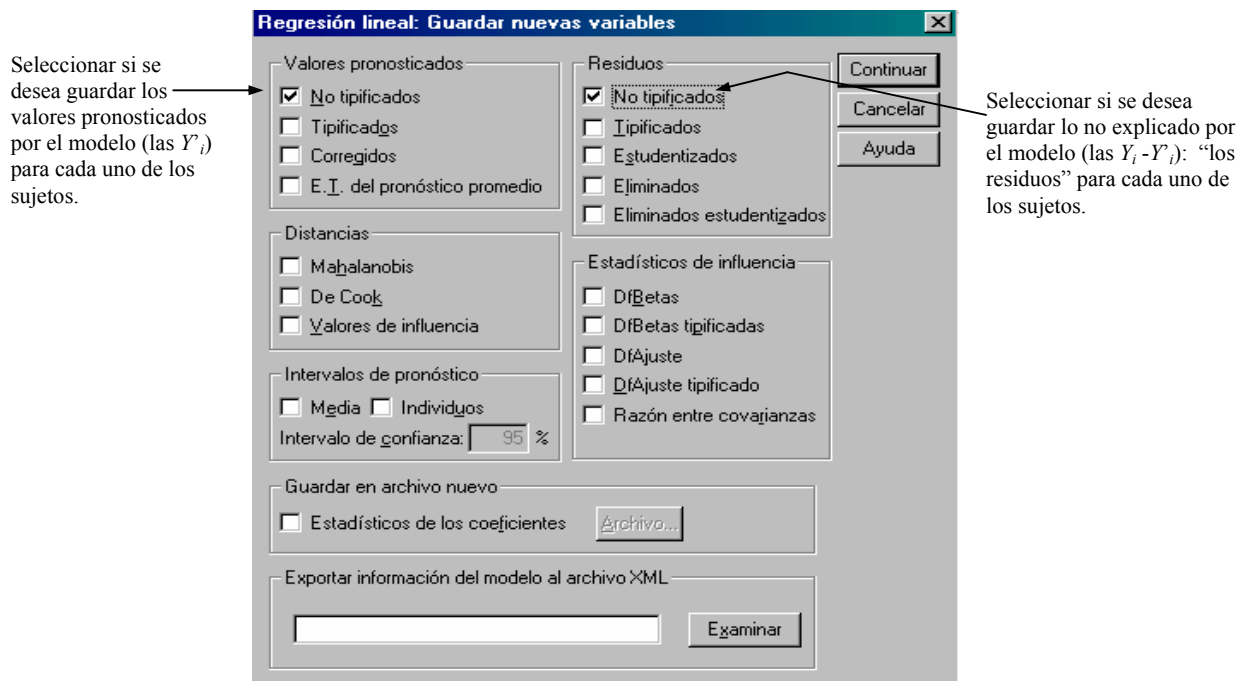


Figura 3.10.

Veamos cuál es el resultado para el ejemplo del cuadro 10.2.a) del libro. Si indicamos todas estas instrucciones, el resultado que ofrece el SPSS es el siguiente (nótese que se ofrecen muchas tablas de resultados. Aquí explicaremos sólo las más importantes y aquellas que el alumno de psicología de primero puede manejar con soltura):

^{*} No vamos a entrar en las restantes opciones del menú ‘Regresión lineal’ porque la mayoría requieren conocimientos de estadística inferencial, que el alumno aún no posee. Dichas opciones se verán en detalle en las prácticas de la asignatura de segundo *Análisis de Datos en Psicología II*.

Regresión

Estadísticos descriptivos

	Media	Desviación típ.	N
Y	2.75	1.71	4
X	6.00	3.65	4

Correlaciones

		Y	X
Correlación de Pearson	Y	1.000	.535
	X	.535	1.000
Sig. (unilateral)	Y	.	.233
	X	.233	.
N	Y	4	4
	X	4	4

Variables introducidas/eliminadas^a

Modelo	Variables introducidas	Variables eliminadas	Método
1	X ^a	.	Introducir

a. Todas las variables solicitadas introducidas
b. Variable dependiente: Y

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregido	Error típ. de la estimación
1	.535 ^a	.286	-.071	1.77

a. Variables predictoras: (Constante), X

ANOVA^b

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2.500	1	2.500	.800	.465 ^a
	Residual	6.250	2	3.125		
	Total	8.750	3			

a. Variables predictoras: (Constante), X
b. Variable dependiente: Y

Coefficientes^a

Modelo		Coefficients no estandarizados		Coefficients estandarizados		t	Sig.
		B	Error típ.	Beta			
1	(Constante)	1.250	1.896			.659	.577
	X	.250	.280	.535		.894	.465

a. Variable dependiente: Y

Esto es el coeficiente de determinación, r^2_{XY} o la proporción de varianza en común entre X e Y.

Esto se verá en la asignatura *Análisis de Datos II*

Los coeficientes A y B del modelo pronosticado en directas se ven en esta columna: siendo A = 1,25 y B = 0,25. Luego:
$$Y_i = 1,25 + 0,25 X_i$$

En esta columna aparece el modelo en típicas:
 $z_Y = r_{XY} z_X$ (en el ejemplo: $z_Y = 0,535 z_X$)

Como se observa, el SPSS ofrece muchas tablas de datos como resultado de la regresión. Por el momento solamente nos fijaremos en dos: La primera se refiere a los coeficientes del modelo y la segunda a su bondad. En cuanto a la primera, se toman los coeficientes no estandarizados. En este caso el mejor modelo para pronosticar Y_i a partir de X_i es $Y_i = 1,25 + 0,25 X_i$

Bondad de ajuste del modelo

Además de la fórmula de la recta de regresión, resulta necesario disponer de información sobre el grado en que el modelo se ajusta a los datos observados (nube de puntos). Una primera aproximación es la interpretación gráfica del problema. Para elaborar la gráfica del ajuste de la recta a los datos observados se pulsa el menú **Gráficos -> Interactivos -> Diagramas de dispersión**. Lo primero es definir la variable del criterio (en el eje de ordenadas) y la variable predictor (en el eje de abscisas) desde la solapa 'Asignar variables'. A continuación se selecciona el método "regresión" desde la solapa 'Ajuste'. El cuadro de diálogo desde el que se hacen estas selecciones tiene el siguiente aspecto:

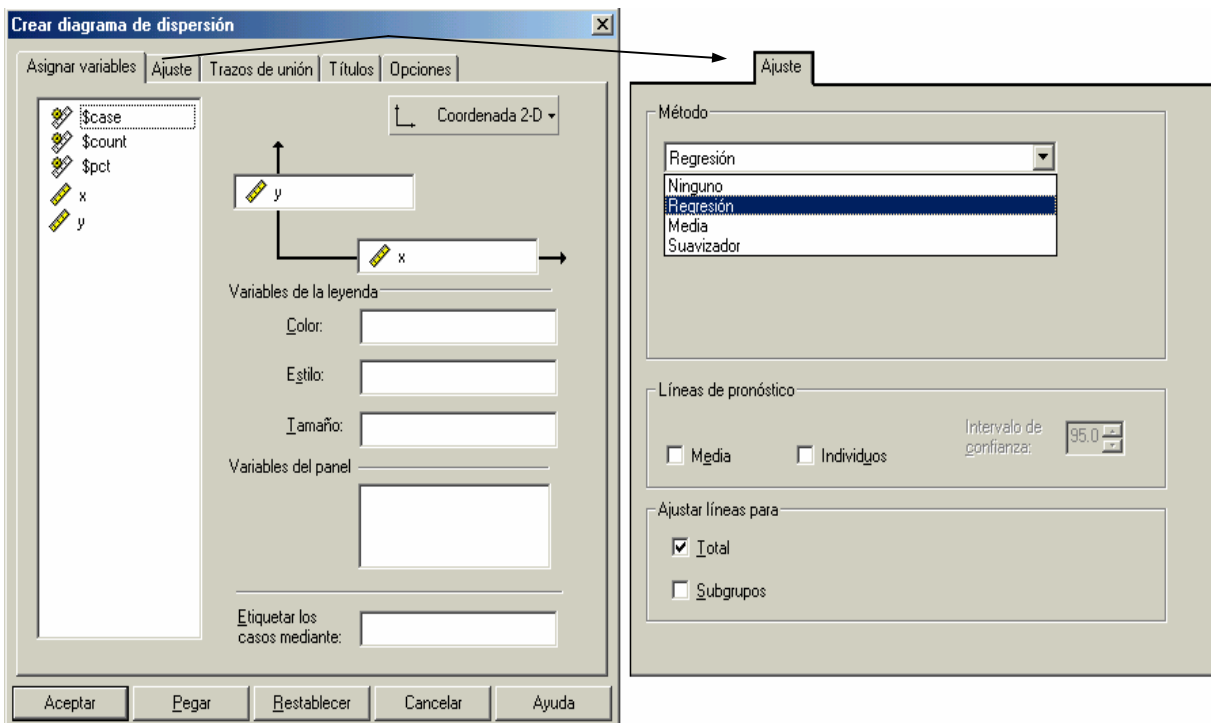
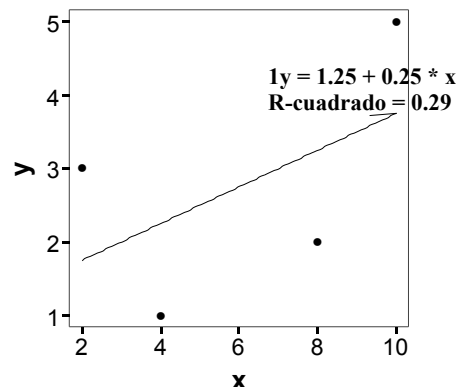


Figura 3.11.

El gráfico obtenido es el siguiente:



Como se observa, los puntos se alejan bastante de la recta, luego no hay un buen ajuste.

Además de la interpretación gráfica del problema, la forma de cuantificar la bondad del modelo es mediante el coeficiente de determinación, r^2_{xy} . Se trata de una medida estandarizada que toma valores entre 0 y 1 y cuya interpretación es muy sencilla: representa la proporción de varianza explicada de la variable del criterio a partir de la predictora. En nuestro ejemplo, $r^2_{xy} = 0,286$ por lo que el modelo no es adecuado para hacer pronósticos de Y a partir de X .

La bondad del modelo también puede valorarse a partir del análisis de los errores en los pronósticos, frecuentemente llamados *residuos* ($Y_i - Y_i$). Nótese que ejecutando las órdenes indicadas en el anterior apartado han resultado dos nuevas variables en el editor de datos (la variable 'pre_1' que se corresponde con las Y_i y lleva la etiqueta 'Unstandardized predicted value' y la variable 'res_1' que se refiere a las $Y_i - Y_i$ y lleva la etiqueta 'Unstandardized residual').

Teniendo los datos para Y_i , Y_i e ($Y_i - Y_i$) podemos evaluar la bondad del modelo a partir de *la descomposición de la varianza del criterio* ($S^2_Y = S^2_{Y'} + S^2_{Y-Y'}$). Para ello entramos en el menú *Analizar -> Estadísticos descriptivos -> Descriptivos: Opciones* (ver figuras 2.11 y 2.12) e indicamos en el cuadro de diálogo que se calcule la varianza para estas tres variables (Y , pre_1 y res_1). Los resultados obtenidos en el visor son los siguientes:

Estadísticos descriptivos

	N	Varianza
Y	4	2.917
Unstandardized Predicted Value	4	.833
Unstandardized Residual	4	2.083
N válido (según lista)	4	

Como se comprueba, la varianza del criterio ($S^2_Y = 2,917$) se descompone en la varianza de los pronósticos ($S^2_{Y'} = 0,833$) y la de los errores ($S^2_{Y-Y'} = 2,083$).

Veamos ahora un ejemplo a partir de los datos de las prácticas. Supóngase que queremos predecir la variable 'peso' a partir de la variable 'estatura'. Para ello construimos la recta de regresión $peso_i = A + B \text{ estatura}_i$. La definición del modelo en el SPSS es la siguiente:

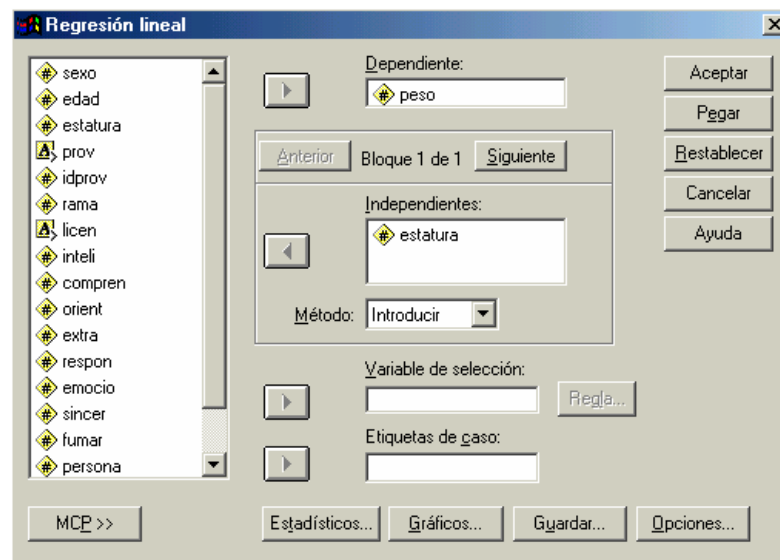


Figura 3.12.

Los resultados obtenidos para el modelo planteado son los siguientes (sólo se muestran las tablas necesarias para la interpretación):

Variables introducidas/eliminadas ^b

Modelo	Variables introducidas	Variables eliminadas	Método
1	Estatura ^a	.	Introducir

a. Todas las variables solicitadas introducidas

b. Variable dependiente: Peso

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.857 ^a	.734	.733	5.3510

a. Variables predictoras: (Constante), Estatura

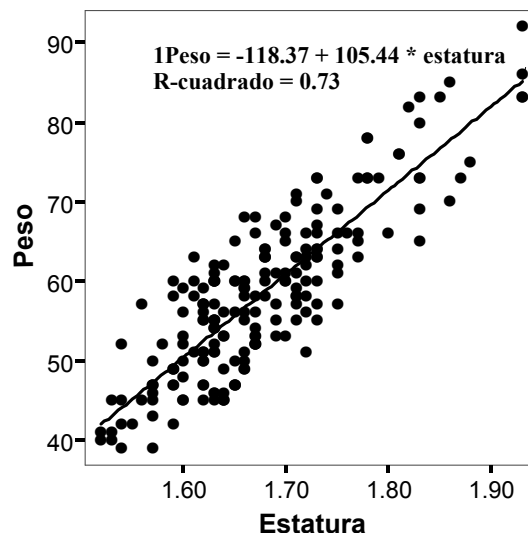
Coefficientes^a

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-118.375	7.565		-15.648	.000
	Estatura	105.437	4.509	.857	23.384	.000

a. Variable dependiente: Peso

El modelo resultante es: $\text{peso}'_i = -118,375 + 105,437 \text{ estatura}_i$. Como se observa, el coeficiente de determinación (R al cuadrado) es 0,734 por lo que el modelo es adecuado para explicar la relación entre la variable 'estatura' y la variable 'peso'. Es decir, la variable 'estatura' tiene una alta capacidad predictiva para explicar la variable del criterio, el 'peso'.

El gráfico obtenido definiendo las opciones del diagrama de dispersión interactivo de la figura 3.11. es el siguiente:



El gráfico muestra que existe un buen ajuste. No obstante, encontrar un resultado similar a éste con variables psicológicas no es frecuente. Para comprobarlo a continuación haremos una regresión de

la variable Responsabilidad sobre la variable Estabilidad emocional, que son dos variables psicológicas del fichero *practica.sav*. Los resultados que ofrece el SPSS son los siguientes:

Variables introducidas/eliminadas^b

Modelo	Variables introducidas	Variables eliminadas	Método
1	Estabilidad emocional ^a	.	Introducir

a. Todas las variables solicitadas introducidas

b. Variable dependiente: Responsabilidad

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.422 ^a	.178	.174	4.2621

a. Variables predictoras: (Constante), Estabilidad emocional

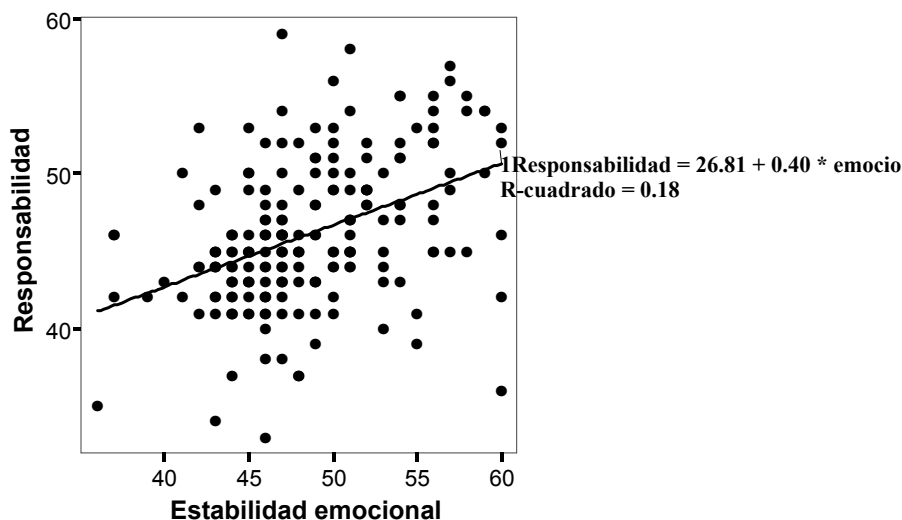
Coefficientes^a

Modelo		Coefficients no estandarizados		Coefficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	26.809	2.968		9.031	.000
	Estabilidad emocional	.397	.061	.422	6.541	.000

a. Variable dependiente: Responsabilidad

Como puede observarse, en este caso la correlación de Pearson entre las variables es 0,42 y la bondad del modelo tan sólo alcanza el valor 0,18. Sin embargo, los psicólogos del equipo podrían concluir que existe un buen ajuste para el modelo que predice la responsabilidad a partir de la estabilidad emocional, pues así lo indican las pruebas inferenciales (que no se han estudiado en esta asignatura pero se verán en la asignatura *Análisis de Datos en Psicología II*).

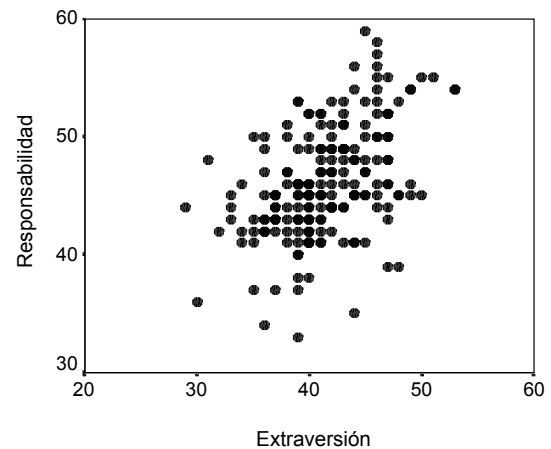
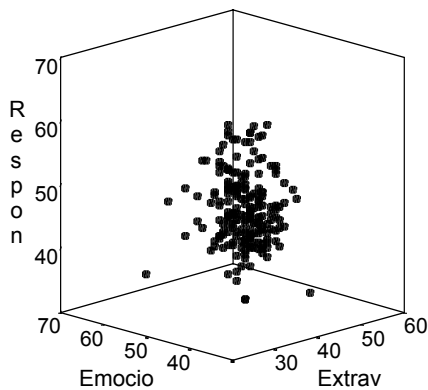
El gráfico inferior muestra el ajuste para el modelo estimado. La comparación con el gráfico de la página anterior pone de manifiesto que las correlaciones entre variables psicológicas adoptan un rango muy inferior al de otras variables observadas, como son el peso y la estatura.



3.4. Análisis de Regresión Lineal Múltiple

Una alternativa a la situación de poco ajuste es incluir una variable predictora adicional (o más de una) y observar si su inclusión produce algún cambio en R cuadrado. Si consideramos dos variables predictoras estamos planteando el siguiente modelo: $Y'_i = A + B_1X_{i1} + B_2X_{i2}$.

Veamos si el ajuste al modelo de regresión de responsabilidad sobre estabilidad emocional mejora al añadir la variable extraversión como segunda variable predictora. El modelo planteado es: $\text{respon}'_i = A + B_1 \text{emocio}_i + B_2 \text{extrav}_i$. La primera aproximación al problema puede hacerse a partir del gráfico de dispersión (ver figura 3.8.). En este caso puede elaborarse un gráfico 3-D para observar la relación entre las dos predictoras y el criterio y también el gráfico simple para observar la relación entre respon y extra:



En el gráfico tridimensional puede verse que en el plano respon-emocio se observa como los puntos indican una relación con tendencia lineal positiva. La gráfica de la derecha muestra el plano respon-extra. Los puntos indican que la variable extraversión tiene una tendencia a la relación lineal positiva con la variable responsabilidad. Comparándola con la gráfica en la que veíamos el diagrama de dispersión entre respon y emocio, en ésta los puntos están algo más alineados, por lo que cabe esperar que la inclusión de la variable extra en el modelo mejore la bondad de ajuste.

Para obtener los coeficientes del modelo y valorar su ajuste se selecciona el menú *Analizar -> Regresión -> Lineal*. De forma similar a la regresión simple, primero se traslada la variable respon a la casilla *Dependiente* y a continuación las variables emocio y extra a la casilla *Independientes*.

Al ejecutar el análisis, en primer lugar aparece la tabla que informa sobre la especificación de las variables en el modelo:

Variables introducidas/eliminadas^b

Modelo	Variables introducidas	Variables eliminadas	Método
1	Extraversión, Estabilidad emocional ^a	.	Introducir

a. Todas las variables solicitadas introducidas

b. Variable dependiente: Responsabilidad

A continuación se muestra la tabla que incluye el coeficiente de determinación y más abajo la tabla de resultados con los coeficientes del modelo (A , B_1 y B_2):

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error tít. de la estimación
1	.534 ^a	.285	.278	3.9848

a. Variables predictoras: (Constante), Extraversión, Estabilidad emocional

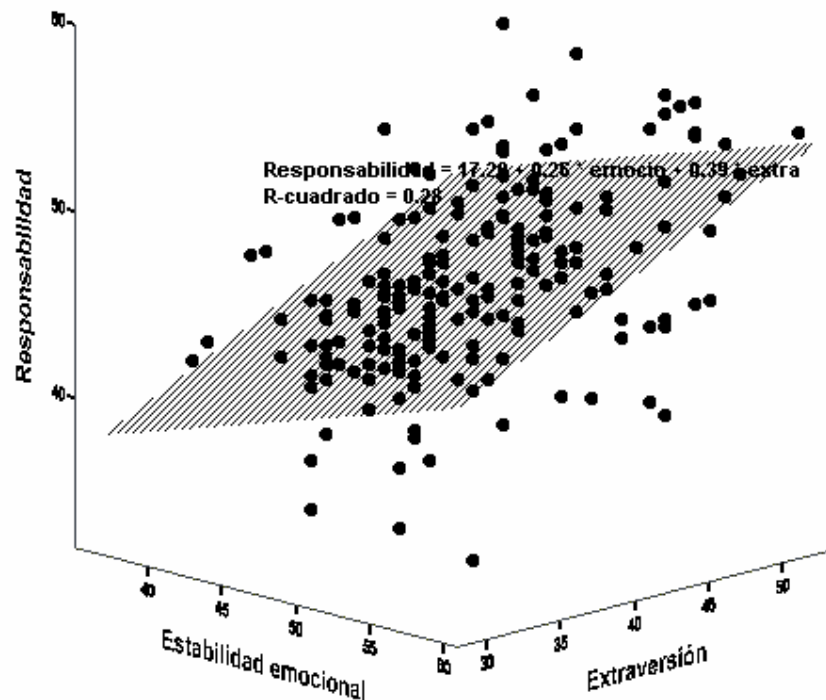
Coefficientes^a

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error tít.	Beta		
1	(Constante)	17.199	3.291		5.226	.000
	Estabilidad emocional	.264	.062	.280	4.273	.000
	Extraversión	.389	.072	.356	5.433	.000

a. Variable dependiente: Responsabilidad

Como se observa, el modelo simple explicaba un 17,8 % de varianza de la variable **responsabilidad** y el modelo múltiple explica el 28,5 %. Por tanto, el modelo resultante (definido mediante: $responsabi_i = 17,2 + 0,26 emocio_i - 0,39 extra_i$) aporta un 11% más en la explicación de la variable del criterio.

En este caso también es posible elaborar el gráfico tridimensional con el ajuste del plano de regresión es el siguiente desde la figura 3.11, activando la opción “Coordenada 3-D”. El resultado es el siguiente:



Se deja como ejercicio comprobar que al añadir la variable edad al modelo de regresión de peso sobre estatura, el ajuste mejora un 0,10%.

3.5. Distribuciones de frecuencias y representaciones gráficas

En este apartado veremos cómo se obtienen distribuciones conjuntas de frecuencias y su representación gráfica para dos o más variables de diversos tipos.

Dos variables cualitativas:

Para describirlas hay que elaborar una *Tabla de contingencia*. La forma de proceder en el SPSS es desde el menú *Analizar -> Estadísticos descriptivos -> Tablas de contingencia*:

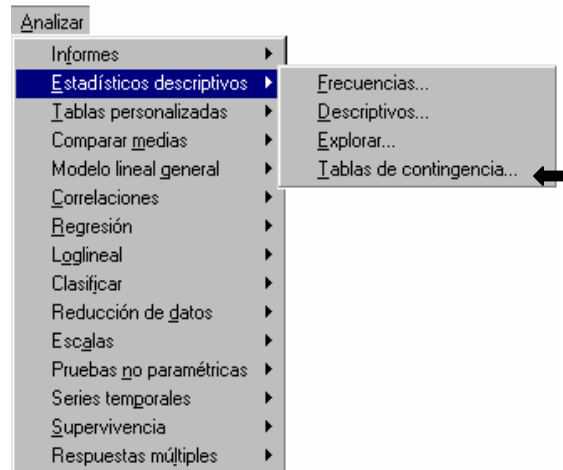


Figura 3.13.

Si se selecciona el procedimiento *Tablas de contingencia* aparece el cuadro de diálogo de la figura 3.14). Supongamos que queremos conocer la distribución conjunta de frecuencias de las variables fumar y sexo con los datos de las prácticas. Para indicar al SPSS esta instrucción, primero se seleccionan las variables. Por ejemplo, las categorías de la variable sexo pueden ser las filas de la tabla de frecuencias y las de la variable fumar las columnas. Este cuadro de diálogo también permite obtener una representación gráfica para las dos variables.

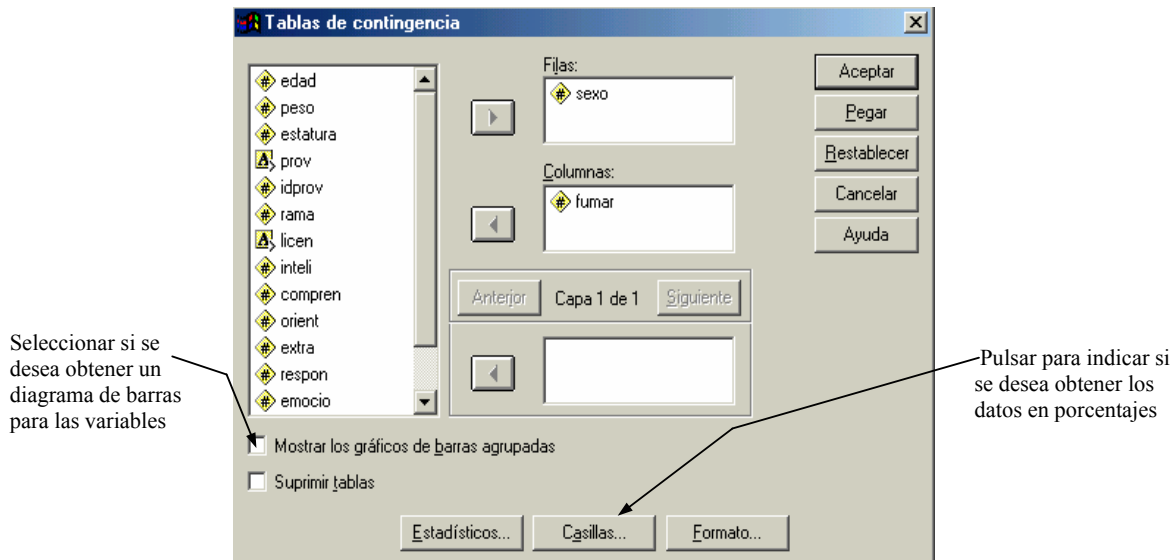
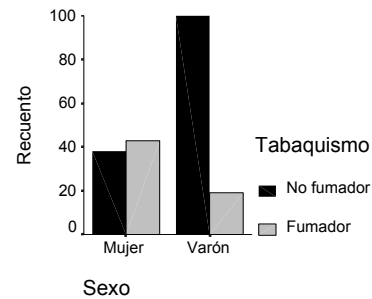


Figura 3.14

Si se pulsa en el botón *Aceptar* el resultado obtenido es el siguiente:

Tabla de contingencia Sexo * Tabaquismo

Recuento		Tabaquismo		Total
		No fumador	Fumador	
Sexo	Mujer	38	43	81
	Varón	100	19	119
Total		138	62	200



En la tabla de resultados aparecen las *distribuciones conjuntas* (n_{ij}) y *marginales* (n_i y n_j) de las variables y la representación gráfica del diagrama de barras conjunto. Como se observa, de los 200 sujetos de la muestra 62 fuman y 138 no fuman. De los fumadores, la mayoría son mujeres y de los no fumadores, varones. Hay 100 no fumadores que son varones y 19 varones que son fumadores.

Las distribuciones conjuntas también pueden obtenerse en frecuencias relativas o porcentajes para ambas variables (filas y columnas). Para ello se selecciona el botón *Casillas* del cuadro de diálogo de la figura 3.14). cuyo aspecto es el siguiente:

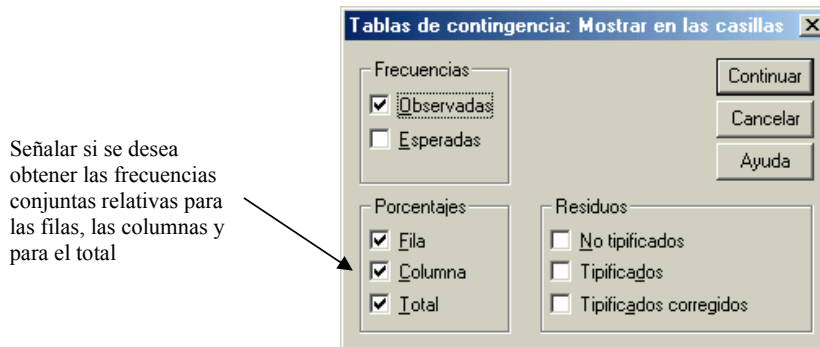


Figura 3.15

Seleccionando las opciones señaladas se obtienen los mismos resultados que antes pero expresados en porcentajes respecto a distintos totales:

Tabla de contingencia Sexo * Tabaquismo

			Tabaquismo		Total
			No fumador	Fumador	
Sexo	Mujer	Recuento	38	43	81
		% de Sexo	46.9%	53.1%	100.0%
		% de Tabaquismo	27.5%	69.4%	40.5%
		% del total	19.0%	21.5%	40.5%
	Varón	Recuento	100	19	119
		% de Sexo	84.0%	16.0%	100.0%
		% de Tabaquismo	72.5%	30.6%	59.5%
		% del total	50.0%	9.5%	59.5%
	Total		Recuento	138	62
		% de Sexo	69.0%	31.0%	100.0%
		% de Tabaquismo	100.0%	100.0%	100.0%
		% del total	69.0%	31.0%	100.0%

Globalmente, los resultados indican que la mayor parte de los sujetos son varones y no fuman (el 50%). Veamos ahora ejemplos de interpretación de *distribuciones condicionales*. Como se observa, hay un 27,5% de los no fumadores que son mujeres (frente al 53,1% de las fumadoras); y de los

varones el 84% son no fumadores. En cuanto a las mujeres, el 46,9% son no fumadoras y dentro de los fumadores el 69,4% son mujeres.

Las tablas de contingencia sólo tienen sentido para variables de tipo cualitativo (o categóricas). Si se desea representar la distribución conjunta de dos variables cuantitativas es necesario agrupar los valores de cada una de dichas variables en intervalos de la misma amplitud (para ello puede utilizarse el menú *Transformar -> Recodificar*). Una vez obtenidos los intervalos para *X* e *Y* puede elaborarse la tabla de contingencia siguiendo el mismo procedimiento que si las variables fueran cualitativas. El gráfico más adecuado para dos variables cuantitativas es el diagrama de dispersión, como vimos en el apartado de análisis de regresión lineal (ver figura 3.8. en página 26).

Una variable cualitativa y otra cuantitativa

Si queremos cruzar una variable *cualitativa* con una *cuantitativa*, se procede igual excepto que para resumir los datos de la cuantitativa hay que emplear estadísticos (p.e. sumas, medias, etc.). Veamos un ejemplo. Tomemos la variable *Fumar* (cualitativa con 2 categorías) y la variable *peso* que por ser cuantitativa la resumimos mediante su *MEDIA*. Para describir conjuntamente estas variables podemos elaborar un gráfico desde el menú *Gráficos -> Líneas -> Simple* (Resúmenes para grupos de casos):

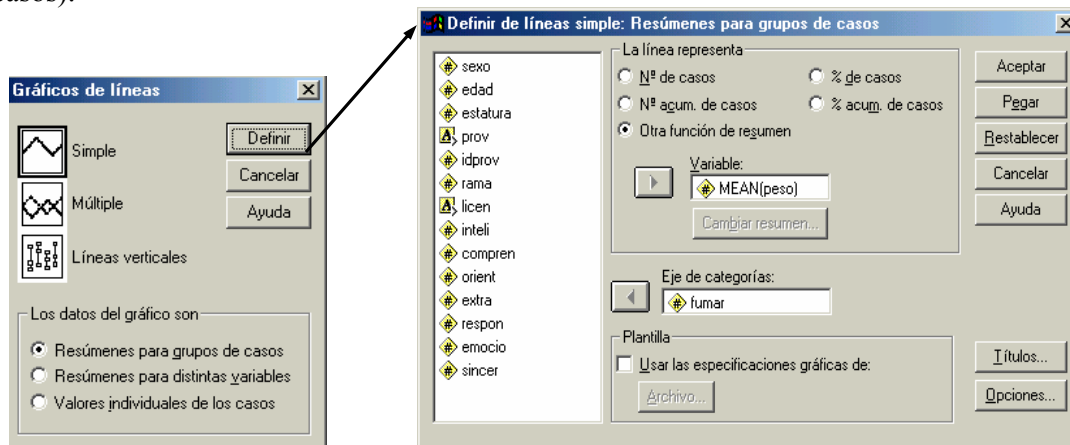
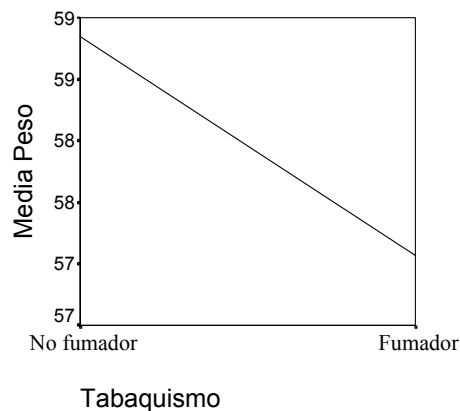


Figura 3.16

Con estas selecciones se obtiene el siguiente resultado en el visor:



Este gráfico representa la media de la variable *peso* para fumadores y no fumadores. Como se observa, los sujetos fumadores tienden a tener un promedio de peso menor (57 kg.) que los que no fuman (59 kg.).

Una variable cuantitativa y dos cualitativas

Es frecuente la representación de una variable dependiente cuantitativa en función de dos independientes de tipo cualitativo. En estos casos también se puede representar gráficamente la relación entre estas tres variables.

Veamos un ejemplo igual al anterior (variable cualitativa Fumar: 2 categorías y variable cuantitativa peso que se expresa mediante su MEDIA) pero añadiendo la variable sexo (cualitativa con 2 categorías). Para describir conjuntamente estas tres variables podemos elaborar un gráfico desde el menú *Gráficos -> Líneas -> Múltiple* (Resúmenes para grupos de casos). Es decir:

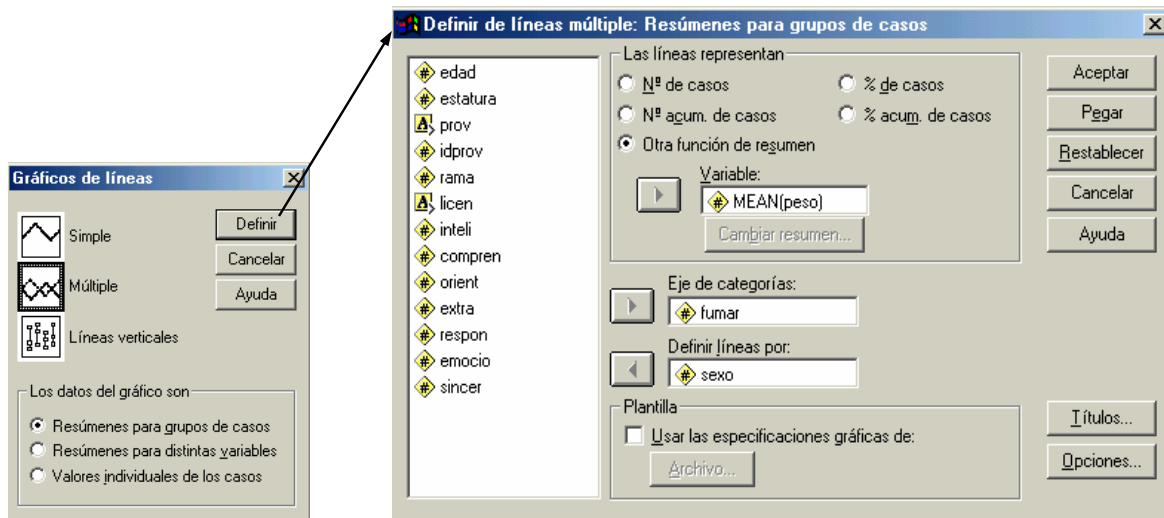
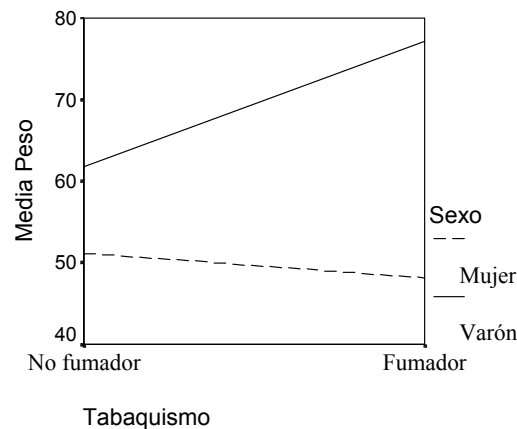


Figura 3.17

Con estas selecciones se obtiene el siguiente resultado en el visor:



Con la gráfica anterior observábamos que los sujetos no fumadores de la muestra tienen un peso medio mayor que los fumadores. Lo que nos indica esta nueva gráfica es lo mismo pero segmentado para la muestra de varones (línea continua) y la de mujeres (línea discontinua). Como puede observarse, en las mujeres se observa que efectivamente el promedio del peso es menor para las fumadoras. Sin embargo, en los varones se produce el efecto contrario. Es decir, los fumadores tienen un promedio en peso mucho mayor que los no fumadores. Por tanto, de no haber elaborado esta gráfica, se hubiera concluido algo erróneo para la muestra de varones. De todos modos, esta interpretación es solamente descriptiva pues para concluir que el género afecta al peso de los fumadores y no fumadores, sería necesario emplear pruebas inferenciales. En el próximo curso se estudiarán las pruebas estadísticas que determinan si cada variable independiente (y ambas en intersección) tiene algún efecto sobre la variable dependiente.

EJERCICIOS:

1. Se desea comparar las características físicas de los sujetos evaluados (todos ellos Españoles) con las de los Norteamericanos. Para ello, utilizaremos algunos datos de las prácticas: la variable *estatura* y la variable *peso*.

El objetivo es comparar nuestros datos descriptivos en estas variables con los de una muestra norteamericana de similares características.

Tenemos el problema de que en Estados Unidos el *peso* y la *estatura* se miden en escalas diferentes a las nuestras. En concreto, el peso en libras y la estatura en pulgadas, donde:

$$\begin{aligned} 1 \text{ LIBRA} &= 453,6 \text{ gramos} \\ 1 \text{ PULGADA} &= 2,54 \text{ cm} \end{aligned}$$

Lo primero es obtener los datos de ambas variables en la escala de medida norteamericana. Después hay que elaborar un informe con los estadísticos descriptivos.

EL INFORME ha de incluir lo siguiente

- a) La media y la varianza de las variables *estatura* y *peso* en la escala norteamericana
 - b) Un gráfico que dé cuenta de la variabilidad en cada una de las variables para los varones y las mujeres
 - c) Información sobre la forma de la distribución (asimetría y curtosis)
 - d) Las puntuaciones típicas para la variable *estatura*
 - e) Una escala derivada T con media 100 y desviación típica 20 para la variable *estatura* y su representación gráfica
 - f) La covarianza entre ambas variables (calcularlo en formato español y norteamericano).
2. Calcule la media y la varianza para la variable que resulta de la suma de las variables *Inteligencia*, *comprensión verbal* y *orientación espacial*.
 3. Se desea predecir las puntuaciones en *estabilidad emocional* (*emocio*) a partir de una de las siguientes variables: *inteligencia*, *extraversión*, *responsabilidad* y *sinceridad*. Seleccione la variable más apropiada como predictora, justificando la elección.
 - a) Represente gráficamente la relación entre la variable predictora escogida y el criterio.
 - b) Obtenga la ecuación de regresión correspondiente (en directas, diferenciales y típicas).
 - c) Descomponga la varianza del criterio para el modelo anterior e interprete la bondad del modelo (de forma descriptiva y con el gráfico interactivo).
 - d) ¿Cuál es la proporción de varianza explicada de la variable *estabilidad emocional* a partir de la variable predictora?
 - e) Si un sujeto obtiene una puntuación de 6 en la variable predictora, ¿Cuál es su puntuación pronosticada en *estabilidad emocional*?
 - f) Si se quisiera predecir las puntuaciones en *estabilidad emocional* a partir de dos variables, ¿cuál añadiría? ¿por qué?. ¿Cuánto mejoraría la bondad del nuevo modelo?
 4. Obtenga la tabla de frecuencias conjuntas para las variables *sexo* y *rama* y la gráfica de barras con ambas variables.
 5. ¿Cuál es el porcentaje de varones de humanidades? ¿y el de los que dentro de los de enseñanzas técnicas, son mujeres?
 6. Elabore la misma tabla de frecuencias que en 4. pero separando entre fumadores y no fumadores
 7. Confeccione una gráfica en la que aparezca la distribución de frecuencias para la variable *edad* en varones y mujeres e interprete los resultados.
 8. Repita lo mismo que en los ejercicios 4 a 7 con las variables *fumar* y *rama* e interprete los resultados obtenidos.
 9. Confeccione una gráfica que represente la relación entre las variables *rama*, *sexo* y *responsabilidad*.

ANEXO I: Notación científica

En muchas ocasiones las tablas de resultados del SPSS incluyen datos en la denominada *notación científica*, debido a que los números escritos de esta forma ocupan menos espacio que en la notación tradicional. Por ejemplo:

Estadísticos descriptivos

	N	Mínimo	Máximo	Media	Desv. típ.
Peso	200	39.00	92.00	58.2950	10.3520
Estatura	200	1.52	1.93	1.6756	8.413E-02

La notación científica se utiliza para señalar la posición del primer valor decimal. Por ejemplo, la desviación típica para la variable **Estatura** aparece escrita en notación científica: 8.413E-02. Esta expresión significa lo siguiente en notación tradicional:

$$8.413E - 02 = 8.413 \cdot 10^{-2} = \frac{8.413}{100} = 0.08413$$

Del mismo modo, el número 5.975E+04 es $5.975 \cdot 10^4 = 59750$.

ANEXO II: Variables del Cuestionario

En este apartado se describe cada una de las variables que contiene el fichero de datos:

1. Sexo: variable nominal, donde: (1): varón; (0): mujer.
2. Edad: variable cuantitativa (medida en años)
3. Peso: variable cuantitativa (en kilogramos)
4. Estatura: variable cuantitativa (en metros)
5. Prov: variable nominal, indicadora de la provincia de residencia
6. Idprov: variable nominal que sirve de identificador numérico de la provincia de residencia:

1 = Almería	18 = Cuenca	35 = Alicante
2 = Cádiz	19 = Guadalajara	36 = Castellón
3 = Córdoba	20 = Toledo	37 = Valencia
4 = Granada	21 = Avila	38 = Badajoz
5 = Huelva	22 = Burgos	39 = Cáceres
6 = Jaén	23 = León	40 = La Coruña
7 = Málaga	24 = Palencia	41 = Lugo
8 = Sevilla	25 = Salamanca	42 = Ourense
9 = Huesca	26 = Segovia	43 = Pontevedra
10 = Teruel	27 = Soria	44 = Baleares
11 = Zaragoza	28 = Valladolid	45 = La Rioja
12 = Asturias	29 = Zamora	46 = Madrid
13 = Las Palmas	30 = Barcelona	47 = Melilla
14 = Santa Cruz de Tenerife	31 = Girona	48 = Murcia
15 = Cantabria	32 = Lleida	49 = Navarra
16 = Albacete	33 = Tarragona	50 = Arava
17 = Ciudad Real	34 = Ceuta	51 = Guipuzkoa
		52 = Bizkaia
7. Rama: variable nominal que indica el área de conocimiento: (0): ciencias experimentales y de la salud, (1): Ciencias sociales y jurídicas, (2): enseñanzas técnicas, (3): humanidades y (4) otros/varios.
8. Licen: variable nominal que indica la titulación del sujeto (licenciado en derecho, etc.).
9. Intelig: inteligencia general: capacidad cognitiva para resolver problemas en general.
10. Compren: comprensión verbal (o capacidad para comprender los mensajes transmitidos verbalmente).
11. Orient: orientación espacial (o capacidad para situarse en el espacio respecto a alguna referencia).
12. Extra: extraversión (valora la cantidad e intensidad de las relaciones personales).
13. Respon: responsabilidad (grado de organización y motivación de la conducta del sujeto).
14. Emocio: estabilidad emocional (refleja el nivel de ajuste emocional de la persona).
15. Sincer: sinceridad (o grado en que responde con honestidad). *
16. Fumar: variable nominal que indica si el sujeto fuma (1) o no fuma (0).

* Las variables 9, 10, 11, 12, 13, 14 y 15 se consideran aquí como puntuaciones (variables cuantitativas).

ANEXO III: Datos para las prácticas

Sujeto:	sexo	edad	peso	estatura	prov	idprov	rama	licen	inteli	compren	orient	extra	respon	emocio	sincer	fumar
1	0	34	46	1.63	MADRID	46	1	LDO. DERECHO	21.00	24.00	29.72	38.00	51.00	49.00	22.00	1
2	0	34	51	1.72	BURGOS	22	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	11.00	25.00	64.72	42.00	46.00	60.00	21.00	1
3	1	25	45	1.53	VALLADOLID	28	1	DDO. BIBLIOTECONOMIA Y DOCUMENTACION	18.00	26.00	13.27	36.00	43.00	49.00	26.00	0
4	1	26	64	1.75	PALENCIA	24	1	LDO. DERECHO	13.00	28.00	10.07	41.00	49.00	45.00	27.00	0
5	1	27	58	1.68	MADRID	46	2	PRIMER CICLO I. INDUSTRIAL	22.00	29.00	27.20	42.00	44.00	51.00	25.00	0
6	0	29	45	1.64	BURGOS	22	3	LDO. FILOLOGIA INGLESA	11.00	25.00	36.65	41.00	52.00	54.00	32.00	1
7	1	28	60	1.70	MADRID	46	2	I. T. TELECOMUNICACION	22.00	26.00	109.24	43.00	49.00	52.00	20.00	0
8	0	27	45	1.64	BURGOS	22	1	DDO. CC. EMPRESARIALES	23.00	30.00	29.82	40.00	43.00	40.00	24.00	1
9	1	33	60	1.70	VALLADOLID	28	3	LDO. FILOLOGIA INGLESA	18.00	29.00	16.77	36.00	42.00	44.00	26.00	0
10	1	28	54	1.63	BURGOS	22	1	LDO. FILOSOFIA Y LETRAS	13.00	17.00	90.01	42.00	47.00	46.00	17.00	0
11	1	28	54	1.63	LEON	23	0	LDO. CC. BIOLOGICAS	15.00	24.00	110.37	40.00	46.00	46.00	24.00	0
12	1	32	54	1.63	BURGOS	22	2	I. TELECOMUNICACION	16.00	23.00	20.67	35.00	37.00	44.00	23.00	0
13	1	35	61	1.71	PALENCIA	24	1	LDO. DERECHO	20.00	31.00	24.28	43.00	51.00	52.00	26.00	0
14	1	28	62	1.72	ZAMORA	29	0	LDO. CC. FISICAS	25.00	27.00	15.40	39.00	46.00	47.00	27.00	0
15	1	30	56	1.65	ASTURIAS	12	3	LDO. FILOLOGIA INGLESA	15.00	24.00	72.14	43.00	41.00	55.00	20.00	0
16	0	35	45	1.64	ALAVA	50	1	DDO. CC. EMPRESARIALES	23.00	26.00	86.93	33.00	44.00	46.00	30.00	1
17	0	34	39	1.57	BURGOS	22	1	LDO. CC. EMPRESARIALES	20.00	25.00	21.92	40.00	41.00	45.00	24.00	1
18	1	28	63	1.73	BARCELONA	30	1	LDO. DERECHO	16.00	20.00	106.99	44.00	41.00	43.00	17.00	0
19	1	26	58	1.67	ZAMORA	29	4	ENSEÑANZA MILITAR GRADO MEDIO	20.00	31.00	8.77	42.00	53.00	42.00	21.00	0
20	0	27	45	1.63	ASTURIAS	12	1	LDO. DERECHO	22.00	21.00	10.38	44.00	49.00	48.00	27.00	1
21	0	31	46	1.64	MADRID	46	1	DDO. CC. DE LA EDUCACION	15.00	25.00	85.80	41.00	43.00	46.00	26.00	1
22	1	26	52	1.60	MADRID	46	0	PRIMER CICLO CC. FISICAS	18.00	30.00	80.28	39.00	46.00	46.00	24.00	0
23	0	35	47	1.65	VIZCAYA	52	3	LDO. GEOGRAFIA / HISTORIA	12.00	23.00	88.11	47.00	52.00	48.00	21.00	1
24	0	33	47	1.65	BARCELONA	30	3	LDO. FILOLOGIA FRANCESA	11.00	18.00	102.41	45.00	47.00	51.00	25.00	1
25	0	34	47	1.65	ZARAGOZA	11	0	LDO. MEDICINA Y CIRUGIA	23.00	30.00	14.23	39.00	53.00	56.00	17.00	1
26	1	28	66	1.76	MADRID	46	0	PRIMER CICLO CC. FISICAS	28.00	29.00	13.86	50.00	45.00	56.00	32.00	0
27	1	31	61	1.70	MADRID	46	2	I. T. INDUSTRIAL	20.00	30.00	79.69	37.00	43.00	49.00	24.00	0
28	1	28	55	1.63	MADRID	46	2	ARQUITECTO TECNICO	16.00	23.00	12.83	30.00	36.00	60.00	12.00	0
29	0	30	42	1.59	MADRID	46	2	I. T. AERONAUTICO	28.00	26.00	15.62	36.00	42.00	46.00	13.00	1
30	1	25	55	1.63	ASTURIAS	12	1	DDO. CC. EMPRESARIALES	16.00	28.00	13.82	42.00	44.00	42.00	17.00	0
31	0	29	45	1.62	BARCELONA	30	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	21.00	29.00	31.04	39.00	40.00	53.00	21.00	1
32	0	26	46	1.63	MADRID	46	3	LDO. FILOLOGIA INGLESA	24.00	29.00	47.74	41.00	41.00	48.00	25.00	1
33	1	28	55	1.63	BURGOS	22	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	18.00	30.00	16.80	42.00	51.00	50.00	28.00	0
34	1	28	56	1.64	NAVARRA	49	2	PRIMER CICLO ARQUITECTO	16.00	21.00	49.12	37.00	43.00	54.00	25.00	0
35	1	28	50	1.57	ASTURIAS	12	2	I. T. MINAS	16.00	24.00	13.39	39.00	41.00	42.00	17.00	0
36	0	30	55	1.73	BURGOS	22	1	LDO. FILOSOFIA Y LETRAS	23.00	25.00	46.26	44.00	54.00	47.00	13.00	1

37	0	27	49	1.66	VALLADOLID	28	2	I. T. FORESTAL	22.00	30.00	84.55	43.00	44.00	47.00	22.00	1
38	1	32	63	1.72	ASTURIAS	12	1	DDO. MAGISTERIO	16.00	31.00	13.36	49.00	54.00	58.00	34.00	0
39	0	31	49	1.66	ASTURIAS	12	1	LDO. CC. ECONOMICAS Y EMPRESARIALES	21.00	31.00	31.98	41.00	46.00	47.00	25.00	1
40	0	29	57	1.75	VALLADOLID	28	1	PRIMER CICLO DERECHO	15.00	22.00	86.08	41.00	43.00	49.00	24.00	1
41	1	38	64	1.73	MADRID	46	0	PRIMER CICLO CC. BIOLÓGICAS	14.00	22.00	13.79	32.00	42.00	39.00	24.00	0
42	1	26	58	1.66	PONTEVEDRA	43	0	LDO. CC. EXACTAS	23.00	28.00	6.90	43.00	51.00	54.00	23.00	0
43	0	32	39	1.54	MADRID	46	2	I. AERONAUTICO	21.00	30.00	39.79	40.00	45.00	45.00	29.00	1
44	0	29	53	1.70	BURGOS	22	3	LDO. FILOLOGIA INGLESA	13.00	8.00	125.98	44.00	45.00	43.00	25.00	1
45	0	24	50	1.66	BURGOS	22	1	DDO. CC. ECONOMICAS Y EMPRESARIALES	14.00	19.00	114.18	39.00	40.00	46.00	20.00	1
46	1	31	53	1.60	BALEARES	44	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	25.00	31.00	9.44	42.00	44.00	45.00	24.00	0
47	1	28	66	1.75	MADRID	46	2	I. T. AERONAUTICO	21.00	32.00	15.24	39.00	46.00	49.00	18.00	0
48	0	25	45	1.60	MADRID	46	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	17.00	29.00	12.63	49.00	45.00	56.00	22.00	1
49	1	31	60	1.68	MADRID	46	1	LDO. CC. POLITICAS / SOCIOLOGIA	20.00	24.00	18.77	38.00	44.00	48.00	13.00	0
50	1	31	61	1.69	MADRID	46	1	LDO. CC. INFORMACION	22.00	24.00	35.92	38.00	41.00	45.00	25.00	0
51	1	28	55	1.62	VALLADOLID	28	0	LDO. CC. FISICAS	16.00	25.00	38.11	41.00	47.00	47.00	27.00	0
52	1	37	55	1.62	MADRID	46	1	LDO. DERECHO	12.00	16.00	52.67	36.00	34.00	43.00	24.00	0
53	1	26	63	1.71	ASTURIAS	12	1	DDO. EDUCACION FISICA	16.00	26.00	37.72	40.00	43.00	47.00	22.00	0
54	1	30	63	1.71	MADRID	46	1	PRIMER CICLO ADMINISTRAC. Y DIRECCION	17.00	29.00	8.48	41.00	45.00	47.00	21.00	0
55	1	26	64	1.72	MADRID	46	1	LDO. DERECHO	16.00	24.00	8.52	46.00	46.00	47.00	25.00	0
56	1	28	52	1.58	VALLADOLID	28	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	19.00	27.00	10.72	43.00	49.00	50.00	30.00	0
57	0	30	45	1.60	MADRID	46	2	I. AERONAUTICO	16.00	26.00	104.50	43.00	44.00	50.00	20.00	1
58	0	33	53	1.69	BURGOS	22	0	LDO. VETERINARIA	24.00	32.00	21.88	46.00	50.00	53.00	24.00	1
59	1	29	59	1.66	VALLADOLID	28	1	LDO. CC. ECONOMICAS Y EMPRESARIALES	24.00	25.00	14.74	36.00	47.00	47.00	24.00	0
60	1	30	59	1.66	MADRID	46	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	17.00	25.00	52.75	39.00	38.00	47.00	13.00	0
61	1	29	61	1.68	ASTURIAS	12	3	PRIMER CICLO FILOLOGIA INGLESA	22.00	18.00	13.06	38.00	46.00	37.00	18.00	0
62	0	28	47	1.62	MADRID	46	1	LDO. CC. INFORMACION	24.00	29.00	29.17	44.00	48.00	47.00	26.00	1
63	1	30	56	1.62	BURGOS	22	1	LDO. CC. ECONOMICAS Y EMPRESARIALES	12.00	17.00	113.11	47.00	48.00	42.00	25.00	0
64	1	42	57	1.63	MADRID	46	2	I. T. TOPOGRAFIA	24.00	25.00	8.64	47.00	43.00	53.00	27.00	0
65	0	25	47	1.62	MADRID	46	1	LDO. DERECHO	16.00	32.00	37.70	46.00	57.00	57.00	33.00	1
66	0	34	55	1.71	MADRID	46	0	LDO. CC. FISICAS	15.00	19.00	86.18	46.00	56.00	57.00	40.00	1
67	0	23	49	1.64	MADRID	46	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	15.00	16.00	18.59	45.00	50.00	41.00	22.00	0
68	0	29	43	1.57	MADRID	46	1	LDO. PERIODISMO	14.00	23.00	73.35	43.00	53.00	49.00	22.00	0
69	1	35	66	1.73	MADRID	46	1	LDO. CC. POLITICAS / SOCIOLOGIA	26.00	30.00	10.71	41.00	42.00	47.00	17.00	0
70	1	26	60	1.66	MADRID	46	1	LDO. ADMINISTRAC. Y DIRECCION EMPRESAS	19.00	25.00	35.59	44.00	48.00	49.00	23.00	0
71	0	30	56	1.72	ASTURIAS	12	1	DDO. CC. EMPRESARIALES	18.00	28.00	42.52	45.00	51.00	49.00	24.00	0
72	0	24	50	1.65	MADRID	46	1	LDO. CC. ECONOMICAS	28.00	30.00	10.77	31.00	48.00	56.00	12.00	0
73	1	41	60	1.66	ASTURIAS	12	4	CAPITAN MARINA MERCANTE	19.00	21.00	31.30	40.00	41.00	46.00	20.00	0
74	1	32	57	1.62	VALLADOLID	28	0	LDO. VETERINARIA	18.00	31.00	14.67	50.00	55.00	54.00	19.00	0
75	1	28	66	1.72	ASTURIAS	12	0	PRIMER CICLO CC. QUIMICAS	14.00	21.00	56.26	44.00	35.00	36.00	23.00	0

76	0	31	57	1.73	MADRID	46	1	LDO. CC. INFORMACION	10.00	19.00	91.69	43.00	48.00	54.00	31.00	0
77	1	29	60	1.65	LEON	23	1	DDO. MAGISTERIO	18.00	12.00	91.57	48.00	45.00	51.00	28.00	0
78	1	27	60	1.65	ASTURIAS	12	2	I. T. INDUSTRIAL	15.00	22.00	10.91	34.00	42.00	46.00	24.00	0
79	1	24	67	1.73	ASTURIAS	12	2	PRIMER CICLO MARINA CIVIL	19.00	32.00	14.93	39.00	46.00	37.00	26.00	0
80	1	25	56	1.60	ASTURIAS	12	2	DDO. MARINA CIVIL	16.00	30.00	5.97	36.00	42.00	50.00	21.00	0
81	1	32	63	1.68	ASTURIAS	12	1	DDO. MAESTRO LENGUA EXTRANJERA	22.00	29.00	9.01	40.00	41.00	50.00	25.00	0
82	0	29	51	1.66	ASTURIAS	12	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	17.00	30.00	35.31	37.00	45.00	47.00	26.00	1
83	0	31	52	1.67	ASTURIAS	12	3	LDO. FILOLOGIA INGLESA	16.00	25.00	109.89	45.00	48.00	52.00	26.00	0
84	1	30	63	1.68	ASTURIAS	12	2	LDO. MARINA CIVIL	23.00	24.00	17.02	40.00	50.00	51.00	29.00	0
85	1	28	63	1.68	ASTURIAS	12	1	DDO. RELACIONES LABORALES	15.00	32.00	72.51	41.00	45.00	43.00	25.00	0
86	1	26	64	1.68	ASTURIAS	12	0	PRIMER CICLO FARMACIA	16.00	24.00	19.42	47.00	48.00	54.00	27.00	0
87	1	26	64	1.68	LEON	23	1	DDO. GESTION Y ADMINISTRACION PUBLICA	14.00	17.00	18.45	40.00	46.00	44.00	26.00	0
88	0	22	52	1.67	ASTURIAS	12	3	PRIMER CICLO FILOLOGIA INGLESA	18.00	28.00	142.29	39.00	33.00	46.00	27.00	0
89	1	29	58	1.61	MADRID	46	1	LDO. CC. ECONOMICAS	19.00	28.00	25.67	48.00	39.00	55.00	20.00	0
90	1	29	52	1.54	MADRID	46	1	LDO. CC. ECONOMICAS Y EMPRESARIALES	19.00	20.00	38.32	48.00	45.00	48.00	22.00	0
91	1	29	59	1.62	MADRID	46	1	LDO. DERECHO	14.00	26.00	17.41	45.00	45.00	58.00	20.00	0
92	0	30	52	1.67	MADRID	46	3	LDO. FILOLOGIA INGLESA	22.00	30.00	13.53	41.00	47.00	46.00	22.00	0
93	0	28	52	1.67	MADRID	46	1	LDO. DERECHO	11.00	28.00	77.23	45.00	47.00	54.00	23.00	1
94	1	38	66	1.70	MADRID	46	1	LDO. PSICOPELAGOGIA	13.00	29.00	69.92	42.00	45.00	48.00	20.00	0
95	1	42	60	1.63	MADRID	46	4	DDO. TURISMO	14.00	23.00	14.98	41.00	51.00	54.00	25.00	0
96	1	28	60	1.63	MADRID	46	1	LDO. CC. ECONOMICAS Y EMPRESARIALES	15.00	27.00	13.29	34.00	41.00	49.00	21.00	0
97	1	37	60	1.63	MADRID	46	2	I. AERONAUTICO	19.00	32.00	8.84	37.00	45.00	43.00	24.00	0
98	0	27	40	1.53	MADRID	46	2	I. T. TELECOMUNICACION	20.00	25.00	117.38	40.00	45.00	45.00	20.00	0
99	0	35	42	1.55	BARCELONA	30	1	LDO. CC. INFORMACION	15.00	27.00	55.62	38.00	47.00	53.00	19.00	0
100	1	33	60	1.63	MADRID	46	0	DDO. ESTADISTICA	13.00	25.00	10.63	37.00	43.00	44.00	16.00	0
101	1	27	61	1.63	MADRID	46	0	LDO. FARMACIA	17.00	20.00	50.09	39.00	53.00	55.00	18.00	0
102	1	30	62	1.64	MADRID	46	3	LDO. BELLAS ARTES	24.00	30.00	12.79	40.00	42.00	43.00	24.00	0
103	1	31	58	1.59	MADRID	46	3	LDO. BELLAS ARTES	23.00	23.00	9.29	41.00	49.00	52.00	24.00	0
104	0	27	40	1.52	MADRID	46	2	I. AERONAUTICO	26.00	29.00	74.67	44.00	45.00	53.00	28.00	1
105	1	28	59	1.60	MADRID	46	2	PRIMER CICLO I. CAMINOS, CANAL Y PUERTO	25.00	29.00	8.13	42.00	47.00	56.00	22.00	0
106	1	29	60	1.61	MADRID	46	0	LDO. CC. QUIMICAS	14.00	23.00	10.37	39.00	42.00	46.00	21.00	0
107	0	23	53	1.67	MADRID	46	1	PRIMER CICLO CC. ECONOMICAS	26.00	29.00	24.93	42.00	44.00	51.00	29.00	0
108	1	25	67	1.69	MADRID	46	0	LDO. CC. FISICAS	24.00	27.00	11.64	47.00	44.00	53.00	22.00	0
109	0	24	53	1.67	MADRID	46	1	DDO. MAESTRO LENGUA EXTRANJERA	18.00	28.00	19.99	43.00	47.00	51.00	23.00	0
110	1	24	68	1.70	MADRID	46	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	13.00	17.00	101.06	46.00	53.00	50.00	23.00	0
111	1	26	62	1.63	MADRID	46	0	LDO. CC. EXACTAS	26.00	28.00	11.62	53.00	54.00	59.00	31.00	0
112	1	30	66	1.67	VIZCAYA	52	1	DDO. CC. EMPRESARIALES	17.00	31.00	15.09	39.00	43.00	45.00	23.00	0
113	1	33	57	1.56	VIZCAYA	52	1	LDO. DERECHO	20.00	25.00	29.45	47.00	50.00	45.00	22.00	0
114	1	29	70	1.71	ZAMORA	29	1	PRIMER CICLO DERECHO	20.00	23.00	17.60	36.00	45.00	48.00	26.00	0

115	1	28	65	1.65	ZAMORA	29	3	LDO. BELLAS ARTES	23.00	29.00	9.89	47.00	55.00	58.00	22.00	0
116	0	34	41	1.53	MADRID	46	1	DDO. CC. ECONOMICAS Y EMPRESARIALES	14.00	19.00	48.41	39.00	49.00	43.00	19.00	0
117	1	29	60	1.59	MADRID	46	1	LDO. CC. POLITICAS / SOCIOLOGIA	10.00	22.00	102.72	39.00	44.00	43.00	25.00	0
118	1	29	71	1.71	VALLADOLID	28	2	PRIMER CICLO I. INDUSTRIAL	25.00	22.00	16.66	44.00	56.00	50.00	29.00	0
119	1	29	68	1.67	VALLADOLID	28	1	DDO. RELACIONES LABORALES	13.00	24.00	15.66	44.00	41.00	46.00	17.00	0
120	1	28	63	1.61	VALLADOLID	28	0	LDO. CC. FISICAS	21.00	21.00	15.70	38.00	47.00	51.00	34.00	0
121	1	25	75	1.88	VALLADOLID	28	0	DDO. ENFERMERIA	21.00	28.00	84.72	40.00	52.00	47.00	24.00	0
122	0	30	42	1.54	BARCELONA	30	1	LDO. CC. COMUNICACION	13.00	32.00	46.94	39.00	45.00	50.00	21.00	1
123	1	46	65	1.83	MURCIA	48	1	DDO. PROFESORADO EGB	20.00	20.00	19.77	40.00	45.00	51.00	30.00	0
124	0	40	55	1.69	ALICANTE	35	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	18.00	26.00	93.36	38.00	42.00	41.00	28.00	0
125	1	28	57	1.71	PALENCIA	24	2	I. T. INDUSTRIAL	14.00	26.00	12.38	40.00	44.00	42.00	25.00	0
126	1	26	70	1.86	MADRID	46	1	LDO. DERECHO	19.00	29.00	9.05	42.00	50.00	50.00	24.00	0
127	1	40	63	1.77	NAVARRA	49	3	LDO. HUMANIDADES	11.00	19.00	85.72	35.00	42.00	37.00	24.00	0
128	1	32	58	1.71	VALLADOLID	28	0	LDO. CC. QUIMICAS	25.00	26.00	15.05	42.00	48.00	46.00	26.00	0
129	0	31	45	1.57	BURGOS	22	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	21.00	27.00	76.81	39.00	45.00	46.00	23.00	1
130	1	32	66	1.80	BURGOS	22	0	LDO. VETERINARIA	16.00	30.00	37.34	45.00	45.00	45.00	26.00	0
131	0	28	58	1.72	BURGOS	22	0	LDO. CC. Y TECNOLOGIA DE LOS ALIMENTOS	21.00	28.00	83.78	38.00	43.00	47.00	24.00	0
132	0	29	47	1.59	BURGOS	22	0	LDO. CC. QUIMICAS	12.00	30.00	30.45	38.00	50.00	47.00	24.00	0
133	1	27	69	1.83	MADRID	46	2	ARQUITECTO TECNICO	17.00	18.00	6.34	36.00	43.00	44.00	25.00	0
134	0	28	47	1.59	BARCELONA	30	1	DDO. BIBLIOTECONOMIA Y DOCUMENTACION	17.00	31.00	71.44	47.00	52.00	56.00	17.00	1
135	0	29	41	1.52	BURGOS	22	1	DDO. CC. EMPRESARIALES	22.00	19.00	80.90	40.00	45.00	44.00	25.00	1
136	1	25	73	1.87	BURGOS	22	4	TECN. EMPRESAS Y ACTIVIDADES TURISTICAS	13.00	21.00	23.74	40.00	42.00	45.00	26.00	0
137	0	24	54	1.67	MADRID	46	1	LDO. CC. ECONOMICAS	18.00	25.00	21.85	47.00	46.00	46.00	25.00	1
138	0	30	48	1.60	BURGOS	22	1	LDO. EDUCACION FISICA	24.00	27.00	23.66	40.00	45.00	51.00	29.00	0
139	1	26	65	1.77	ASTURIAS	12	3	LDO. GEOGRAFIA / HISTORIA	20.00	29.00	45.13	29.00	44.00	50.00	20.00	0
140	1	29	56	1.66	ASTURIAS	12	1	LDO. DERECHO	20.00	32.00	86.94	47.00	52.00	60.00	29.00	0
141	1	27	66	1.77	ASTURIAS	12	2	DDO. MARINA CIVIL	20.00	26.00	17.02	48.00	53.00	60.00	18.00	0
142	1	33	63	1.73	LAS PALMAS	13	2	DDO. NAUTICA	18.00	22.00	23.72	40.00	38.00	46.00	17.00	0
143	0	30	61	1.75	LAS PALMAS	13	1	LDO. DERECHO	13.00	25.00	23.51	47.00	46.00	47.00	26.00	1
144	1	33	64	1.73	LAS PALMAS	13	1	PRIMER CICLO PSICOLOGIA	20.00	25.00	9.64	37.00	44.00	46.00	24.00	0
145	1	29	73	1.83	SANTA CRUZ	14	0	LDO. CC. FISICAS	19.00	27.00	19.28	40.00	46.00	44.00	23.00	0
146	1	30	73	1.83	SANTA CRUZ	14	2	DDO. MARINA CIVIL	17.00	24.00	34.84	41.00	45.00	51.00	23.00	0
147	1	27	83	1.93	BALEARES	44	1	DDO. CC. EMPRESARIALES	23.00	30.00	40.44	39.00	46.00	48.00	24.00	0
148	1	32	83	1.93	LEON	23	3	LDO. GEOGRAFIA / HISTORIA	16.00	29.00	17.49	42.00	42.00	60.00	25.00	0
149	0	32	50	1.62	LEON	23	2	I. T. INDUSTRIAL	23.00	31.00	17.80	41.00	44.00	44.00	19.00	0
150	0	31	50	1.62	MADRID	46	1	LDO. CC. ECONOMICAS Y EMPRESARIALES	18.00	20.00	87.70	53.00	54.00	59.00	22.00	0
151	0	29	51	1.63	MADRID	46	1	LDO. DERECHO	10.00	27.00	80.15	42.00	49.00	51.00	18.00	1
152	0	27	51	1.63	LEON	23	3	LDO. TRADUCCION E INTERPRETACION	17.00	23.00	86.35	43.00	49.00	46.00	27.00	1
153	1	30	57	1.62	LEON	23	3	LDO. FILOLOGIA INGLES	21.00	24.00	24.91	33.00	43.00	44.00	23.00	0

154	0	29	45	1.56	LEON	23	1	LDO. ADMINISTRAC. Y DIRECCION EMPRESAS	20.00	28.00	114.86	46.00	58.00	51.00	25.00	1
155	0	32	58	1.71	VALENCIA	37	1	DDO. CC. EMPRESARIALES	15.00	28.00	121.23	47.00	39.00	49.00	16.00	0
156	1	37	57	1.62	ASTURIAS	12	2	PRIMER CICLO MARINA CIVIL	27.00	31.00	41.25	40.00	42.00	44.00	20.00	0
157	0	37	46	1.57	ASTURIAS	12	3	LDO. FILOLOGIA HISPANICA	29.00	28.00	62.63	40.00	52.00	50.00	13.00	0
158	1	36	69	1.75	MADRID	46	2	DDO. MARINA CIVIL	9.00	25.00	14.66	42.00	44.00	43.00	19.00	0
159	1	28	65	1.70	ASTURIAS	12	3	LDO. TRADUCCION E INTERPRETACION	26.00	29.00	18.02	37.00	44.00	43.00	24.00	0
160	1	34	73	1.79	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	18.00	24.00	32.66	47.00	50.00	59.00	15.00	1
161	1	29	73	1.78	ASTURIAS	12	0	LDO. CC. QUIMICAS	24.00	31.00	13.60	46.00	50.00	57.00	23.00	1
162	1	31	73	1.78	BARCELONA	30	0	LDO. VETERINARIA	26.00	30.00	9.61	35.00	43.00	48.00	24.00	1
163	0	30	60	1.73	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	21.00	31.00	46.77	41.00	41.00	44.00	26.00	0
164	0	29	62	1.75	ASTURIAS	12	1	DDO. CC. EMPRESARIALES	15.00	26.00	46.57	39.00	43.00	45.00	25.00	0
165	0	29	57	1.69	ASTURIAS	12	2	I. INDUSTRIAL	19.00	20.00	38.36	36.00	49.00	57.00	29.00	1
166	0	30	51	1.62	ASTURIAS	12	1	LDO. DERECHO	15.00	23.00	109.19	43.00	46.00	44.00	31.00	0
167	0	28	51	1.62	ASTURIAS	12	1	LDO. PSICOLOGIA	21.00	27.00	22.62	40.00	44.00	44.00	28.00	0
168	0	28	52	1.63	ASTURIAS	12	0	LDO. CC. QUIMICAS	17.00	17.00	53.51	46.00	55.00	56.00	43.00	1
169	1	27	86	1.93	ASTURIAS	12	2	LDO. MARINA CIVIL	17.00	29.00	17.02	45.00	41.00	47.00	22.00	1
170	0	30	59	1.71	ASTURIAS	12	1	DDO. GRADUADO SOCIAL	11.00	17.00	104.56	39.00	43.00	45.00	27.00	0
171	1	24	69	1.73	ASTURIAS	12	1	DDO. MAGISTERIO	16.00	21.00	15.61	51.00	55.00	54.00	29.00	1
172	0	26	53	1.64	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	19.00	26.00	49.23	49.00	46.00	47.00	20.00	1
173	0	28	53	1.64	ASTURIAS	12	1	DDO. MAGISTERIO	14.00	30.00	98.59	40.00	43.00	47.00	25.00	0
174	0	25	47	1.57	ASTURIAS	12	2	I. T. INFORMATICA	17.00	29.00	29.01	40.00	49.00	50.00	24.00	0
175	1	26	76	1.81	ASTURIAS	12	1	DDO. GESTION Y ADMINISTRACION PUBLICA	18.00	25.00	9.50	42.00	49.00	52.00	25.00	1
176	0	25	47	1.57	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	23.00	26.00	49.60	45.00	59.00	47.00	14.00	0
177	1	25	76	1.81	MADRID	46	0	LDO. CC. EXACTAS	19.00	32.00	7.80	46.00	54.00	51.00	20.00	1
178	0	25	47	1.57	MADRID	46	0	PRIMER CICLO CC. FISICAS	26.00	30.00	29.81	34.00	46.00	45.00	13.00	0
179	1	30	73	1.77	ASTURIAS	12	1	LDO. DERECHO	19.00	29.00	32.27	46.00	52.00	46.00	26.00	1
180	1	34	71	1.74	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	15.00	27.00	101.57	37.00	37.00	48.00	22.00	1
181	1	31	80	1.83	ASTURIAS	12	2	I. T. MINAS	20.00	31.00	11.65	41.00	52.00	52.00	34.00	1
182	1	31	83	1.85	ASTURIAS	12	0	PRIMER CICLO CC. QUIMICAS	21.00	27.00	14.35	41.00	48.00	52.00	23.00	1
183	1	28	73	1.73	ASTURIAS	12	2	I. T. MINAS	20.00	18.00	30.40	37.00	45.00	50.00	27.00	1
184	1	27	78	1.78	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	21.00	25.00	13.68	37.00	42.00	43.00	24.00	1
185	0	25	60	1.72	ASTURIAS	12	2	DDO. NAUTICA	24.00	26.00	85.70	46.00	48.00	49.00	23.00	0
186	0	26	55	1.66	MADRID	46	3	LDO. FILOLOGIA INGLESА	16.00	26.00	22.16	45.00	53.00	45.00	23.00	0
187	1	25	85	1.86	ASTURIAS	12	3	PRIMER CICLO HISTORIA DEL ARTE	15.00	24.00	10.94	35.00	41.00	44.00	24.00	1
188	1	33	68	1.66	MADRID	46	0	LDO. CC. BIOLOGICAS	27.00	31.00	23.18	33.00	45.00	44.00	14.00	1
189	1	30	83	1.83	ASTURIAS	12	1	DDO. CC. EMPRESARIALES	25.00	29.00	12.49	35.00	50.00	49.00	15.00	1
190	0	29	49	1.59	MADRID	46	0	LDO. CC. QUIMICAS	16.00	32.00	55.12	44.00	46.00	51.00	26.00	1
191	0	28	49	1.59	ASTURIAS	12	3	LDO. FILOLOGIA INGLESА	14.00	23.00	102.19	36.00	50.00	45.00	22.00	0
192	0	26	49	1.59	ASTURIAS	12	1	LDO. PERIODISMO	20.00	27.00	78.66	40.00	41.00	44.00	14.00	0

193	0	27	56	1.67	ASTURIAS	12	1	DDO. CC. EMPRESARIALES	16.00	26.00	9.99	46.00	44.00	47.00	18.00	0
194	1	26	92	1.93	ASTURIAS	12	3	LDO. FILOLOGIA INGLESA	14.00	25.00	19.29	49.00	54.00	56.00	23.00	1
195	1	26	73	1.73	ASTURIAS	12	4	DDO. TURISMO	18.00	22.00	54.34	43.00	52.00	56.00	14.00	1
196	0	25	50	1.60	ALAVA	50	1	LDO. CC. INFORMACION	25.00	27.00	22.16	39.00	46.00	49.00	24.00	0
197	1	24	82	1.82	ASTURIAS	12	2	LDO. MARINA CIVIL	18.00	30.00	9.56	46.00	45.00	57.00	17.00	1
198	0	26	51	1.61	ASTURIAS	12	3	LDO. FILOLOGIA INGLESA	25.00	29.00	21.60	43.00	49.00	52.00	17.00	0
199	1	24	78	1.78	ASTURIAS	12	0	LDO. CC. BIOLOGICAS	21.00	26.00	9.18	39.00	37.00	48.00	17.00	0
200	0	27	45	1.54	MADRID	46	0	LDO. CC. BIOLOGICAS	14.00	21.00	43.92	44.00	45.00	47.00	24.00	0

ANEXO IV: EJERCICIO PRÁCTICO A RESOLVER CON EL SPSS

En este anexo se incluyen 2 prácticas que se realizarán en las clases prácticas de Análisis de datos en Psicología I con el SPSS.

Cuando tu profesor te lo indique, tendrás que llevar a clase este documento y contestar a las preguntas que se incluyen en él.

La práctica se refiere a UN CUESTIONARIO diferente al que estamos utilizando en las clases prácticas. Se trata de una *Encuesta sobre Ocio y Tiempo libre* que consta de 20 preguntas y fue contestada por 717 jóvenes. Las preguntas de dicho cuestionario se incluyen en este anexo y los datos correspondientes se encuentran en el fichero **ejercicios.sav** que se puede descargar desde:

http://www.uam.es/personal_pdi/psicologia/carmenx/Practicas2.html

Para resolver esta práctica en las clases, es imprescindible que lleves este documento y tu diquette 3 ½ con el fichero **ejercicios.sav**.

- **IMPORTANTE:** Antes de contestar a cada pregunta de las prácticas debes leer detenidamente la denominación (en mayúsculas) y descripción de cada pregunta del cuestionario (ver las dos siguientes páginas). Por ejemplo, la variable OCIO oscila entre 1 y 6 (Donde 2: hacer deporte, 6: leer, etc.). No debe confundirse con la variable HORAS, que se refiere a las horas destinadas al ocio por semana.

CUESTIONARIO

Durante el año 1999, se realizó una Encuesta sobre la Juventud Española. La muestra era de 10.000 jóvenes, de la que hemos seleccionado tan sólo aquellos entrevistados con estudios universitarios (717 sujetos). De las 60 preguntas de que constaba el cuestionario, el archivo de datos (**ejercicios.sav**) sólo recoge las 16 preguntas que aparecen a continuación, tal como se plantearon a los entrevistados.

Los nombres que tienen las variables en el archivo de datos aparecen **EN MAYÚSCULA y NEGRITA**.

A partir de la práctica nº 2, el alumno recibirá una hoja con una serie de preguntas que tendrá que responder, utilizando SPSS, y que versarán sobre los conceptos utilizados en dicha práctica.

OCIO

P.1 Para empezar y refiriéndonos a lo que haces en tus días de ocio, quisiéramos saber ¿cuál es la actividad que más te gusta hacer fuera de casa cuando dispones de tiempo libre?

- | | | |
|--------------------------------|---|-------------|
| - Beber, ir de copas, bailar | 1 | |
| - Hacer deporte | 2 | |
| - Viajes, excursiones | 3 | OCIO |
| - Ir al cine, al teatro | 4 | |
| - Ir a museos, ir a conciertos | 5 | |
| - Leer | 6 | |

P.2 En total, ¿cuántas horas libres tienes a la semana para tu ocio o diversión?

_____ **HORAS**

P.3 Aproximadamente y por término medio, ¿cuántas horas semanales dedicas a ver la televisión?

_____ **TV**

FAMILIA

P.4 ¿Cómo dirías que ha sido tu infancia, la definirías como...

- | | | |
|-------------------------------------|---|-----------------|
| - La etapa más feliz de tu vida | 1 | |
| - Una etapa más feliz que otras | 2 | |
| - Una etapa igual de feliz que otra | 3 | |
| - Una etapa menos feliz que otras | 4 | INFANCIA |
| - La etapa menos feliz de tu vida | 5 | |

P.5 ¿Podrías decirme cuál es tu grado de dedicación en las siguientes tareas de tu hogar?

1. No suelo hacerlo nunca
 2. Sólo lo hago en ocasiones
 3. Comparto esta tarea con otra/s personas
 4. Recae en mí toda la responsabilidad, y suelo hacerlo siempre
- | | | | | | |
|---|---|---|---|---|----------------|
| - Hacer la cama | 1 | 2 | 3 | 4 | CAMA |
| - Limpiar la casa | 1 | 2 | 3 | 4 | LIMPIAR |
| - Cuidar de los hijos o hermanos pequeños | 1 | 2 | 3 | 4 | NIÑOS |

P.6 ¿Cuántos hijos crees que llegarás a tener?

P.7 Y en el plano ideal, ¿cuántos hijos te gustaría llegar a tener?

- | | <u>Llegará a tener</u> | <u>Le gustaría</u> | |
|-----------|------------------------|--------------------|-----------------------------|
| - Uno | 1 | 1 | |
| - Dos | 2 | 2 | |
| - Tres | 3 | 3 | HIJOS1 HIJOS2 |
| - Cuatro | 4 | 4 | |
| - Cinco | 5 | 5 | |
| - Ninguno | 0 | 0 | |

P.8 ¿En cuál de las siguientes situaciones te encuentras?

- | | | |
|--|---|-------------|
| - Tienes novio/a formal | 1 | |
| - Ahora no tienes novio/a formal | 2 | |
| - Hasta ahora sólo has tenido relaciones afectivas pasajeras | 3 | AMOR |
| - Nunca has tenido una relación afectiva especial con un chico/a | 4 | |

PRACTICAS SPSS.**ANALISIS DE DATOS I.***CULTURA*

P.9 Vamos a hablar ahora de tus estudios. ¿En qué centro realizaste la totalidad o la mayor parte de tus estudios?

- En un centro estatal, público 1
- En un centro privado no religioso 2 **COLEGIO**
- En un centro privado religioso 3

P.10 Excluyendo los libros de texto, ¿cuántos libros has leído en los últimos doce meses?

_____ **LIBROS**

ACTITUDES

P.11 ¿Cómo te defines en materia religiosa?

- Católico practicante 1
- Católico no practicante 2
- Creyente de otra religión 3 **RELIGION**
- No creyente 4
- Indiferente 5

P.12 Cuando se habla de política se utilizan normalmente las expresiones izquierda y derecha. ¿En qué casilla te colocarías? **ESCALA**

	+-----+ -----	
Izda.		Dcha.
	+-----+ -----	
	01 02 03 04 05 06 07 08 09 10	

INGRESOS

P.13 ¿Qué cantidad aproximadamente de dinero (en miles de pesetas) ingresas al mes por cada uno de los conceptos siguientes?

Ingresos personales _____ **INGRESO1**
 Ingresos de tu pareja _____ **INGRESO2**
 Aportaciones familiares _____ **INGRESO3**

P.14 Actualmente, entre todos los miembros de tu hogar y por todos los conceptos, ¿de cuántos ingresos netos (sin descuentos) disponéis por término medio en tu hogar al mes?

- Menos o igual a 50.000 pts. 01
- De 50.001 a 100.000 pts. 02
- De 100.001 a 150.000 pts. 03
- De 150.001 a 200.000 pts. 04
- De 200.001 a 300.000 pts. 05 **INGRESOT**
- De 300.001 a 400.000 pts. 06
- De 400.001 a 500.000 pts. 07
- De 500.001 a 750.000 pts. 08
- De 750.001 a 1 Millón de pts. 09
- Más de 1 Millón de pts. 10

DATOS

P.15 Sexo:

- Hombre 1 **GENERO**
- Mujer 2

P.16 ¿Cuántos años cumpliste en tu último cumpleaños?

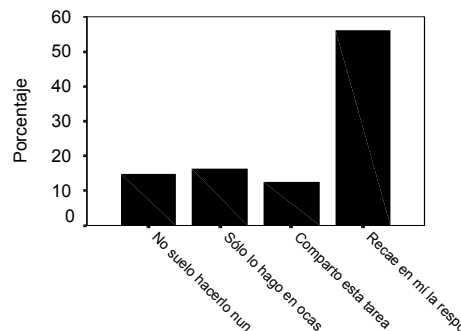
_____ años **EDAD**

PRÁCTICA 1. ANÁLISIS ESTADÍSTICO DE UNA VARIABLE.

ADVERTENCIA: Esta práctica consta de **preguntas obligatorias (en negrita)**, de *preguntas opcionales (en cursiva)* y de preguntas "algo difíciles" (en letra normal).

Comience por las preguntas en negrita y continúe con las que están escritas en cursiva. Por último, intente responder a las de letra normal.

1. **¿Cuántas variables tiene el archivo de datos?** _____
2. **¿Qué variable ocupa el lugar nº 10?** _____
3. **¿Cuántos sujetos forman parte de la muestra?** _____
4. **¿Qué puntuación obtuvo el sujeto nº 24 en la variable ESCALA?** _____
5. **Indique el nivel de medida de las siguientes variables (para responder consulte la descripción de la pregunta correspondiente en el cuestionario, pues en la vista de variables del SPSS no se muestra correctamente):**
 - OCIO** _____
 - HORAS** _____
 - INGRESOT** _____
6. *Escriba el nombre de una variable, distinta a las tres anteriores medida a nivel...*
 - Nominal* _____
 - Ordinal* _____
 - Cuantitativa discreta* _____
 - Cuantitativa continua* _____
7. **Elabore la distribución de frecuencias de la variable INGRESOT**
 - 7.1 **¿Cuántas familias tienen ingresos de cómo máximo 200.000 pts?** _____
 - 7.2 **¿Qué % de familias tienen ingresos entre 200.001 y 400.000 pts?** _____
 - 7.3 **¿Cuántas familias tienen ingresos superiores a 500.000 pts?** _____
8. **Elabore la distribución de frecuencias de la variable EDAD sólo para las mujeres de la muestra**
 - 8.1 **¿Cuántas mujeres tienen cómo máximo 26 años?** _____
 - 8.2 **¿Qué % de mujeres tienen entre 21 y 27 años?** _____
9. *Elabore un gráfico de barras para cada una de las variables de la pregunta 5 del cuestionario.*
 - 9.1 *¿A qué variable corresponde el siguiente gráfico?*



- 9.2 *Elabore la distribución de frecuencias de dicha variable*

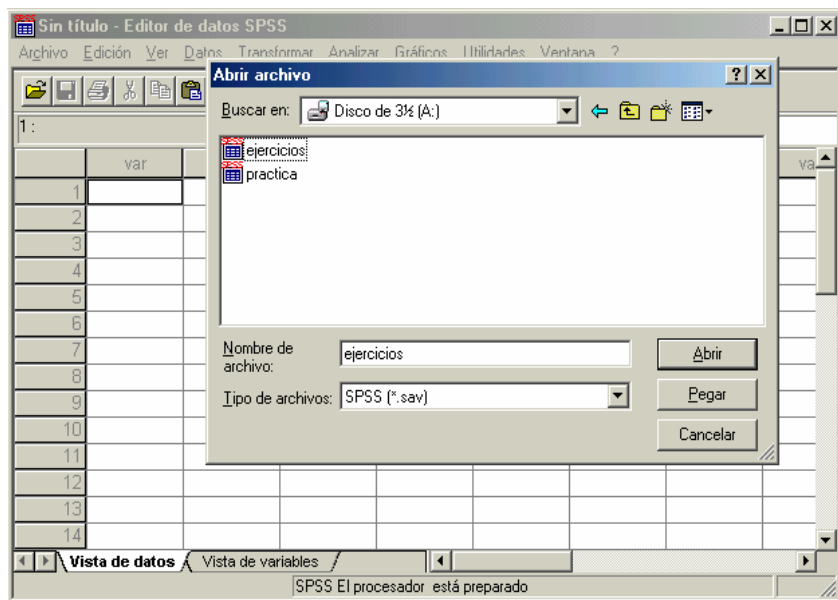
10. ¿Qué % de entrevistados tienen o han tenido algún tipo de relación afectiva? (pregunta 8 del cuestionario)
11. Conteste a las siguientes preguntas, todas ellas referidas a la variable **TV** (pregunta 3 del cuestionario)
- 11.1 ¿Cuál es el índice de asimetría de dicha variable? _____
- 11.2 ¿Qué porcentaje de individuos ve menos de 18 horas la televisión? _____
- 11.3 El 82% de la muestra ve como máximo ____ horas de televisión. Dicho valor es el centil ____
- 11.4 Por término medio, ¿cuántas horas ven semanalmente la televisión? _____
- 11.5 Divida la distribución en 4 grupos iguales de sujetos e indique qué valores de la variable TV separan estos 4 grupos _____, _____, _____.
- 11.6 ¿Cuáles son los valores de las tres medidas de tendencia central? _____, _____, _____
- 11.7 **Convierta la variable TV (“horas/semana dedicadas a ver la TV”) en la variable DIAS (“días/semana dedicados a ver la TV”).**
La media de DIAS es _____.
- 11.8 ¿Quién ve más la televisión, los varones o las mujeres? _____
¿En qué estadístico se ha basado para responder? _____
- 11.9 Indique cuál sería la media y la varianza de la nueva variable X que resulta de sumar a las puntuaciones en la variable TV el valor 5.
- 11.10 ¿Qué valores, en el grupo de varones, comprende el 5º tallo del diagrama de tallo y hojas?
- 11.11 ¿Qué grupo (varones o mujeres) es más heterogéneo? _____
¿En qué indicador de variabilidad se ha basado para responder? _____
- 11.12 Represente gráficamente la variabilidad en TV para varones y mujeres.
12. ¿Qué porcentaje de entrevistados nacieron antes del año 1978? _____
13. Con las variables de la pregunta 13 del cuestionario (INGRESO1, INGRESO2, INGRESO3) elabore la nueva variable **TOTALES = INGRESO1 + INGRESO2 + INGRESO3**.
- 13.1 La media de la nueva variable **TOTALES** es: _____ y la varianza es: _____.
- 13.2 Compruebe, mediante las propiedades de la media, si la media en **TOTALES** resulta igual si se calcula a partir de las medias en las variables **INGRESO1**, **INGRESO2** e **INGRESO3**.
- 13.3 Convierta la variable **TOTALES** (que está en pesetas) en la variable **EUROS** (1 euro = 166,386 pts) y diga cuál es la media de **EUROS**: _____

ESQUEMA DE MENÚS DEL SPSS PARA LA PRÁCTICA 1

ENTRAR EN SPSS: Inicio > Programas > SPSS 13.0 para Windows

I. ABRIR ARCHIVO

Archivo > Abrir > Datos



Al abrir el archivo se obtiene el siguiente resultado en el editor de datos:

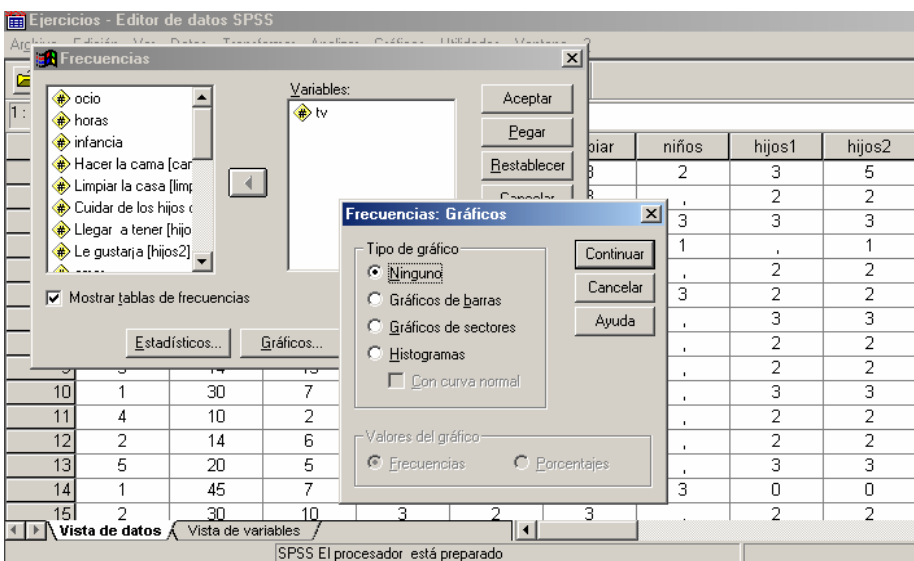
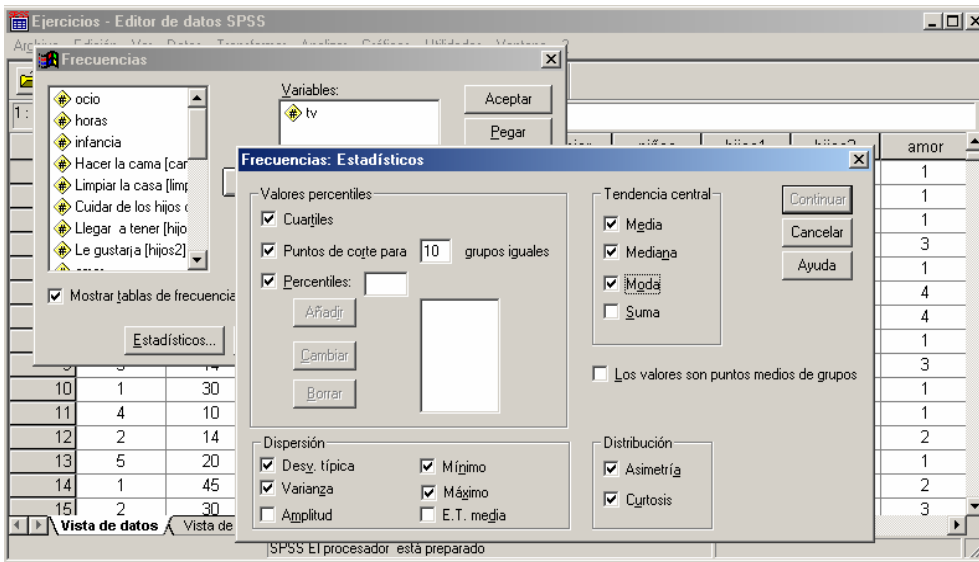
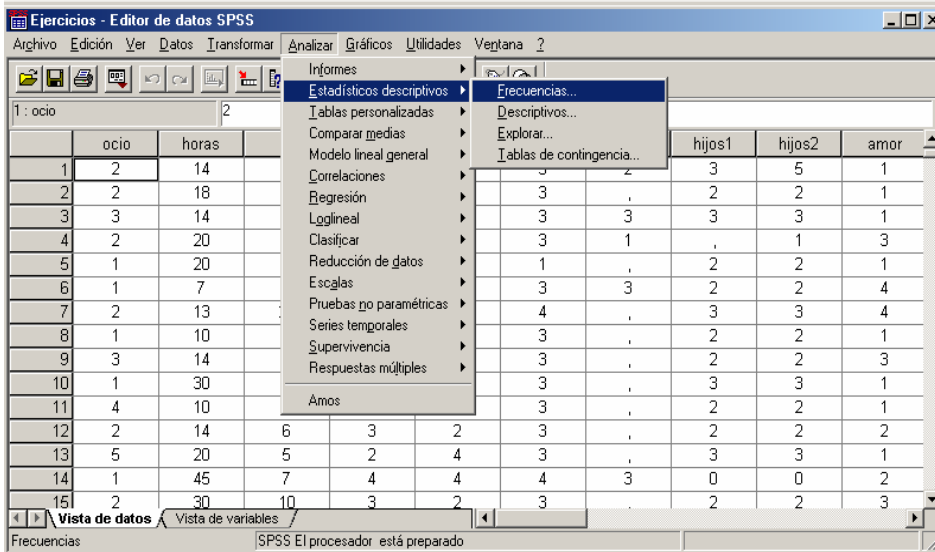
	ocio	horas	tv	infancia	cama	limpiar	niños	hijos1	hijos2	amor	colegio	libros	religion	escala	ingreso1
1	2	14	14	3	4	3	2	3	5	1	1	3	1	5	0
2	2	18	6	1	4	3	.	2	2	1	1	15	2	7	40
3	3	14	4	3	4	3	3	3	3	1	1	4	1	6	0
4	2	20	6	3	4	3	1	.	1	3	3	4	1	5	0
5	1	20	10	2	1	1	.	2	2	1	1	.	5	8	0
6	1	7	3	3	1	3	3	2	2	4	1	12	1	5	0
7	2	13	20	3	4	4	.	3	3	4	1	.	5	98	0
8	1	10	6	3	4	3	.	2	2	1	1	12	4	3	0
9	3	14	15	2	4	3	.	2	2	3	1	4	2	5	50
10	1	30	7	3	3	3	.	3	3	1	1	8	2	6	0
11	4	10	2	3	1	3	.	2	2	1	1	12	5	4	0
12	2	14	6	3	2	3	.	2	2	2	1	2	4	3	0
13	5	20	5	2	4	3	.	3	3	1	1	25	1	4	0
14	1	45	7	4	4	4	3	0	0	2	1	15	3	98	36
15	2	30	10	3	2	3	.	2	2	3	1	12	2	8	0

Cada variable del fichero corresponde a cada pregunta del cuestionario. Por ejemplo, la variable 'ocio' corresponde a la pregunta 1, que tiene 6 opciones de respuesta, la variable 'horas' a la pregunta 2, que es una pregunta abierta en la que se indica el número de horas que se dedican al ocio, etcétera.

Es conveniente, tener el cuestionario delante para comprobar a qué variable se refiere cada pregunta, pues el cuestionario indica en mayúsculas la variable de SPSS que corresponde a cada pregunta.

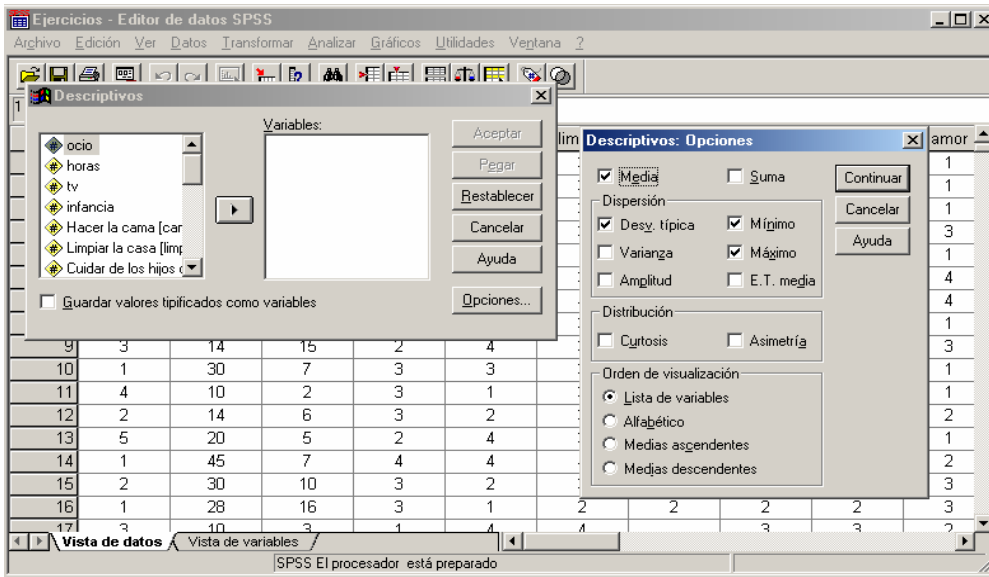
II. FRECUENCIAS

Analizar > Estadísticos Descriptivos > Frecuencias



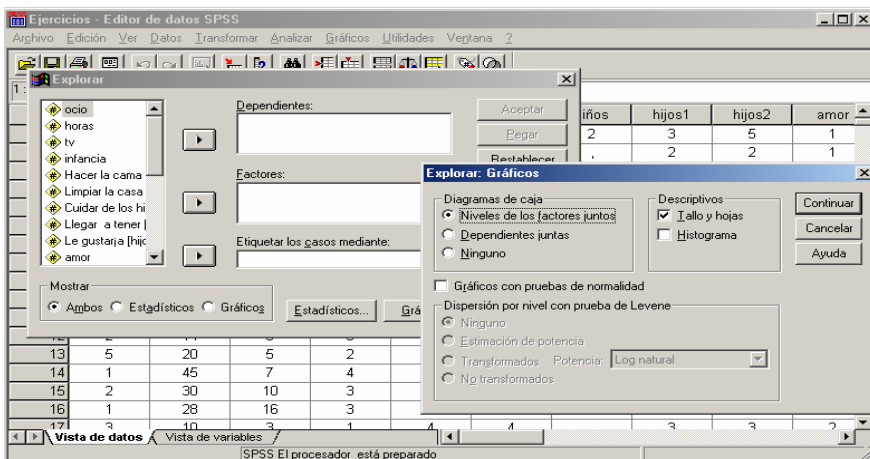
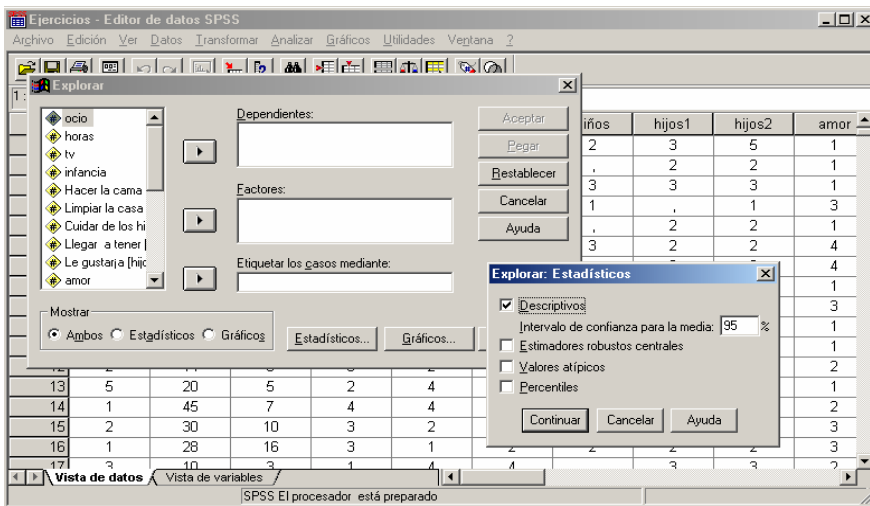
III. DESCRIPTIVOS

Analizar > Estadísticos Descriptivos > Descriptivos



IV. EXPLORAR

Analizar > Estadísticos Descriptivos > Explorar



PRÁCTICA 2. ANÁLISIS ESTADÍSTICO DE DOS VARIABLES.**1. REPASO DE ESTADÍSTICA UNIVARIADA**

Las siguientes preguntas se refieren a la variable **ESCALA** del archivo *Ejercicios.sav*

- 1.1. Calcule los estadísticos de tendencia central de dicha variable _____, _____, _____.
- 1.3. Calcule los índices de asimetría y curtosis de dicha variable _____, _____.
- 1.3. Queremos hacer 4 grupos iguales en dicha variable. Los grupos son: izquierda, centro-izquierda, centro-derecha y derecha. ¿Qué valores de la variable **ESCALA** delimitan cada uno de los grupos?
_____, _____, _____.
- 1.4. Calcule los centiles 20, 40 y 80 de dicha distribución: $C_{20} = \underline{\hspace{2cm}}$ $C_{40} = \underline{\hspace{2cm}}$ $C_{80} = \underline{\hspace{2cm}}$

2. MATRICES DE VARIANZA-COVARIANZA Y CORRELACIONES

Seleccione las variables de la pregunta 13 del cuestionario (**INGRESO1**, **INGRESO2** e **INGRESO3**) del archivo *ejercicios.sav* y ejecute la matriz de varianzas-covarianzas y la matriz de correlaciones entre las tres variables (menú *Analizar -> Correlaciones -> bivariadas*).

Según los resultados obtenidos en el visor del SPSS conteste a las siguientes preguntas:

- 2.1. ¿Cuál de las tres variables es más homogénea? _____
- 2.2. Elabore la matriz de varianzas-covarianzas entre las variables **INGRESO1** e **INGRESO2**

	INGRESO 1	INGRESO 2
INGRESO 1		
INGRESO 2		

- 2.3. Elabore la matriz de correlaciones entre las variables **INGRESO2** e **INGRESO3**

	INGRESO 2	INGRESO 3
INGRESO 2		
INGRESO 3		

- 2.4. Compruebe cómo se obtendría la correlación de Pearson entre **INGRESO1** e **INGRESO2** si solo se utilizan los datos que aparecen en la matriz de varianzas-covarianzas obtenida en el apartado 2.2.

3. COMBINACIÓN LINEAL DE VARIABLES

Calcule (mediante el menú *Transformar -> Calcular* del SPSS) una nueva variable que sea la suma de los ingresos personales del entrevistado, y los ingresos de su pareja (**INGRESO1** + **INGRESO2**). A la nueva variable la llamaremos **INPAREJA**. La etiqueta de dicha variable será “ingresos personales y de la pareja”.

- 3.1. La media de la variable **INPAREJA** es: _____ la varianza: _____ y el índice de asimetría: _____
- 3.2. Calcule a partir de la matriz de varianzas-covarianzas entre las variables **INGRESO1** e **INGRESO2** (obtenida en el apartado 2.2.), el valor de la varianza de la variable **INPAREJA**.

4. MODELOS LINEALES DE REGRESIÓN

Queremos elaborar un modelo de regresión lineal que nos permita pronosticar las horas que un entrevistado ve la televisión a la semana (TV) a partir de las horas semanales de ocio de dicho entrevistado (HORAS).

En dicho modelo ...

- 4.1. La variable predictora es: _____ y de la variable criterio es: _____
- 4.2. La ecuación de regresión en puntuaciones directas es: _____, en puntuaciones diferenciales es: _____ y en puntuaciones típicas es: _____
- 4.3. ¿Cuál es la bondad de dicho modelo de regresión? _____
¿En qué estadístico se basa y cuál es su valor? _____
- 4.4. ¿Cuántas horas semanales de televisión pronosticaría a 3 sujetos si cada uno de ellos ha obtenido los siguientes valores en la variable predictora? ...
Sujeto 1: Valor en la predictora: 30 ¿Valor en el pronóstico? _____
Sujeto 2: Valor en la predictora: 10 ¿Valor en el pronóstico? _____
Sujeto 3: Valor en la predictora: 6 ¿Valor en el pronóstico? _____
- 4.5. Descomponga la varianza del criterio. Según lo obtenido, la varianza explicada es: _____ y la no explicada es: _____.

5. ELABORACIÓN DE TABLAS DE CONTINGENCIA

Se pretende estudiar la relación existente entre las actividades de ocio (**OCIO**) y la situación afectiva del entrevistado (**AMOR**).

	OCIO				
AMOR					

- 5.1. Elabore la distribución marginal de frecuencias de la variable **OCIO**
- 5.2. Elabore la distribución marginal de porcentajes de la variable **AMOR**
- 5.3. Elabore la distribución condicional, en porcentajes, de la variable **OCIO** respecto a los que tienen novio formal en la variable **AMOR**
- 5.4. Elabore la distribución condicional, en porcentajes, de la variable situación afectiva (**AMOR**) respecto a los que en su tiempo de ocio se dedican a viajar (**OCIO**)

6. INTERPRETACIÓN DE TABLAS DE CONTINGENCIA

Elabore una tabla de contingencia con las variables **RELIGIÓN** (columna) y **COLEGIO** (fila).

	RELIGIÓN				
COLEGIO					

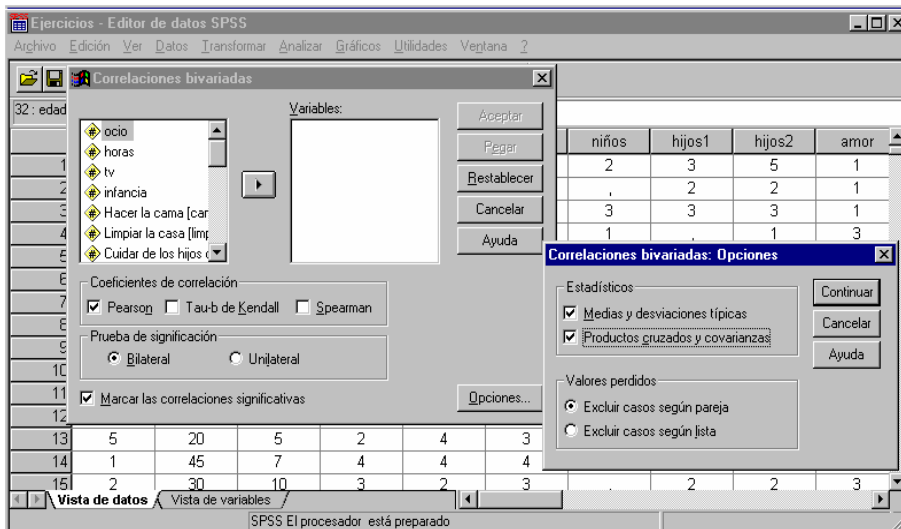
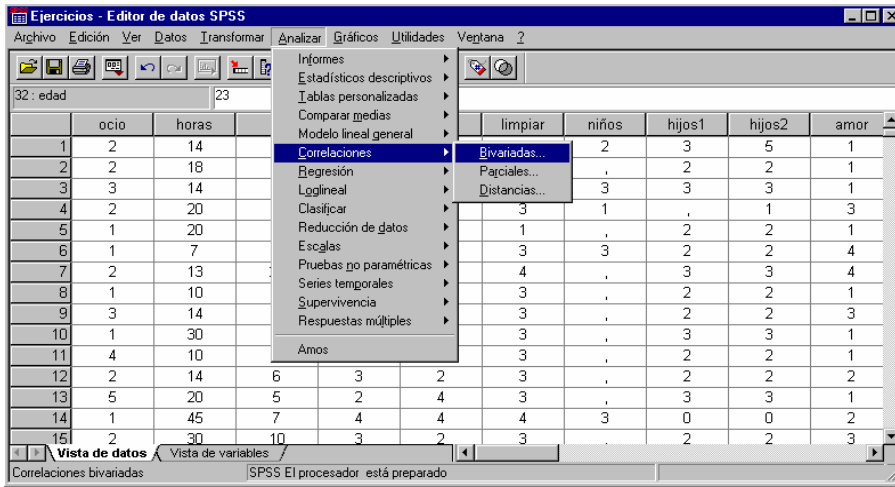
Elabore las tablas de distribuciones condicionales (% columna y % fila) entre ambas variables y conteste a las siguientes preguntas:

- 6.1. % de entrevistados que se declaran no-creyentes _____
- 6.2. % de entrevistados que han cursado estudios en un colegio público _____
- 6.3. % de no-creyentes que han cursado estudios en un colegio privado religioso _____
- 6.4. % de alumnos de un colegio público que se declaran creyentes de otra religión _____

ESQUEMA DE MENÚS DEL SPSS PARA LA PRÁCTICA 2

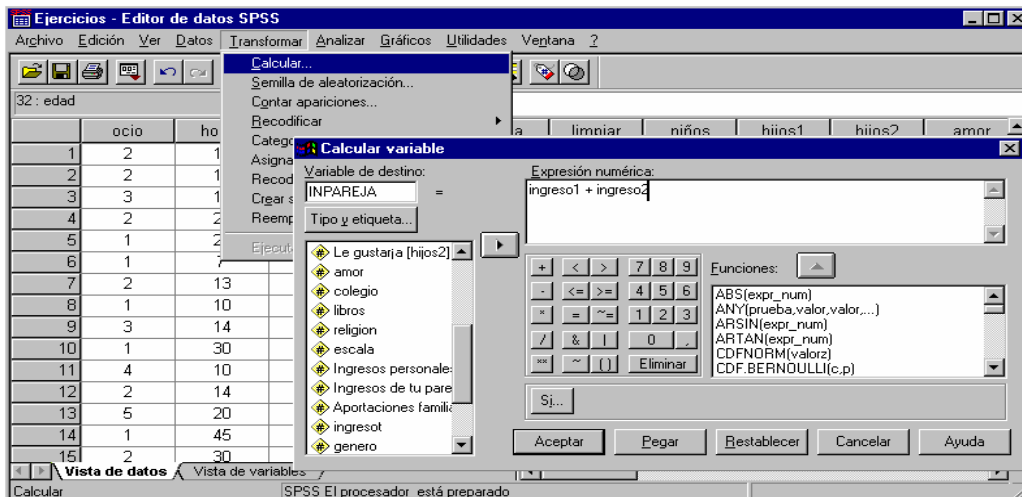
I. CORRELACIONES

Archivo > Correlaciones > Bivariadas



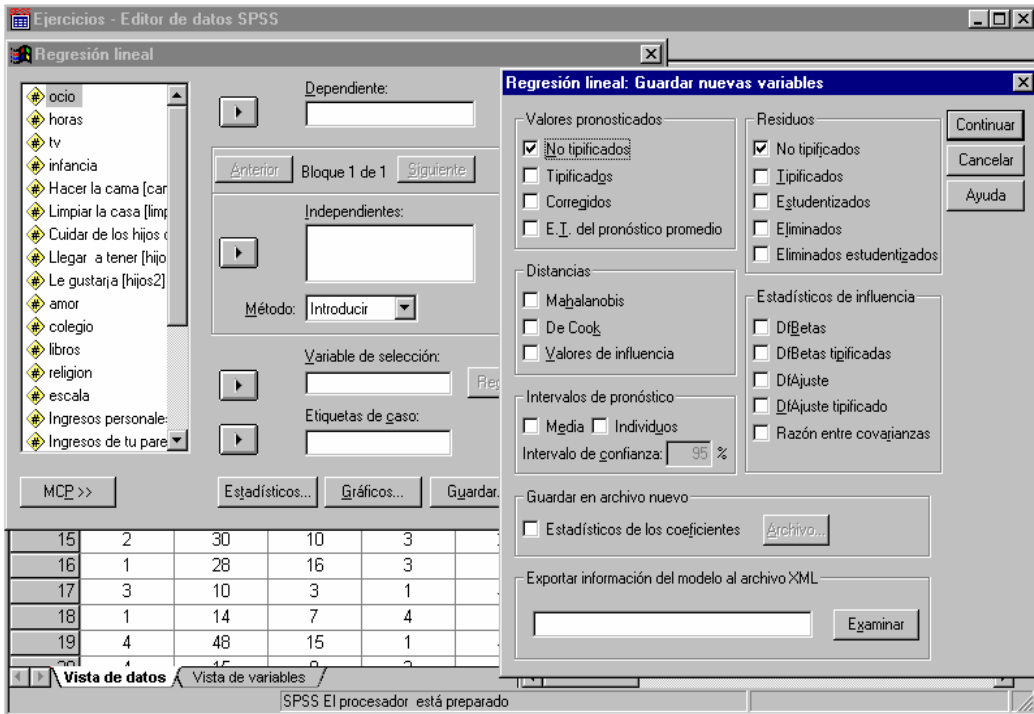
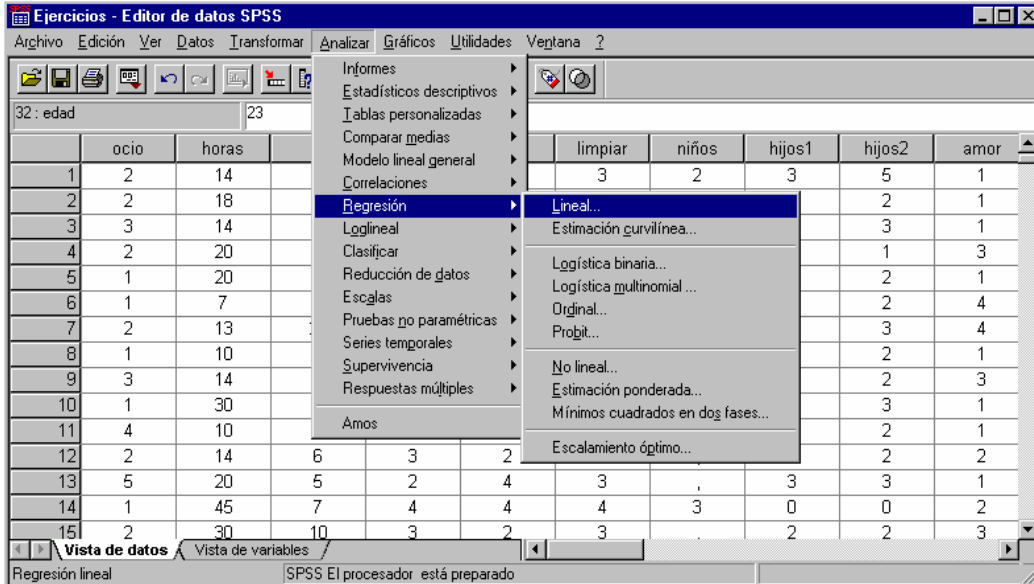
III. CALCULAR COMBINACIONES LINEALES

Transformar > Calcular



III. REGRESIÓN

Analizar > Regresión > Lineal



IV. TABLAS DE CONTINGENCIA

Analizar > Estadísticos Descriptivos > Tablas de contingencia

