

Práctica 3.

Cálculo de estadísticos de tendencia central y dispersión a partir de ficheros de encuesta.

En esta práctica se obtendrán estadísticos de resumen de variables de intervalo procedentes de encuestas de opinión. La práctica tiene dos fases, en la primera se explicará cómo descargar un fichero con microdatos de una encuesta del CIS y en una segunda parte se calcularán e interpretarán los principales estadísticos de tendencia central y de dispersión. La realización de la práctica resulta central para conocer los instrumentos básicos para el tratamiento de grandes cantidades de información. En concreto analizaremos algunas variables del Barómetro de Opinión de Enero de 2012 (Estudio CIS 2927).

Primera Parte: Elaboración del fichero activo

En la práctica anterior se explicó como grabar manualmente los datos de una variable. Ahora nos situaremos en el caso en que los datos ya han sido grabados en soporte informático para su uso. Por una parte tenemos los datos de los cuestionarios grabados en un fichero, en nuestro caso será el fichero DA2927, y por otra parte tenemos el programa para realizar los cálculos: el PSPP-. El fichero de datos cargado para su uso por el programa se conoce con el nombre de fichero activo –tiene extensión SAV-, y lo visualizaremos en el formato de hoja de cálculo. Para realizar dicha operación necesitamos un fichero de sintaxis que contiene las instrucciones para la lectura de los datos. El fichero de sintaxis, contiene las llamadas a los datos y la especificación de las posiciones de las variables y la ejecución de dichas instrucciones permite la generación de un fichero de trabajo, sobre el que haremos las distintas demandas de información. De forma esquemática podemos representar el funcionamiento de los programas de análisis de encuestas:

Fichero de datos DA2927.TXT

292702166104625060	0	0	01	0422422	1	822	1	8	
292702167104625060	0	0	01	0522523	81622	812			
292702168104625060	0	0	01	0431311	111	6	813		
292702169104625060	0	0	01	0533322	111	8	8	3	
292702170104625060	0	0	01	0433423	11611	111			
292702171104625060	0	0	01	0421422	8	1	4	8	3
292702172104625060	0	0	01	0333522	1	8	099	0	
292702173104625060	0	0	01	0422421	1	8	01216		
292702174104625060	0	0	01	0538428	121	621	3		
292702175104625060	0	0	01	0523523	1	8	0	1	8
2927021761046	030	0	01	0438321	1	0	020	0	
2927021771046	030	0	01	0431322	1	822	116		
2927021781046	030	0	01	0421422	1	0	099	0	
2927021791046	030	0	01	0521311	196	0	1	0	
2927021801046	030	0	01	053252213	8	6	1	0	
2927021811046	030	0	01	0431311	1	822	7	8	
2927021821046	030	0	02	7332311	1	8	0	1	8
292702183104625060	0	0	01	0433431	11322	1	0		
292702184104625060	0	0	02	36331422	111	0	122		
292702185104625060	0	0	01	0531888	111	0	813		
2927021861046	030	0	01	0531311	1	8	0	1	0

Fichero de sintaxis ES2927.SPS

```

*ES2927 --- PSPPIRE Editor de sintaxis
Archivo  Editar  Ejecutar  Ventanas  Ayuda
1  DATA LIST FILE= DA2927
2  / CUES 5-9 A1 1-9(A) CCAA PROV 10-13 MUN 14-16 TAMUNI 17 AREA 18
3  DISTR 19-20 SECCION 21-23 ENTREV 24-27
4  P0 28 P0A 29-31 P1 TO P6 32-37 P701 TO P703 38-43
5  P7A 38-39 P801 TO P803 44-49 P8A 44-45
6  P901 TO P910 P10 TO P13 P1401 TO P1409 P1501 TO P1508 50-80
7  P16 P16A P1701 TO P1714 81-112 p18 P1901 TO P1906 P20
8  P2101 TO P2103 P21A01 TO P21A03 113-126 P2201 TO P2213 127-152
9  P23 P24 153-154 P25 TO P27 155-160 P28 161 P28A 162-163
10 P29 164 P30 165-166 P31 167 P31A 168-169 P32 P32A P33 TO P37
11 P37A P37B 170-178 P38 179-181 P39 P39A 182-183 P40 184-185
12 P41 P41A P41B P42 P4301 TO P4304 186-193 P4401 TO P4405
13 P4501 TO P4505 194-223(A) P46 TO P49 224-227
14 FINAL 228-275(A) P26A 276-277 VOTOSIM 278-279 RECUERDO 280-281
15 ESTUDIOS 282 OCUPA 283-284 CONDICION 285-286 ESTATUS 287.
16
17 MISSING VALUES P0 TO P16a P18 to p21a03 p23 to p4304 P46 TO P49 P26A
18 VOTOSIM RECUERDO ESTUDIOS OCUPA CONDICION ESTATUS(0).
19
20 VARIABLE LABEL
21 CCAA 'Comunidad autónoma'
22 /PROV 'Provincia'
23 /TAMUNI 'Tamaño de hábitat'
24 /AREA 'Área metropolitana'
25 /MUN 'Municipio'
26 /p901 'El paro'
27 /p902 'La inseguridad ciudadana'
28 /p903 'El terrorismo'
29 /p904 'Las drogas'
30 /p905 'El medio ambiente'
31 /p906 'La asistencia sanitaria'
32 /p907 'La vivienda'

```

Fichero Activo ENCUESTA.SAV

	CUES	A1	CCAA	PROV	MUN	TAMUNI	AREA	DISTR	SECCION	ENTREV	P0	P0A	P1	P2	P3	P4
1	2162	292702162	10	46	250	6	0	0	0	0	1	0	4	2	3	4
2	2163	292702163	10	46	250	6	0	0	0	0	1	0	4	3	3	5
3	2164	292702164	10	46	250	6	0	0	0	0	1	0	4	2	1	3
4	2165	292702165	10	46	250	6	0	0	0	0	1	0	5	3	3	4
5	2166	292702166	10	46	250	6	0	0	0	0	1	0	4	2	2	4
6	2167	292702167	10	46	250	6	0	0	0	0	1	0	5	2	2	5
7	2168	292702168	10	46	250	6	0	0	0	0	1	0	4	3	1	3
8	2169	292702169	10	46	250	6	0	0	0	0	1	0	5	3	3	3
9	2170	292702170	10	46	250	6	0	0	0	0	1	0	4	3	3	4
10	2171	292702171	10	46	250	6	0	0	0	0	1	0	4	2	1	4
11	2172	292702172	10	46	250	6	0	0	0	0	1	0	3	3	3	5
12	2173	292702173	10	46	250	6	0	0	0	0	1	0	4	2	2	4
13	2174	292702174	10	46	250	6	0	0	0	0	1	0	5	3	8	4
14	2175	292702175	10	46	250	6	0	0	0	0	1	0	5	2	3	5
15	2176	292702176	10	46	0	3	0	0	0	0	1	0	4	3	8	3
16	2177	292702177	10	46	0	3	0	0	0	0	1	0	4	3	1	3
17	2178	292702178	10	46	0	3	0	0	0	0	1	0	4	2	1	4
18	2179	292702179	10	46	0	3	0	0	0	0	1	0	5	2	1	3
19	2180	292702180	10	46	0	3	0	0	0	0	1	0	5	3	2	5
20	2181	292702181	10	46	0	3	0	0	0	0	1	0	4	3	1	3
21	2182	292702182	10	46	0	3	0	0	0	0	2	7	3	3	2	3
22	2183	292702183	10	46	250	6	0	0	0	0	1	0	4	3	3	4

a) Obtención y preparación de los ficheros

Comenzaremos obteniendo los datos desde el Banco de Datos del CIS. Desde la dirección www.cis.es podemos llegar a localizar los datos desde diversos lugares (Estudios, Barómetros...). En este caso recomendamos el acceso por **Banco de Datos**.



Dentro de **Banco de datos** en la zona de **Búsquedas de estudios** seleccionaremos **Índice anual**, y en los correspondientes al año 2012 buscaremos el **Barómetro Enero 2012**.

04/01/12	2927	BARÓMETRO ENERO 2012	Resultados_PDF Resultados_HTML Cruces Ficha_técnica Cuestionario Fichero_datos
--------------------------	----------------------	--------------------------------------	---

Una vez localizado el estudio en el Banco de Datos, seleccionamos el icono **Fichero_datos**

Para descargar el fichero el CIS nos solicitará nuestra conformidad de las condiciones de utilización. Debemos indicar nuestros datos, marcar la casilla de aceptación y pulsar en descargar. (De momento la descarga es sólo posible para

CIS
Centro de Investigaciones Sociológicas

Descarga del archivo: MD2927.zip

Datos del solicitante

Nombre:

Apellidos:

Universidad/Centro/Empresa:

E-mail:

Objeto de la solicitud

Investigación/informe/libro
 Trabajo de curso
 Docencia

Tesis/tesina/máster
 Medios de comunicación
 Otros

Proyectos/planes/campañas
 Conferencias/ponencias

Condiciones de utilización:

El grado de **fiabilidad y exactitud** de la información obtenida a partir de los datos por elaboración propia del usuario es **responsabilidad exclusiva** del mismo. Los datos que se van a descargar se encuentran sometidos a los derechos de Copyright del CIS que, sin perjuicio de las excepciones en la legislación vigente, **no podrán ser reproducidos o transmitidos**, total o parcialmente, en ninguna forma ni por ningún medio electrónico, mecánico o de otro tipo, **no podrá hacerse uso comercial** de los datos suministrados, **ni cedernos a terceros** por cualesquiera procedimientos de difusión, sin la **autorización expresa y escrita del CIS**. Además el usuario tiene obligación de **indicar el origen de los datos** cuando haga uso público de ellos en cualquier forma.

De conformidad con lo establecido en la [Ley Orgánica 15/1999](#), de 13 de diciembre, de **Protección de Datos de Carácter Personal**, los datos de este formulario pasarán a formar parte de un fichero informático, responsabilidad del CIS, mantenido con fines exclusivamente estadísticos, sobre el que pueden ejercitarse los derechos de acceso, rectificación, cancelación y oposición.

El/la solicitante conoce y acepta los términos anteriores

Por favor, lea las [INSTRUCCIONES](#) sobre el contenido y utilización del fichero comprimido antes de realizar la descarga.

Descargar

CIS - Montalbán, 8 - 28014 Madrid - España | e-mail: cis@cis.es

El fichero que descargamos es un fichero comprimido MD2927.ZIP. La mejor opción es que guardemos esta carpeta en un dispositivo USB, de forma que conozcamos claramente el “*path*” o lugar en el que está localizable dicha carpeta.

El “*path*” literalmente inglés es el camino nos indica la ruta de acceso física a los ficheros. Los discos duro de los ordenadores suelen denominarse con la letra C seguida de dos puntos. Así “C:” significa disco duro C. Los dispositivos USB utilizan letras por lo general desde la E hasta la K.

Después de la unidad se indica la carpeta o fichero –siempre utilizando un *slash* o barra inversa “\” como separador, por ejemplo si grabamos nuestro fichero en una memoria flash denominada K, el *path* será: **K:\MD2927.ZIP**

Al abrir la carpeta nos encontraremos con los siguientes ficheros.

La carpeta MD2927 contiene los siguientes ficheros:

codigo2927.pdf	Es el libro de códigos del estudio. Muestra la posición de las variables (columnas) en el fichero de datos. También nos muestra los valores permitidos para cada una de las variables.
cues2927.pdf	Es la reproducción del cuestionario utilizado
DA2927	Son los datos de la encuesta grabados (en formato ASCII)
ES2927	Es el fichero de sintaxis para SPSS o PSPP. Básicamente es el libro de códigos convertido en comandos para la lectura del fichero de datos.
Ft2927.pdf	Es la Ficha Técnica del Estudio
Sas2927	Es el fichero de sintaxis para el programa SAS (Para nuestra práctica este fichero no tiene utilidad ninguna)
tarjetas2927.pdf	Es el fichero que contiene una reproducción de las tarjetas que utilizaron los encuestadores durante la realización de la entrevista.

Antes de continuar conviene que se familiarice con el cuestionario utilizado. Seleccione alguna pregunta y busque su correspondencia en el libro de códigos.

Para continuar vamos a trasladar el fichero de datos DA2927 y el fichero de sintaxis ES2927 a una nueva carpeta, alojada en el directorio raíz [K:]¹. La nueva carpeta la denominaremos CIS – por ejemplo- y en ella alojaremos los dos ficheros. Aprovecharemos también para añadir la extensión al nombre de los ficheros.

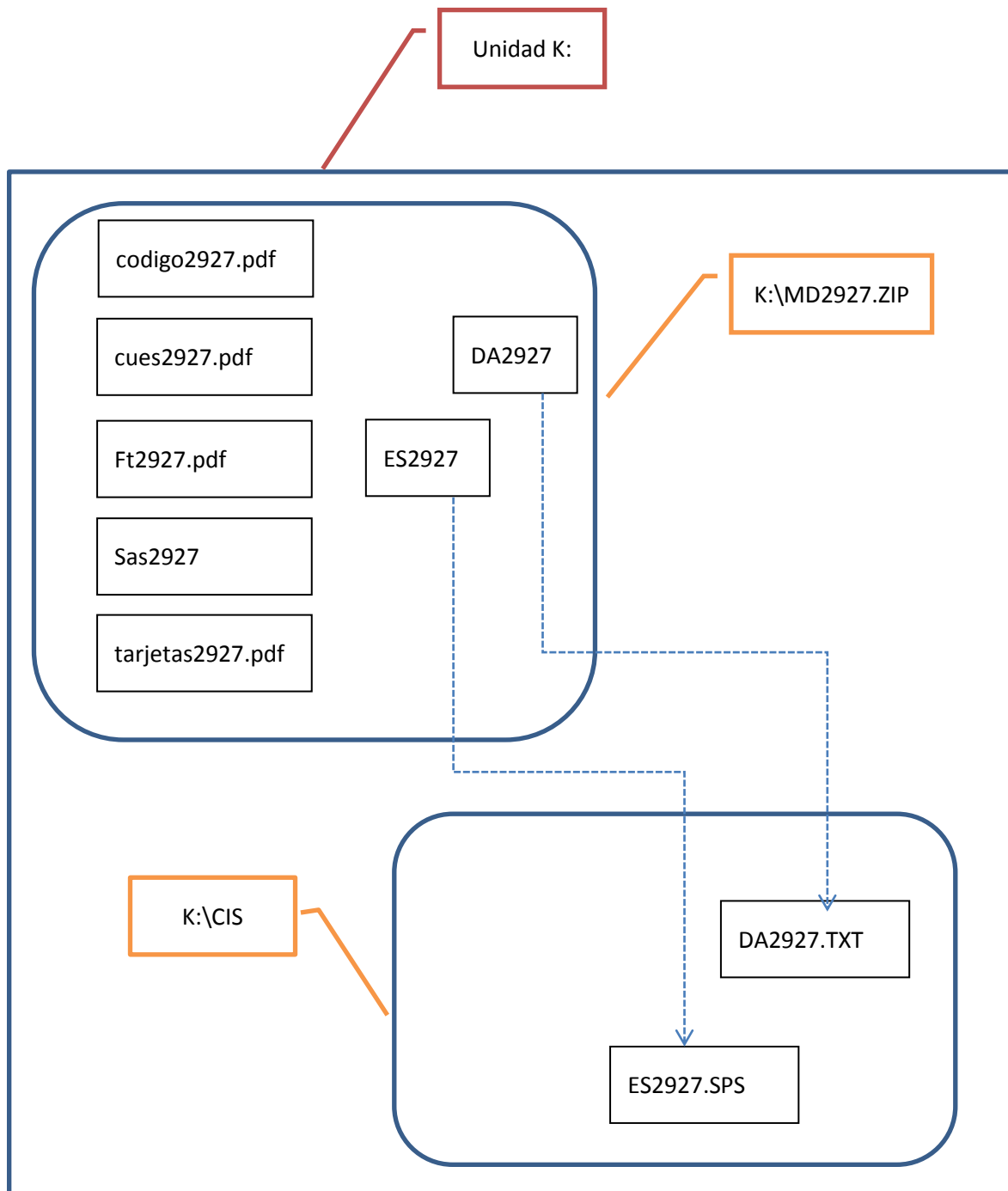
(En cada sistema operativo o versión de Windows varía el procedimiento por ello indicaremos los pasos de forma genérica):

- a) Crear en el directorio raíz de (por ejemplo en [K:]) una carpeta denominada CIS.
[K:\CIS]
- b) Trasladar el fichero DA2927 desde la carpeta [K:\DA2927.ZIP] a [K:\CIS]
- c) Trasladar el fichero ES2927 desde la carpeta [K:\DA2927.ZIP] a [K:\CIS]
- d) Renombrar el fichero [K:\CIS\DA2927] con la extensión TXT de forma que quede como DA2927.TXT situado como [K:\CIS\DA2927.TXT]

¹ El símbolo de los corchetes [] se utiliza para diferenciar dentro del texto las expresiones de denominación de los *paths* donde están alojados los ficheros.

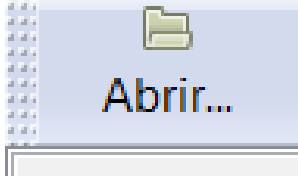
- e) Renombrar el fichero [K:\CIS\ES2927] con la extensión SPS de forma que quede como [K:\CIS\ES2927.SPS]

El siguiente esquema ayuda a comprender las operaciones realizadas

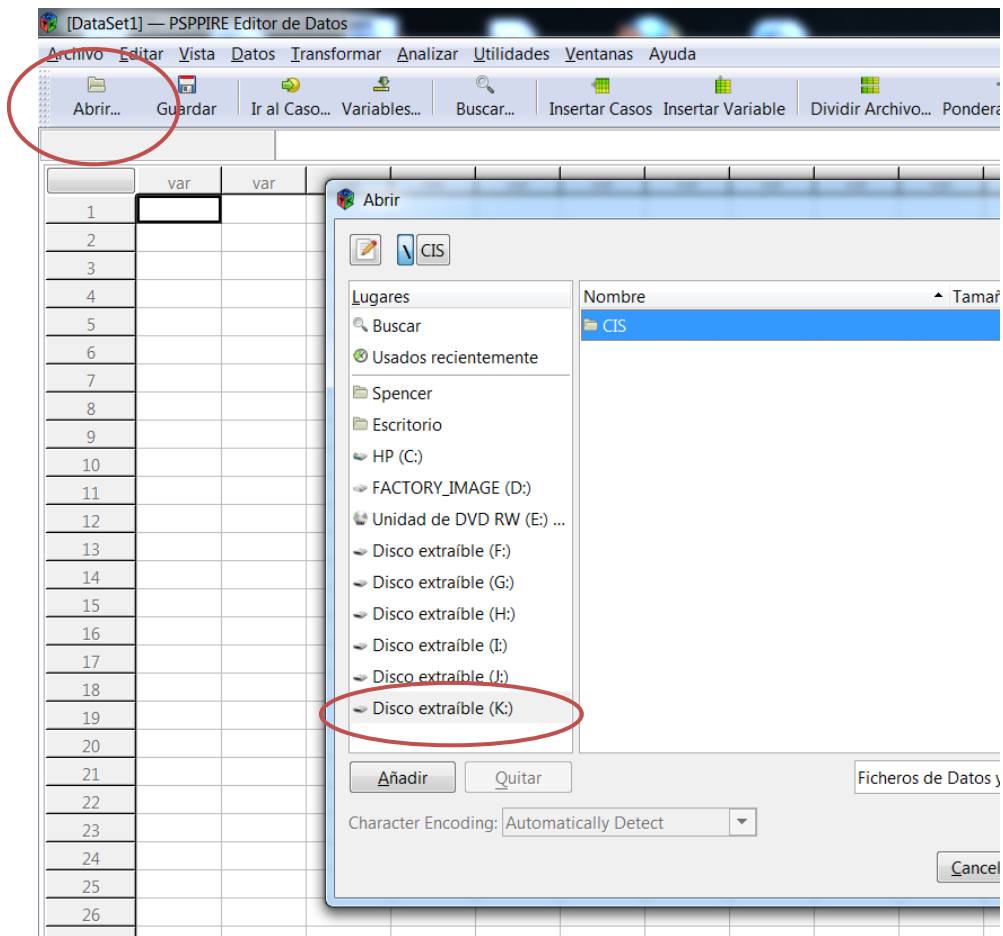


b) Generación del fichero activo

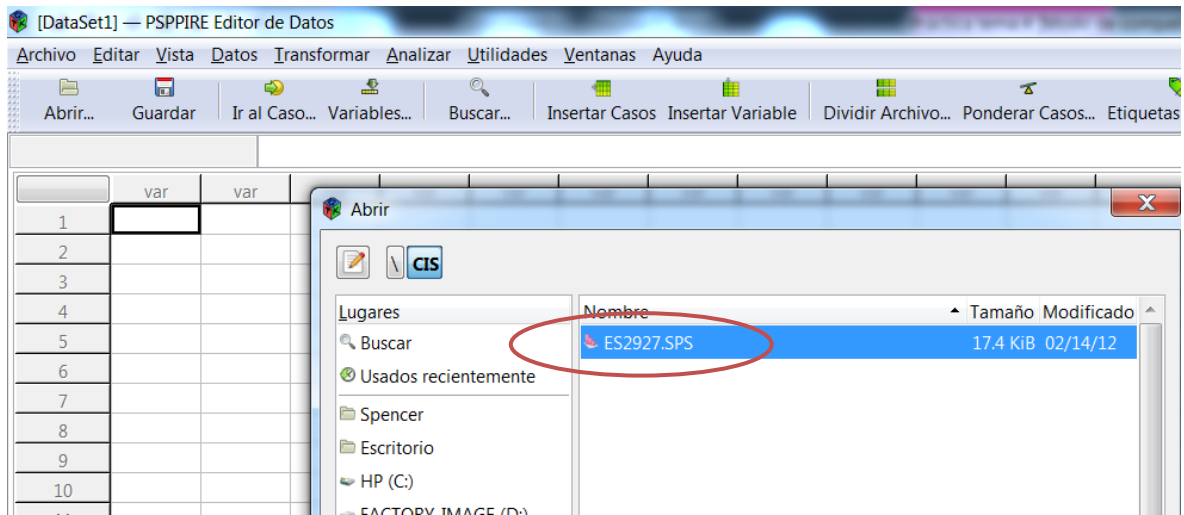
Una vez descargados y descomprimidos los ficheros necesarios, abriremos el programa PPSS. En primer lugar, llamaremos mediante el botón **Abrir...**



al fichero que contiene las instrucciones de sintaxis ES2927.SPS que está en:
[K:\CIS\ES2927.SPS]



Navegaremos por los menús, seleccionando la unidad [K:], la carpeta [\CIS] y el fichero: ES2927.SPS



Después de seleccionar el fichero se abrirá una nueva ventana con las instrucciones que contiene el fichero de sintaxis.

```

*ES2927.SPS — PSPPIRE Editor de sintaxis
Archivo Editar Ejecutar Ventanas Ayuda
1 DATA LIST FILE='DA2927'
2 / CUES 5-9 A1 1-9(A) CCAA PROV 10-13 MUN 14-16 TAMUNI 17 AREA 18
3 DISTR 19-20 SECCION 21-23 ENTREV 24-27
4 P0 28 P0A 29-31 P1 TO P6 32-37 P701 TO P703 38-43
5 P7A 38-39 P801 TO P803 44-49 P8A 44-45
6 P901 TO P910 P10 TO P13 P1401 TO P1409 P1501 TO P1508 50-80
7 P16 P16A P1701 TO P1714 81-112 p18 P1901 TO P1906 P20
8 P2101 TO P2103 P21A01 TO P21A03 113-126 P2201 TO P2213 127-152
9 P23 P24 153-154 P25 TO P27 155-160 P28 161 P28A 162-163
10 P29 164 P30 165-166 P31 167 P31A 168-169 P32 P32A P33 TO P37
11 P37A P37B 170-178 P38 179-181 P39 P39A 182-183 P40 184-185
12 P41 P41A P41B P42 P4301 TO P4304 186-193 P4401 TO P4405
13 P4501 TO P4505 194-223(A) P46 TO P49 224-227
14 FINAL 228-275(A) P26A 276-277 VOTOSIM 278-279 RECUERDO 280-281
15 ESTUDIOS 282 OCUPA 283-284 CONDICION 285-286 ESTATUS 287.
16
17 MISSING VALUES P0 TO P16a P18 to p21a03 p23 to p4304 P46 TO P49 P26A
18 VOTOSIM RECUERDO ESTUDIOS OCUPA CONDICION ESTATUS(0).
19
20 VARIABLE LABEL
21 CCAA'Comunidad autónoma'
22 /PROV'Provincia'
23 /TAMUNI'Tamaño de hábitat'
24 /AREA'Área metropolitana'
25 /MUN'Municipio'
26 /p901 'El paro'
27 /p902 'La inseguridad ciudadana'
28 /p903 'El terrorismo'
29 /p904 'Las drogas'
30 /p905 'El medio ambiente'
31 /p906 'La asistencia sanitaria'
32 /p907 'La vivienda'
33 /p908 'La inmigración'
34 /p909 'La educación'
35 /p910 'Las pensiones'
36 /p1401 'El trabajo'
37 /p1402 'La familia'
38 /p1403 'La política'
39 /p1404 'El bienestar económico'
40 /p1405 'El tiempo libre/ocio'
41 /p1406 'La salud'

```

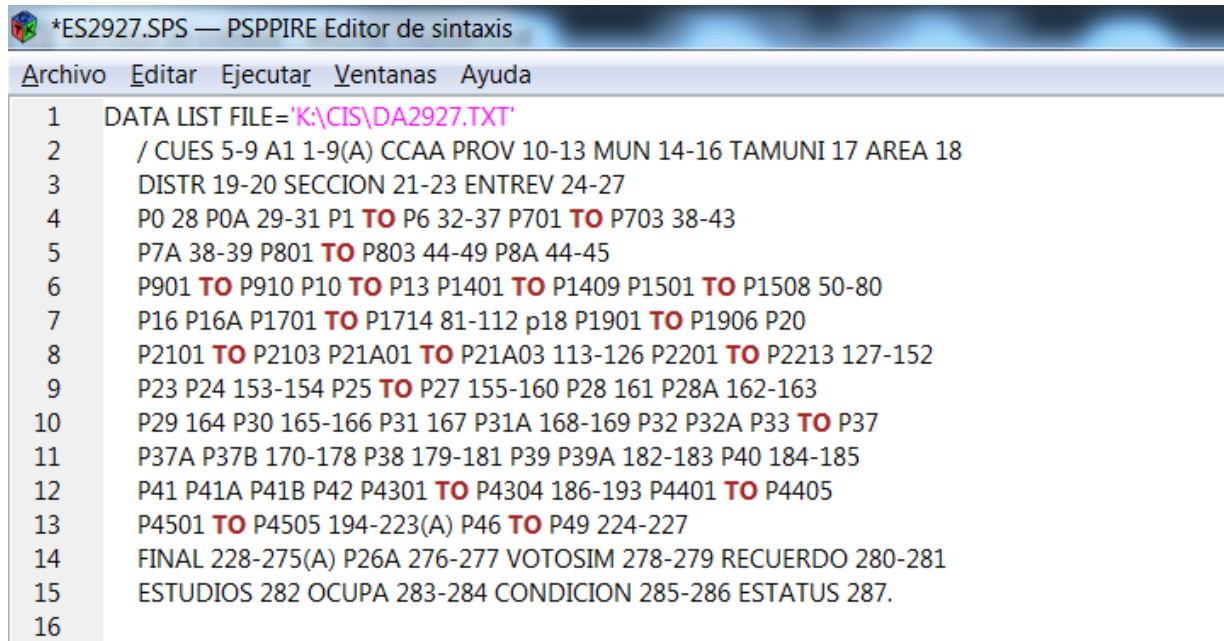
Si nos desplazamos por el fichero vamos a encontrar cinco instrucciones distintas. Las instrucciones comienzan con las siguientes expresiones y terminan con “.” (un punto):

Línea 1	DATA LIST
Línea 17	MISSING VALUES
Línea 20	VARIABLE LABEL
Línea 104	VALUE LABEL
Línea 735	FREQ

Instrucción **DATA LIST**

La primera instrucción DATA LIST sirve para indicar al programa la posición “path” del fichero que contiene los datos, y la posición de cada una de los datos en el fichero.

Lo primero que vamos a hacer es indicar el *path* correcto, que recordemos que es [K:\CIS\DA2927.TXT]



```
*ES2927.SPS — PSPPIRE Editor de sintaxis
Archivo  Editar  Ejecutar  Ventanas  Ayuda
1  DATA LIST FILE='K:\CIS\DA2927.TXT'
2  / CUES 5-9 A1 1-9(A) CCAA PROV 10-13 MUN 14-16 TAMUNI 17 AREA 18
3  DISTR 19-20 SECCION 21-23 ENTREV 24-27
4  P0 28 P0A 29-31 P1 TO P6 32-37 P701 TO P703 38-43
5  P7A 38-39 P801 TO P803 44-49 P8A 44-45
6  P901 TO P910 P10 TO P13 P1401 TO P1409 P1501 TO P1508 50-80
7  P16 P16A P1701 TO P1714 81-112 p18 P1901 TO P1906 P20
8  P2101 TO P2103 P21A01 TO P21A03 113-126 P2201 TO P2213 127-152
9  P23 P24 153-154 P25 TO P27 155-160 P28 161 P28A 162-163
10 P29 164 P30 165-166 P31 167 P31A 168-169 P32 P32A P33 TO P37
11 P37A P37B 170-178 P38 179-181 P39 P39A 182-183 P40 184-185
12 P41 P41A P41B P42 P4301 TO P4304 186-193 P4401 TO P4405
13 P4501 TO P4505 194-223(A) P46 TO P49 224-227
14 FINAL 228-275(A) P26A 276-277 VOTOSIM 278-279 RECUERDO 280-281
15 ESTUDIOS 282 OCUPA 283-284 CONDICION 285-286 ESTATUS 287.
16
```

La expresión **DATA LIST FILE='K:\CIS\DA2927.TXT'** nos indica el lugar donde se alojan nuestros datos. Obsérvese que el *path* va indicado entre comillas. (Se pueden utilizar las dobles comillas "" o también como es habitual en inglés, las simples.)

La barra "/" es el comienzo de la segunda parte del comando que está destinado a nombrar las variables e indicar su posición. Si se consulta el fichero **cues2927.pdf** que contiene el cuestionario podemos interpretar fácilmente las siguientes líneas. Recuérdese que en el cuestionario se indican las columnas de las variables entre paréntesis.

La expresión **CUES 5-9** significa que el número de cuestionario, lo vamos a llamar CUES y dicho dato ocupan las columnas 5 a 9. (Tiene una longitud de 5 dígitos)

Comunidad Autónoma _____	<input type="text"/>	(10)(11)	Nº ESTUDIO	Nº CUESTIONARIO
Provincia _____	<input type="text"/>	(12)(13)	<input type="text" value="2 9 2 7"/>	<input type="text"/>
Municipio _____ (nombre municipio)	<input type="text"/>	(14)(15)(16)	(1)(2)(3)(4)	(5)(6)(7)(8)(9)
Tamaño de hábitat _____	<input type="text"/>	(17)(18)		
Distrito _____	<input type="text"/>	(19)(20)		
Sección _____	<input type="text"/>	(21)(22)(23)		
Entrevistador/a _____	<input type="text"/>	(24)(25)(26)(27)		

La expresión **A1 1-9(A)** define una variable que llamamos A1 y ocupa las nueve primeras columnas. La (A) entre paréntesis señala que una variable que no es numérica. Esta variable no la utilizaremos, es de uso interno del CIS. Es la composición del número de Estudio (columnas 1-4 que en este caso es siempre 2927 para toda la muestra más el número de cuestionario).

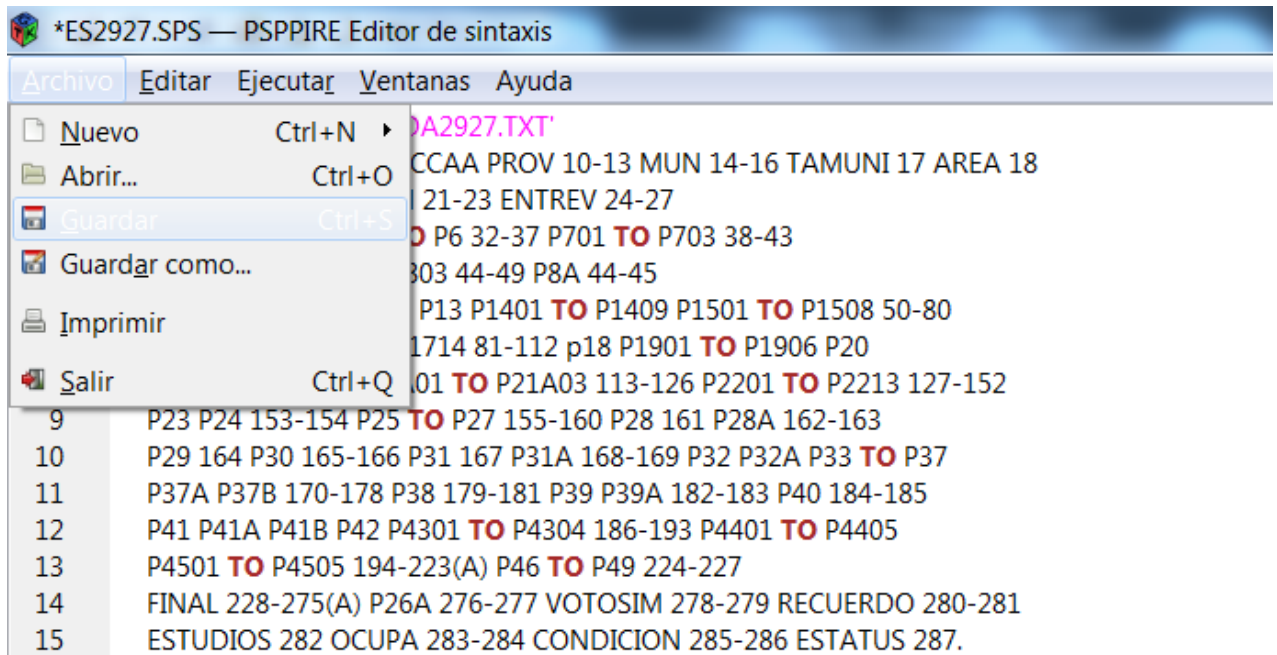
La expresión **CCAA PROV 10-13** define dos variables a la vez CCAA indica la Comunidad Autónoma y Prov la provincia. Las cuatro columnas se reparten de forma equitativa las dos primeras 10-11 para CCAA y las dos últimas para PROV. Podría haberse escrito también de forma equivalente **CCAA 10-11 PROV 12-13**.

En la línea 7 encontramos la expresión **P1 TO P6 32-37** con el uso de TO podemos de forma abreviada de nombrar varias variables a la vez, en este caso a 6 variables: P1, P2, P3, P4, P5, P6. Consecuentemente P1 ocupa la columna 32, P2 la 33... y P7 la 37.

<p>P.0 En primer lugar querría preguntarle si tiene Ud. la nacionalidad...</p> <ul style="list-style-type: none"> - Española 1 (28) - Española y otra 2 - Otra nacionalidad 3 → FIN DE LA ENTREVISTA <p>P.0a ¿Cuál?</p> <p>_____ (29)(30)(31)</p> <p>N.C. 999</p>	<p>P.4 Y refiriéndonos ahora a la <u>situación política</u> general de España, ¿cómo la calificaría Ud.: muy buena, buena, regular, mala o muy mala?</p> <ul style="list-style-type: none"> - Muy buena 1 - Buena 2 - Regular 3 - Mala 4 (35) - Muy mala 5 - N.S. 8 - N.C. 9
<p>P.1 Para empezar, refiriéndonos a la <u>situación económica</u> general de España, ¿cómo la calificaría Ud.: muy buena, buena, regular, mala o muy mala?</p> <ul style="list-style-type: none"> - Muy buena 1 - Buena 2 - Regular 3 - Mala 4 (32) - Muy mala 5 - N.S. 8 - N.C. 9 	<p>P.5 ¿Y cree Ud. que la situación política actual del país es mejor, igual o peor que <u>hace un año</u>?</p> <ul style="list-style-type: none"> - Mejor 1 - Igual 2 - Peor 3 (36) - N.S. 8 - N.C. 9
<p>P.2 Y, ¿cree Ud. que la situación económica actual del país es mejor, igual o peor que <u>hace un año</u>?</p> <ul style="list-style-type: none"> - Mejor 1 - Igual 2 - Peor 3 (33) - N.S. 8 - N.C. 9 	<p>P.6 Y, ¿cree Ud. que <u>dentro de un año</u> la situación política del país será mejor, igual o peor que ahora?</p> <ul style="list-style-type: none"> - Mejor 1 - Igual 2 - Peor 3 (37) - N.S. 8 - N.C. 9
<p>P.3 Y, ¿cree Ud. que <u>dentro de un año</u> la situación económica del país será mejor, igual o peor que ahora?</p> <ul style="list-style-type: none"> - Mejor 1 - Igual 2 - Peor 3 (34) - N.S. 8 - N.C. 9 	<p>P.7 ¿Cuál es, a su juicio, el principal problema que existe actualmente en España? ¿Y el segundo? ¿Y el tercero?</p> <p>_____ (38)(39)</p> <p>_____ (40)(41)</p> <p>_____ (42)(43)</p> <p>N.S. 98</p> <p>N.C. 99</p>

El comando DATA LIST termina en la línea 15 con un punto “.”.

Antes de seguir explicando los siguientes comandos vamos a ejecutar el DATA LIST. Asegúrese de haber escrito correctamente el *path*. A continuación grabe el fichero. Para ello localice lo botones **Archivo, Guardar**



Una vez guardado el fichero, para ejecutar un comando marque el bloque (líneas desde la 1 a la 15) y seleccione el botón **Ejecutar** y la opción **Selección**



La ejecución del comando abrirá una tercera ventana de resultados, en la que se van registrando las ordenes de la sesión y los resultados de la misma. En este caso ofrece una tabla con el nombre de las variables y las columnas que ocupan así como el formato. Por ejemplo F5.0 señala que es una variable numérica con cinco dígitos y sin decimales.

DATA LIST

DATA LIST

DATA LIST FILE='K:\CIS\DA2927.TXT'

/ CUES 5-9 A1 1-9(A) CCAA PROV 10-13 MUN 14-16 TAMUNI 17 AREA 18

DISTR 19-20 SECCION 21-23 ENTREV 24-27

P0 28 P0A 29-31 P1 TO P6 32-37 P701 TO P703 38-43

P7A 38-39 P801 TO P803 44-49 P8A 44-45

P901 TO P910 P10 TO P13 P1401 TO P1409 P1501 TO P1508 50-80

P16 P16A P1701 TO P1714 81-112 P18 P1901 TO P1906 P20

P2101 TO P2103 P21A01 TO P21A03 113-126 P2201 TO P2213 127-152

P23 P24 153-154 P25 TO P27 155-160 P28 161 P28A 162-163

P29 164 P30 165-166 P31 167 P31A 168-169 P32 P32A P33 TO P37

P37A P37B 170-178 P38 179-181 P39 P39A 182-183 P40 184-185

P41 P41A P41B P42 P4301 TO P4304 186-193 P4401 TO P4405

P4501 TO P4505 194-223(A) P46 TO P49 224-227

FINAL 228-275(A) P26A 276-277 VOTOSIM 278-279 RECUERDO 280-281

ESTUDIOS 282 OCUPA 283-284 CONDICION 285-286 ESTATUS 287.

Leyendo 1 registro de 'K:\CIS\DA2927.TXT'.

Variable	Registro	Columnas	Formato
CUES	1	5- 9	F5.0
A1	1	1- 9	A9
CCAA	1	10- 11	F2.0
PROV	1	12- 13	F2.0
MUN	1	14- 16	F3.0
TAMUNI	1	17- 17	F1.0
AREA	1	18- 18	F1.0
DISTR	1	19- 20	F2.0
SECCION	1	21- 23	F3.0
ENTREV	1	24- 27	F4.0
P0	1	28- 28	F1.0
P0A	1	29- 31	F3.0
P1	1	32- 32	F1.0
P2	1	33- 33	F1.0
P3	1	34- 34	F1.0
P4	1	35- 35	F1.0
P5	1	36- 36	F1.0
P6	1	37- 37	F1.0

Si vamos a la pantalla inicial ahora encontramos la rejilla de datos rellena. Podemos ver que el primer cuestionario del fichero es el 2162, pertenece a la Comunidad Autónoma 10, provincia 46. Si consultamos el libro de códigos, fichero **codigo2927.pdf**, podemos comprobar que es una entrevista realizada en Valencia. A la pregunta número 1 responde con 4, es decir que considera la situación económica de España como Mala. El entrevistado siguiente 2163 la ha considerado como Muy Mala.

	CUES	A1	CCAA	PROV	MUN	TAMUNI	AREA	DISTR	SECCION	ENTREV	P0	P0A	P1
1	2162	292702162	10	46	250	6	0	0	0	0	1	0	4
2	2163	292702163	10	46	250	6	0	0	0	0	1	0	4
3	2164	292702164	10	46	250	6	0	0	0	0	1	0	4
4	2165	292702165	10	46	250	6	0	0	0	0	1	0	5
5	2166	292702166	10	46	250	6	0	0	0	0	1	0	4
6	2167	292702167	10	46	250	6	0	0	0	0	1	0	5
7	2168	292702168	10	46	250	6	0	0	0	0	1	0	4
8	2169	292702169	10	46	250	6	0	0	0	0	1	0	5
9	2170	292702170	10	46	250	6	0	0	0	0	1	0	4
10	2171	292702171	10	46	250	6	0	0	0	0	1	0	4
11	2172	292702172	10	46	250	6	0	0	0	0	1	0	3
12	2173	292702173	10	46	250	6	0	0	0	0	1	0	4
13	2174	292702174	10	46	250	6	0	0	0	0	1	0	5
14	2175	292702175	10	46	250	6	0	0	0	0	1	0	5

Resulta evidente, que la lectura de los datos mediante los códigos resulta engorrosa. Para ello tenemos dos comandos que nos facilitan dicha labor. Si volvemos al fichero de sintaxis, podemos observar otros comandos (Variable Label y Value Label):

```
VARIABLE LABEL
  CCAA 'Comunidad autónoma'
  /PROV 'Provincia'
  /TAMUNI 'Tamaño de hábitat'
  /AREA 'Área metropolitana'
  /MUN 'Municipio'
  /p901 'El paro'
  /p902 'La inseguridad ciudadana'
  /p903 'El terrorismo'
  /p904 'Las drogas'
  /p905 'El medio ambiente'
  /p906 'La asistencia sanitaria'
  /p907 'La vivienda'
  /p908 'La inmigración'
  /p909 'La educación'

...

/RECUERDO 'Recuerdo de voto'
/ESTUDIOS 'Estudios del entrevistado'
/OCUPA 'Ocupación del entrevistado'
/CONDICION 'Condición socioeconómica'
/ESTATUS 'Estatus socioeconómico'.
```

VARIABLE LABEL, que nos rotula a las variables. Por ejemplo, en los resultados en vez de que aparezca CCAA aparezca “Comunidad Autónoma”. La sintaxis de dicho comando es simple. Nombre de la variable y entre comillas el rótulo, las variables se separan con la barra “/”. (Nótese como es habitual que el comando termina con un punto).

VALUE LABEL

CCAA

01'Andalucía' 02'Aragón' 03'Asturias (Principado de)'
04'Baleares (Islas)' 05'Canarias' 06'Castabria'
07'Castilla La Mancha' 08'Castilla y León'
09'Cataluña' 10'Comunidad Valenciana' 11'Extremadura'
12'Galicia' 13'Madrid (Comunidad de)' 14'Murcia (Región de)'
15'Navarra (Comunidad Foral de)' 16'País Vasco' 17'Rioja (La)'

/PROV

01'Álava' 02'Albacete' 03'Alicante' 04'Almería' 33'Asturias'
05'Ávila' 06'Badajoz' 07'Baleares' 08'Barcelona' 09'Burgos'
10'Cáceres' 11'Cádiz' 39'Castabria' 12'Castellón de la Plana'
13'Ciudad Real' 14'Córdoba' 15'Coruña (A)' 16'Cuenca'
17'Girona' 18'Granada' 19'Guadalajara' 20'Guipúzcoa'
21'Huelva' 22'Huesca' 23'Jaén' 24'León' 25'Lleida' 27'Lugo' 28'Madrid'
29'Málaga' 30'Murcia' 31'Navarra' 32'Ourense' 34'Palencia'
35'Palmas (Las)' 36'Pontevedra' 26'Rioja (La)' 37'Salamanca'
38'Santa Cruz de Tenerife' 40'Segovia' 41'Sevilla' 42'Soria'
43'Tarragona' 44'Teruel' 45'Toledo' 46'Valencia' 47'Valladolid'
48'Vizcaya' 49'Zamora' 50'Zaragoza'

•••

/CONDICION

1 'Directores y profesionales'
2 'Técnicos y cuadro medios'
3 'Pequeños empresarios'
4 'Agricultores'
5 'Empleados de oficinas y servicios'
6 'Obreros cualificados'
7 'Obreros no cualificados'
8 'Jubilados y pensionistas'
9 'Parados'
10 'Estudiantes'
11 'Trabajo doméstico no remunerado'
12 'No clasificables'

/ESTATUS

1 'Clase alta/ media-alta'
2 'Nuevas clases medias'
3 'Viejas clases medias'
4 'Obreros cualificados'
5 'Obreros no cualificados'
9 'No consta'.

VALUE LABEL, que nos rotula los valores de las variables. La sintaxis es también sencilla, después de cada nombre de variable se van colocando los distintos valores con sus rótulos entre comillas.

Después de la ejecución de ambos comandos –marcaremos con el ratón las líneas de cada uno de los comandos y en **Ejecutar** pulsaremos la opción **Selección**, nos encontramos con la siguiente rejilla de datos, que en vez de códigos muestra las etiquetas. En la barra superior tenemos el botón de **Etiquetas de Valor** que nos permitirá mostrar bien los códigos, bien las etiquetas de valor.

	CUES	A1	CCAA	PROV	MUN	TAMUNI	AREA	DISTR	SECCION	ENTREV	P0	P0A	P1	P2
1	2162	292702162	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Igual
2	2163	292702163	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Peor
3	2164	292702164	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Igual
4	2165	292702165	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Muy mala	Peor
5	2166	292702166	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Igual
6	2167	292702167	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Muy mala	Igual
7	2168	292702168	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Peor
8	2169	292702169	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Muy mala	Peor
9	2170	292702170	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Peor
10	2171	292702171	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	Igual
11	2172	292702172	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Regular	Peor

Recuérdese que también pueden cambiarse las etiquetas de valor y de variable de forma manual, accediendo por la pestaña de **Vista de Variables**. Sin embargo para grandes conjuntos de datos, resulta más cómodo el uso de ficheros de sintaxis con las etiquetas.

20	2181	292702181	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	
21	2182	292702182	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Regular	
22	2183	292702183	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Mala	
23	2184	292702184	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Regular	
24	2185	292702185	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Muy mala	
25	2186	292702186	Iad Valenciana	Valencia	250300 habitantes	0	0	0	0	0	0	0	Muy mala	

Una vez que hemos ejecutado los comandos DATA LIST, VARIABLE LABEL y VALUE LABEL podemos generar nuestro fichero activo, es decir el fichero sobre el que pediremos los resultados. Para ello pulsamos sobre **Archivo** y seleccionamos la opción **Guardar como**.

Untitled — PSPPIRE Editor de Datos

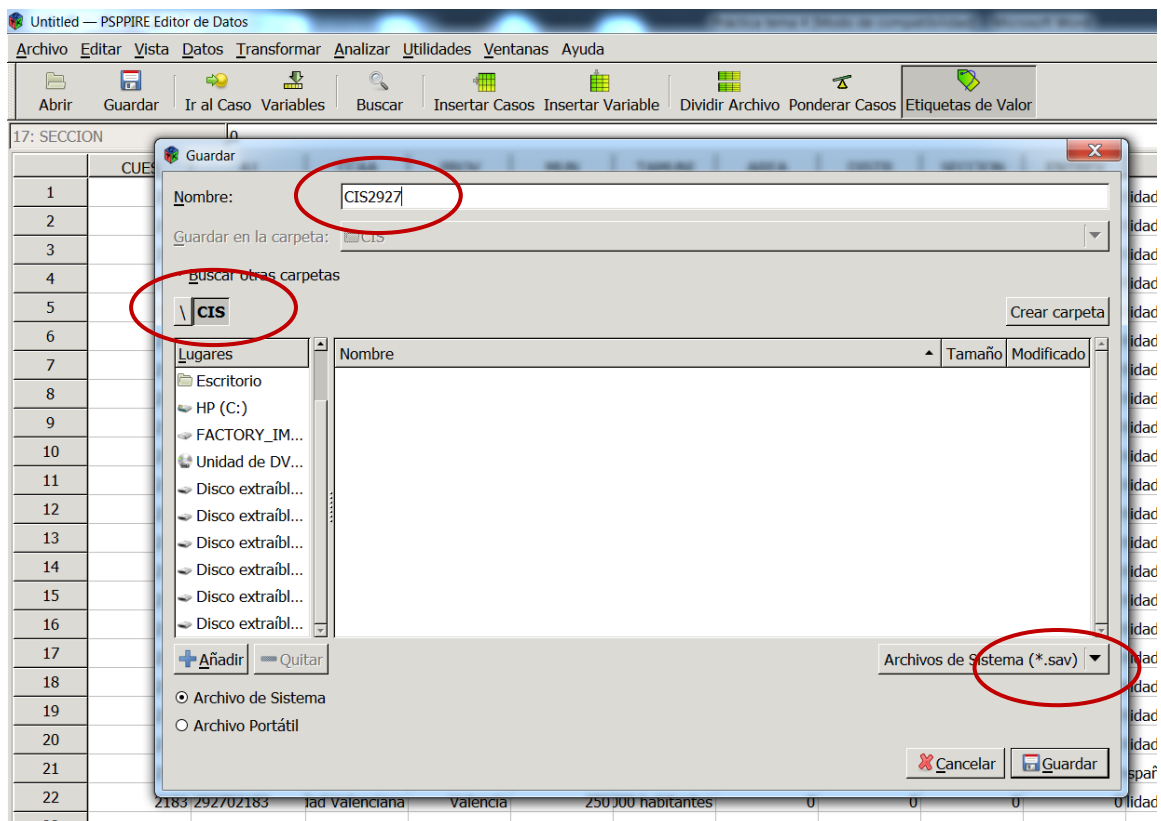
Archivo Editar Vista Datos Transformar Analizar Utilidades Ventanas Ayuda

Nuevo Ctrl+N
 Abrir Ctrl+O
 Importar datos de texto delimitados
 Guardar Ctrl+S
 Guardar como
 Muestra información del archivo de datos
 Datos utilizados recientemente
 Archivos utilizados recientemente
 Salir Ctrl+Q

Insertar Casos Insertar Variable Dividir Archivo Ponderar Casos Etiquetas de Valor

	PROV	MUN	TAMUNI	AREA	DISTR	SECCION	ENTR
5	Valencia	250300 habitantes	0	0	0	0	
6	Valencia	250300 habitantes	0	0	0	0	
7	Valencia	250300 habitantes	0	0	0	0	
8	Valencia	250300 habitantes	0	0	0	0	
9	Valencia	250300 habitantes	0	0	0	0	
10	Valencia	250300 habitantes	0	0	0	0	
11	Valencia	250300 habitantes	0	0	0	0	

En la ventana que se abrirá seleccionamos la carpeta en la que estamos trabajando. Recordemos que estamos guardando nuestros datos en la carpeta CIS en la unidad [K:]



Al fichero activo lo vamos a llamar CIS2927, tendrá como extensión por defecto "SAV". Después de pulsar la opción **Guardar** habremos generado el fichero activo. En próximas sesiones simplemente abriendo [K:\CIS\CIS2927.SAV] accederemos a la matriz ya con datos y etiquetas.

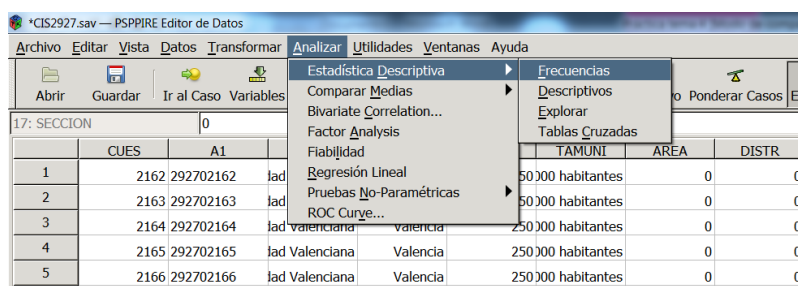
Segunda Parte: Obtención de estadísticos de resumen

Desde la parrilla del fichero activo podemos comenzar a solicitar datos sobre la encuesta. En concreto estamos interesados en conocer las valoraciones dadas a diferentes políticos. En concreto la pregunta 17, en la que se valoran desde 0 a 10 a un total de 14 políticos.

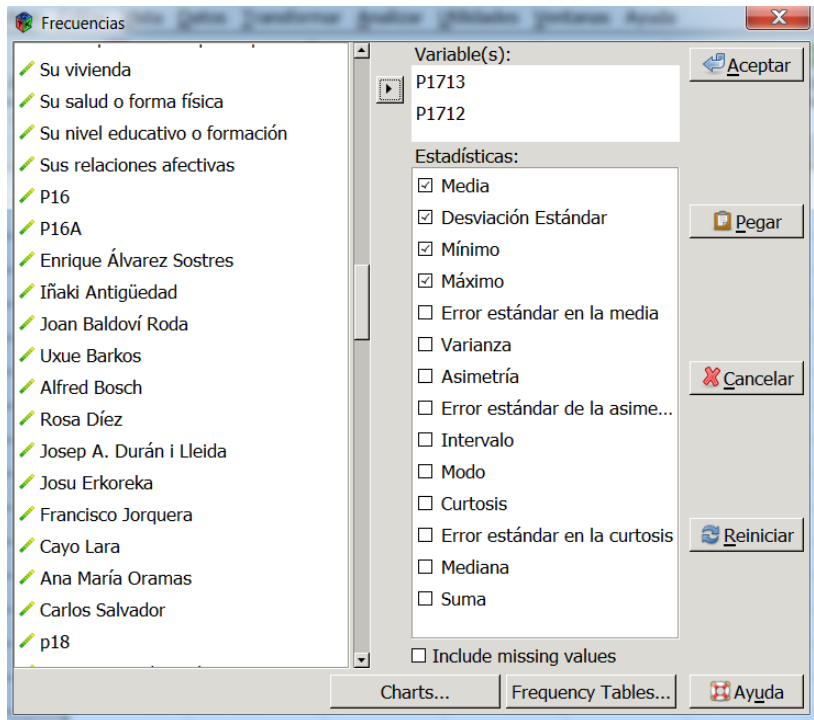
	No conoce	Valoración	NS	NC	
	97	00-10	98	99	
- Enrique Álvarez Sostres.....	97	___	98	99	(85)(86)
- Iñaki Antigüedad	97	___	98	99	(87)(88)
- Joan Baldoví Roda	97	___	98	99	(89)(90)
- Uxue Barkos	97	___	98	99	(91)(92)
- Alfred Bosch.....	97	___	98	99	(93)(94)
- Rosa Díez	97	___	98	99	(95)(96)
- Josep A. Durán i Lleida	97	___	98	99	(97)(98)
- Josu Erkoreka	97	___	98	99	(99)(100)
- Francisco Jorquera.....	97	___	98	99	(101)(102)
- Cayo Lara.....	97	___	98	99	(103)(104)
- Ana María Oramas	97	___	98	99	(105)(106)
- Alfredo Pérez Rubalcaba	97	___	98	99	(107)(108)
- Mariano Rajoy.....	97	___	98	99	(109)(110)
- Carlos Salvador	97	___	98	99	(111)(112)

Antes de solicitar cualquier estadístico conviene observar los datos, para ello vamos a solicitar la distribución de frecuencias. Como forma expositiva, vamos a analizar los datos de Mariano Rajoy y de Alfredo Pérez Rubalcaba. Conviene que el alumno, realice la práctica con la valoración de otros dos políticos.

Para la obtención de frecuencias nos iremos –desde la ventana del fichero activo- a **Analizar**, opción **Estadística Descriptiva**, para seleccionar **Frecuencias**



En la ventana que se abre seleccionamos a ambos políticos y los enviamos a la caja de **Variable(s)**, en dicha caja aparecerán con el nombre de variable P1713 que corresponde a Mariano Rajoy (pregunta 17, opción 13) y P1712 que corresponde a Alfredo Pérez Rubalcaba (pregunta 17 opción 12).



Pulsaremos al botón **A**ceptar y los resultados aparecerán en la ventana de resultados. Con la opción **A**ceptar obtenemos inmediatamente los resultados. Si utilizamos el botón **P**egar en vez de obtener resultados, el programa nos añadirá en el fichero de sintaxis un comando, que deberíamos ejecutar desde dicho fichero. Esto será útil en algunas ocasiones.

FREQ **COMANDO FREQUENCIES**

En el fichero de sintaxis ES2927.SPS ya venía en la última línea un comando para el cálculo de frecuencias:

```
FREQ VAR=CCAA TO P49 P26A VOTOSIM RECUERDO ESTUDIOS OCUPA
CONDICION ESTATUS.
```

Dicho comando pide las distribuciones de frecuencias de la mayor parte de las variables del cuestionario; desde la variable de CCAA hasta la pregunta 49 –partícula TO-, la última del cuestionario, más distintas variables que han sido generadas a partir de los datos como ESTUDIOS por ejemplo.

La sintaxis resulta sencilla el comando FREQUENCIES se puede abreviar escribiendo únicamente las tres primeras letras FRE, en este caso utiliza 4, FREQ y especificando en el subcomando VAR= la lista de variables.

En el fichero ES2927.SPS resulta conveniente borrar dicha instrucción.

En el caso que nos ocupa podemos escribir mediante sintaxis la demanda de frecuencias para las valoraciones de ambos políticos

```
FRE VAR=P1713 P1712.
```

(No hay que olvidar el punto final)

Los resultados que vamos a obtener son:

Mariano Rajoy

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	0	350	14,11	14,11	14,11
	1	82	3,31	3,31	17,42
	2	126	5,08	5,08	22,50
	3	182	7,34	7,34	29,84
	4	200	8,06	8,06	37,90
	5	456	18,39	18,39	56,29
	6	241	9,72	9,72	66,01
	7	214	8,63	8,63	74,64
	8	190	7,66	7,66	82,30
	9	76	3,06	3,06	85,36
	10	96	3,87	3,87	89,23
No conoce	97	10	,40	,40	89,64
N.S.	98	205	8,27	8,27	97,90
N.C	99	52	2,10	2,10	100,00
<i>Total</i>		2480	100,0	100,0	

Mariano Rajoy

<i>N</i>	<i>Válido</i>	2480
	<i>Perdidos</i>	0
<i>Media</i>		14,63
<i>Desv Std</i>		29,15
<i>Mínimo</i>		,00
<i>Máximo</i>		99,00

Alfredo Pérez Rubalcaba

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	0	379	15,28	15,28	15,28
	1	84	3,39	3,39	18,67
	2	160	6,45	6,45	25,12
	3	175	7,06	7,06	32,18
	4	240	9,68	9,68	41,85
	5	465	18,75	18,75	60,60
	6	226	9,11	9,11	69,72
	7	220	8,87	8,87	78,59
	8	153	6,17	6,17	84,76
	9	62	2,50	2,50	87,26
	10	54	2,18	2,18	89,44
No conoce	97	24	,97	,97	90,40
N.S.	98	188	7,58	7,58	97,98
N.C	99	50	2,02	2,02	100,00
<i>Total</i>		2480	100,0	100,0	

Alfredo Pérez Rubalcaba

<i>N</i>	<i>Válido</i>	2480
	<i>Perdidos</i>	0
<i>Media</i>		14,17
<i>Desv Std</i>		28,97
<i>Mínimo</i>		,00
<i>Máximo</i>		99,00

Las tablas anteriores nos muestran la distribución de frecuencias para cada uno de los valores desde 0 a 10 más los códigos No conoce (97) No sabe (98) No contesta (99). Además de las frecuencias absolutas, tenemos el porcentaje, los porcentajes válidos –que explicaremos más adelante su significado- y la distribución porcentual acumulada basada en los porcentajes válidos.

Después de la tabla de frecuencias aparece una tabla con diferentes estadísticos. Nos dice que N, el número de casos, ha sido de 2480 para las dos variables, los 2480 han sido válidos porque no hay ningún caso perdido. Es decir, para todos los cuestionarios se ha relleno dicha variable, bien con una valoración, bien con un código de conocimiento o de no respuesta. Sin embargo, hay situaciones en que puede que no exista para algunos casos valor en dicha variable, por ejemplo porque no es procedente la respuesta o no está definida la variable (por ejemplo lugar de trabajo para personas en paro).

A continuación nos ofrece las medias y las desviaciones típicas, así como los mínimos y máximos.

Probablemente el lector haya pensado que algo falla. Por ejemplo M. Rajoy tiene una valoración media de 14,63 y A. Pérez de 14,17. En ambos casos la media de la puntuación

entre 0 y 10 es mayor de 10. Resulta evidente que no tiene sentido dicho resultado. El problema está en que los códigos 97,98 y 99 se han computado también como valores.

Para calcular los estadísticos de la valoración deberemos seleccionar únicamente a aquellos entrevistados que han valorado a los políticos. Para realizar dicha operación lo haremos en dos fases; primero haremos una copia de la variable, y segundo sobre la variable copiada, declararemos inválidos los valores 97,98 y 99.

Dicho conjunto de operaciones lo vamos a realizar con comandos de sintaxis. Los siguientes comandos los escribiremos al final de nuestro fichero de sintaxis (ES2927.SPS)

En primer lugar para copiar las variables utilizamos el comando **COMPUTE**

```
COMPUTE P1713V=P1713.
```

```
COMPUTE P1712V=P1712.
```

La sintaxis del comando **COMPUTE** es sencilla, la nueva variable = vieja variable. (Observe que cada comando como es habitual termina con un punto). En una columna nueva hemos copiado los valores de la variable P1713 y a esa nueva variable la hemos denominado P1713V. (Hemos elegido el mismo nombre de variable añadiendo la V para indicar que dicha variable será la que contenga los valores).

Si ejecutamos ambos comandos, veremos que al final de la rejilla del fichero activo se han añadido dos nuevas columnas después de la variable STATUS que era la última. De momento todos los casos tienen el valor 0.

Observemos que en la barra de abajo aparece la siguiente indicación:

ESTUDIOS	OCUPA	CONDICION	ESTATUS	P1713V	P1712V
3	5	9	4	,00	,00
6	3	5	2	,00	,00
2	3	8	2	,00	,00
3	3	9	2	,00	,00
2	7	6	4	,00	,00
5	2	2	1	,00	,00
3	2	1	1	,00	,00
6	2	2	1	,00	,00
4	3	9	2	,00	,00
4	8	6	4	,00	,00
2	4	5	2	,00	,00
3	1	8	3	,00	,00
2	7	8	3	,00	,00
2	9	9	5	,00	,00
2	7	11	4	,00	,00
2	9	7	5	,00	,00
1	9	8	5	,00	,00
2	6	5	3	,00	,00
5	4	9	2	,00	,00
4	4	5	2	,00	,00
2	9	9	5	,00	,00
5	3	9	1	,00	,00
3	9	7	5	,00	,00
4	9	7	5	,00	,00
2	7	11	4	,00	,00

Transformaciones pendientes SIN Filtro SIN Ponderar NO Dividido

Cuando encontremos el siguiente texto, deberemos en la barra superior seleccionar la opción **Transformar** y elegir **Ejecutar Transformaciones pendientes**

Después de esta operación veremos que ambas columnas contienen una copia de los valores de las variables originales.

Una vez copiadas las variables, vamos a eliminar los valores 97,98 y 99. Para ello utilizaremos el comando **MISSING VALUES**

MISSING VALUES P1713V P1712V (97,98,99).

El comando MISSING VALUES especifica las variables y entre paréntesis el conjunto de valores que se declaran perdidos separados por comas. Como todos los comandos termina con un punto.

En el fichero de sintaxis (ES2917.SPS) el segundo comando es MISSING VALUE.

```
MISSING VALUES P0 TO P16a P18 to p21a03 p23 to p4304 P46 TO P49  
P26A
```

```
VOTOSIM RECUERDO ESTUDIOS OCUPA CONDICION ESTATUS(0).
```

La ejecución de dicho comando declara como valor perdido el código 0 para el conjunto de variables especificado.

Para terminar nuestra operación podemos añadir un comando variable label para etiquetar las nuevas variables. A continuación escribimos la secuencia de sintaxis. Añadiremos también el comando *. Este comando significa comentario y termina también con un punto. Desde el asterisco hasta el el punto podemos introducir el texto y anotaciones que queramos en el fichero de sintxis. (Habitualmente para que sea más visible en vez de un asterisco se utilizan más).

****Sintaxis para la transformación de las variables de valoración de políticos para la obtención de puntuaciones medias. Primero con COMPUTE se hace una copia de las variables. Después sobre las nuevas variables se declaran los códigos perdidos. Finalmente se etiquetan las nuevas variables.

```
COMPUTE P1713V=P1713.  
COMPUTE P1712V=P1712.
```

```
MISSING VALUES P1713V P1712V (97,98,99).
```

```
VARIABLE LABEL P1713V "Valoración de Mariano Rajoy"  
/ P1712V "Valoración de Alfredo Pérez Rubalcaba".
```

Ahora simplemente es cuestión de solicitar la distribución de frecuencias de P1713V y de P1712V. Recordemos que desde la parrilla de datos, seleccionamos **Analizar: Estadística Descriptiva: Frecuencias**, e incluimos las dos variables, que las encontraremos al final del listado. Los resultados ahora serán:

Valoración de Mariano Rajoy

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	,00	350	14,11	15,82	15,82
	1,00	82	3,31	3,71	19,52
	2,00	126	5,08	5,69	25,21
	3,00	182	7,34	8,22	33,44
	4,00	200	8,06	9,04	42,48
	5,00	456	18,39	20,61	63,08
	6,00	241	9,72	10,89	73,97
	7,00	214	8,63	9,67	83,64
	8,00	190	7,66	8,59	92,23
	9,00	76	3,06	3,43	95,66
	10,00	96	3,87	4,34	100,00
	97,00	10	,40	Perdidos	
	98,00	205	8,27	Perdidos	
	99,00	52	2,10	Perdidos	
<i>Total</i>		2480	100,0	100,0	

Valoración de Mariano Rajoy

<i>N</i>	<i>Válido</i>	2213
	<i>Perdidos</i>	267
<i>Media</i>		4,55
<i>Desv Std</i>		2,86
<i>Mínimo</i>		,00
<i>Máximo</i>		10,00

Valoración de Alfredo Pérez Rubalcaba

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	,00	379	15,28	17,09	17,09
	1,00	84	3,39	3,79	20,87
	2,00	160	6,45	7,21	28,09
	3,00	175	7,06	7,89	35,98
	4,00	240	9,68	10,82	46,80
	5,00	465	18,75	20,96	67,76
	6,00	226	9,11	10,19	77,95
	7,00	220	8,87	9,92	87,87
	8,00	153	6,17	6,90	94,77
	9,00	62	2,50	2,80	97,57
	10,00	54	2,18	2,43	100,00
	97,00	24	,97	Perdidos	
	98,00	188	7,58	Perdidos	
	99,00	50	2,02	Perdidos	
<i>Total</i>		2480	100,0	100,0	

Valoración de Alfredo Pérez Rubalcaba

<i>N</i>	<i>Válido</i>	2218
	<i>Perdidos</i>	262
<i>Media</i>		4,25
<i>Desv Std</i>		2,75
<i>Mínimo</i>		,00
<i>Máximo</i>		10,00

Ahora podemos comprender la diferencia entre la distribución de porcentajes y la de "Porcentajes Válidos". Por ejemplo los 456 entrevistados que han valorado con un 5 a Mariano Rajoy, son el 18,39% de la muestra (456/2480) pero suponen el 20,61% (456/2213) de quienes han valorado a dicho político que han sido 2213 entrevistados.

Los porcentajes acumulados, están basados en los porcentajes de casos válidos. Podemos observar por ejemplo que un 83,64% han valorado a Rajoy con 7 o menos, mientras que dicha cifra es mayor para Rubalcaba 87,87%. La distribución señala que en el caso de Rajoy hay relativamente más puntuaciones altas que en el caso de Rubalcaba.

Ahora las medias son reales, están hechas sobre el número de valoraciones. Podemos observar que desde 0 (como valor mínimo) a 10 (como valor máximo). La valoración de Rajoy es 4,55 ligeramente superior al 4,25 obtenido por Rubalcaba. La tabla también nos ofrece la desviación típica para ambas distribuciones.

En la tabla siguiente encontrará los resultados de valoración al conjunto de políticos incluidos en la pregunta 17. Seleccione un par de ello y reproduzca el prodimiento anterior para encontrar estos resultados.

	Media	Desviación típica	(N)
Enrique Álvarez Sostres	3.71	2.35	(93)
Iñaki Antigüedad	3.08	2.63	(124)
Joan Baldoví Roda	3.54	2.51	(69)
Uxue Barkos	4.41	2.48	(276)
Alfred Bosch	4.02	2.49	(116)
Rosa Díez	4.75	2.45	(1483)
Josep A. Durán i Lleida	4.13	2.44	(1296)
Josu Erkoreka	3.38	2.47	(421)
Francisco Jorquera	3.89	2.42	(121)
Cayo Lara	4.17	2.44	(1146)
Ana María Oramas	4.09	2.61	(136)
Alfredo Pérez Rubalcaba	4.25	2.75	(2218)
Mariano Rajoy	4.55	2.86	(2213)
Carlos Salvador	3.71	2.52	(84)

Para terminar la práctica nos centraremos en el cálculo de otras medidas de tendencia central como son la moda y la mediana. Para ello utilizaremos la pregunta 30, que contiene la edad de los entrevistados. Observemos como es habitual que el código 99 está reservado a los No Contesta. Por lo tanto, copiamos la variable en una nueva –por ejemplo EDAD- y declaramos *missing* el valor 99.

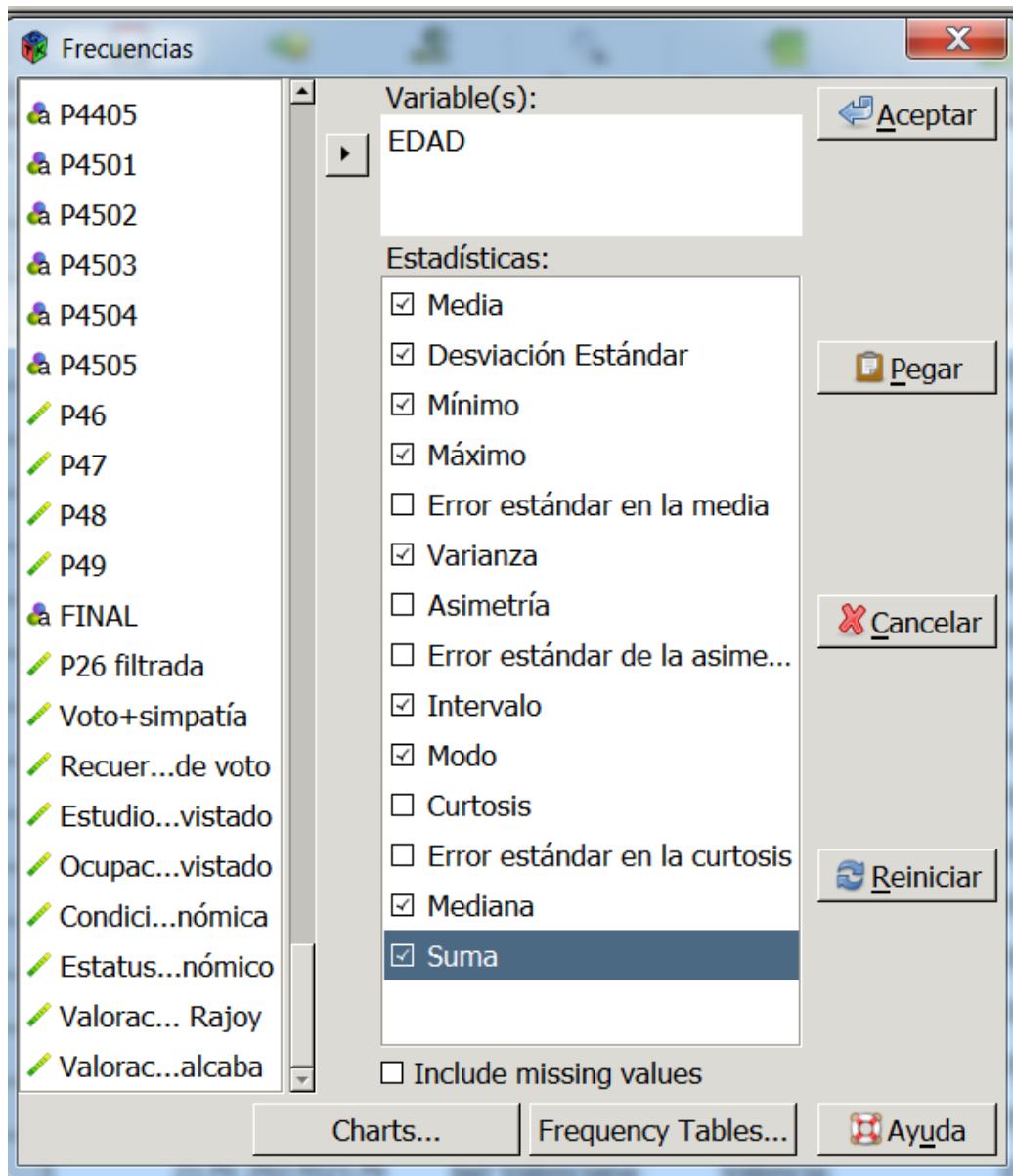
****Instrucciones para generar la variable EDAD.

COMPUTE EDAD=P30.

MISSING VALUE EDAD (99).

Una vez ejecutados los comandos anteriores pedimos los resultados **Analizar: Estadística Descriptiva: Frecuencias**, como variable incluimos EDAD y marcamos todos los estadísticos que hemos estudiado:

Media
Desviación Estándar
Mínimo
Máximo
Varianza
Intervalo
Modo²
Mediana
Suma



Obtenemos la tabla de frecuencias y a continuación los estadísticos:

² Es un error en la tipografía de SPSS quiere decir Moda.

EDAD

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	18,00	28	1,1	1,1	1,1
	19,00	36	1,5	1,5	2,6
	20,00	24	1,0	1,0	3,6
	21,00	41	1,7	1,7	5,2
	22,00	27	1,1	1,1	6,3
	23,00	35	1,4	1,4	7,7
	24,00	30	1,2	1,2	8,9
	25,00	61	2,5	2,5	11,4
	26,00	43	1,7	1,7	13,1
	27,00	34	1,4	1,4	14,5
	28,00	45	1,8	1,8	16,3
	29,00	26	1,0	1,0	17,4
	30,00	41	1,7	1,7	19,0
	31,00	54	2,2	2,2	21,2
	32,00	52	2,1	2,1	23,3
	33,00	55	2,2	2,2	25,5
	34,00	73	2,9	2,9	28,5
	35,00	48	1,9	1,9	30,4
	36,00	48	1,9	1,9	32,3
	37,00	44	1,8	1,8	34,1
	38,00	43	1,7	1,7	35,8
	39,00	43	1,7	1,7	37,6
	40,00	57	2,3	2,3	39,9
	41,00	44	1,8	1,8	41,6
	42,00	46	1,9	1,9	43,5
	43,00	56	2,3	2,3	45,8
	44,00	76	3,1	3,1	48,8
	45,00	56	2,3	2,3	51,1
	46,00	54	2,2	2,2	53,3
	47,00	47	1,9	1,9	55,2
	48,00	31	1,3	1,3	56,4
	49,00	36	1,5	1,5	57,9
	50,00	45	1,8	1,8	59,7
	51,00	39	1,6	1,6	61,3
	52,00	38	1,5	1,5	62,8
	53,00	50	2,0	2,0	64,8
	54,00	39	1,6	1,6	66,4
	55,00	30	1,2	1,2	67,6
	56,00	28	1,1	1,1	68,7
	57,00	31	1,3	1,3	70,0
	58,00	31	1,3	1,3	71,2
	59,00	34	1,4	1,4	72,6
	60,00	39	1,6	1,6	74,2
	61,00	26	1,0	1,0	75,2
	62,00	33	1,3	1,3	76,6
	63,00	36	1,5	1,5	78,0
	64,00	42	1,7	1,7	79,7
	65,00	41	1,7	1,7	81,4
	66,00	26	1,0	1,0	82,4
	67,00	30	1,2	1,2	83,6
	68,00	30	1,2	1,2	84,8
	69,00	23	,9	,9	85,8
	70,00	40	1,6	1,6	87,4
	71,00	28	1,1	1,1	88,5
	72,00	16	,6	,6	89,1
	73,00	19	,8	,8	89,9
	74,00	25	1,0	1,0	90,9
	75,00	24	1,0	1,0	91,9
	76,00	21	,8	,8	92,7
	77,00	27	1,1	1,1	93,8
	78,00	18	,7	,7	94,6
	79,00	16	,6	,6	95,2
	80,00	17	,7	,7	95,9
	81,00	20	,8	,8	96,7
	82,00	17	,7	,7	97,4
	83,00	9	,4	,4	97,7
	84,00	13	,5	,5	98,3
	85,00	12	,5	,5	98,7
	86,00	8	,3	,3	99,1
	87,00	5	,2	,2	99,3
	88,00	7	,3	,3	99,6
	89,00	5	,2	,2	99,8
	90,00	2	,1	,1	99,8
	91,00	1	,0	,0	99,9
	92,00	1	,0	,0	99,9
	94,00	2	,1	,1	100,0
	Total	2478	99,9	100,0	
Perdidos	99,00	2	,1		
Total		2480	100,0		

EDAD

<i>N</i>	<i>Válido</i>	2478
	<i>Perdidos</i>	2
<i>Media</i>		47,48
<i>Modo</i>		44,00
<i>Desv Std</i>		17,79
<i>Varianza</i>		316,31
<i>Err.Est.Curt.</i>		,10
<i>Err.Est.Asim.</i>		,05
<i>Intervalo</i>		76,00
<i>Mínimo</i>		18,00
<i>Máximo</i>		94,00
<i>Suma</i>		117662,0
<i>Percentiles</i>	50 (Mediana)	45,50

En total hay dos casos perdidos, siendo 2478 los casos válidos –de los que tenemos información- que intervienen en el cálculo de los estadísticos.

Media: Nos dice que la edad media de la muestra es de 47,48 años (Téngase en cuenta que el universo está compuesto por mayores de 18 años.) Es el resultado de la Suma (117662) entre N Válido (2478).

Moda: El mayor número de entrevistados tienen 44 años. En total hay 76.

Desviación Estándar: Aunque el programa la denomina así se trata de la desviación típica muestral o cuasidesviación. (Consulte el tema VIII del texto *Estadística para la Investigación Social*). Para estos tamaños muestrales dado que la diferencia de cálculo es en el denominador, que utiliza N-1 en vez de N, no hay diferencias en el resultado.

Varianza: Es el cuadrado de la Desviación Estándar.

Intervalo: Diferencia entre el mayor valor 94 y el menor 18. Es el Recorrido de la variable.

Suma: Es la suma de todos los valores –edades- de la muestra. En total suman 117662 años.

Percentil 50 o Mediana: 45,5 años es el valor que separa la muestra en dos grupos iguales. Los diferentes programas utilizan algoritmos distintos para obtener la mediana u otros Cuantiles³.

³Rob J. Hyndman and Yanan Fan en su artículo “Sample Quantiles in Statistical Packages” publicado en la revista *The American Statistician*, Vol. 50, No. 4 (Nov., 1996), pp. 361-365, repasan los algoritmos utilizados en los principales paquetes estadísticos (BMDP, GLIM, SAS, Splus, SPSS, etc.) para calcular la mediana o los cuantiles y encuentran hasta 9 sistemas diferentes, en gran parte de los cuales no se hace público su sistema de cálculo. Los autores denuncian la falta de estandarización de algoritmos que provocan confusión en los usuarios.