

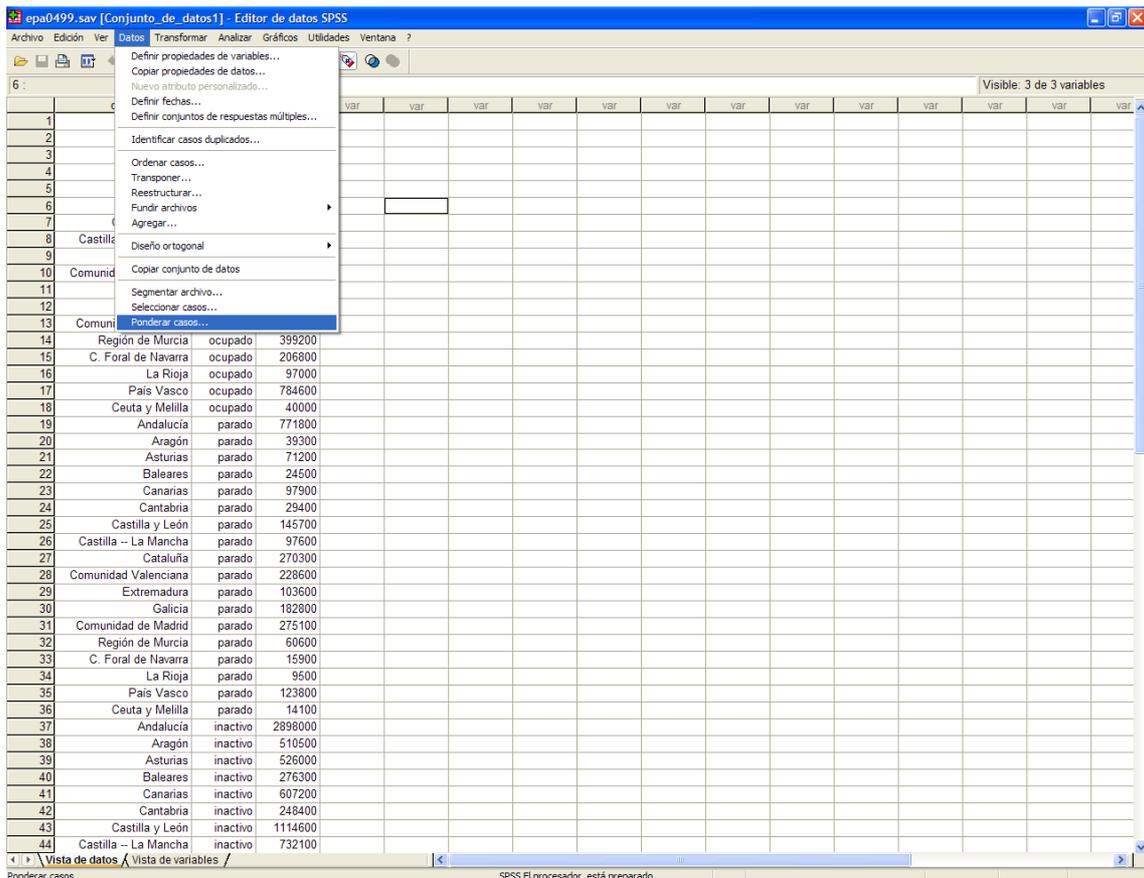
Análisis de correspondencias

utilizando el SPSS

Vamos a realizar el Análisis de correspondencias sobre el mismo ejemplo (ejemplo 6.1 de los apuntes), sobre el que ya se obtuvo el análisis mediante el R, pero ahora empleando el paquete estadístico SPSS.

Lo primero que conviene recordar es que la hoja de datos de SPSS mantiene la estructura de filas, que representan a los individuos, y columnas, que representan a las variables. En consecuencia, no se puede situar en la hoja de datos una tabla de contingencia donde las filas sean las categorías de una variable, y las columnas las categorías de la otra variable. La estructura de individuos por variables, obligaría a introducir tantas filas como individuos, pero eso se puede evitar haciendo uso de

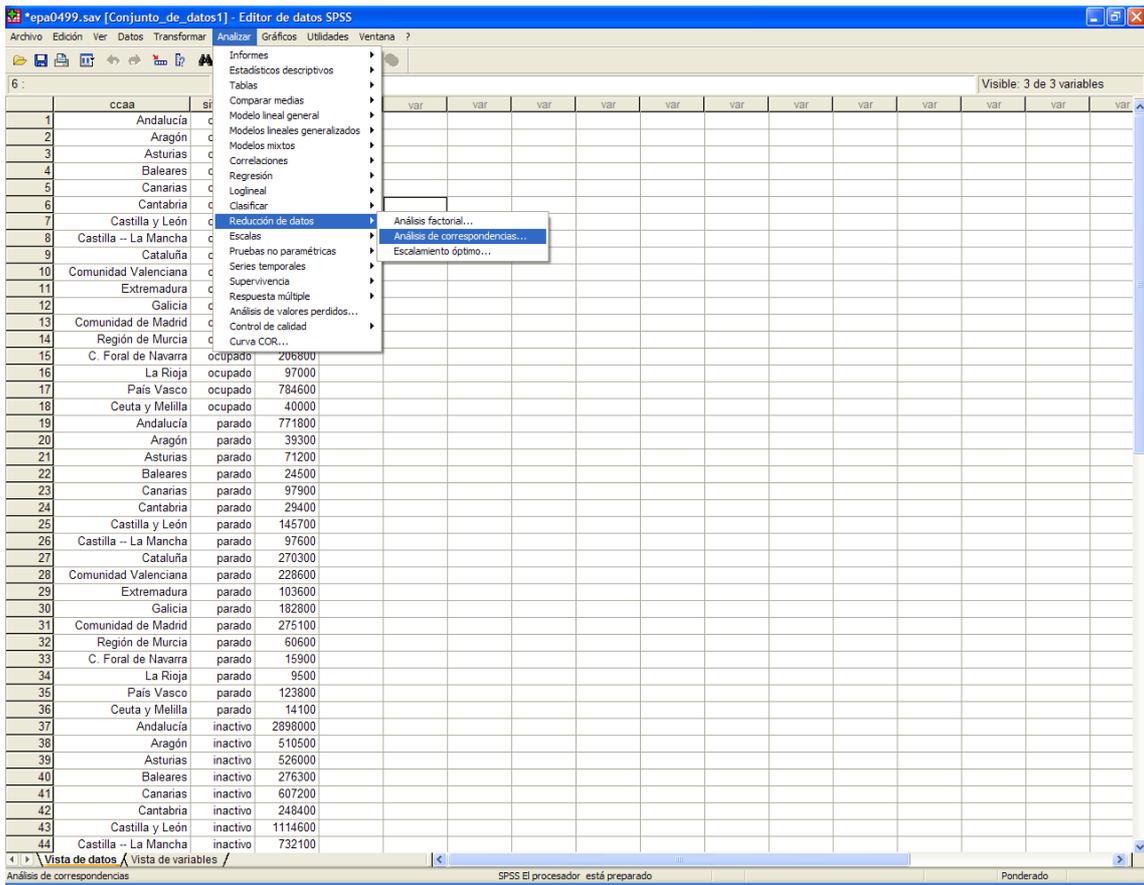
Datos – Ponderar casos ...



En el menú de diálogo que se abre tras efectuar la selección anterior, se escogería la variable que contiene las frecuencias de cada combinación de niveles de las variables categóricas.

Una vez situados los datos en la hoja, y efectuada la ponderación (lo cual deja una leyenda en el margen inferior, en la zona derecha), para obtener un Análisis de correspondencias se pincha en

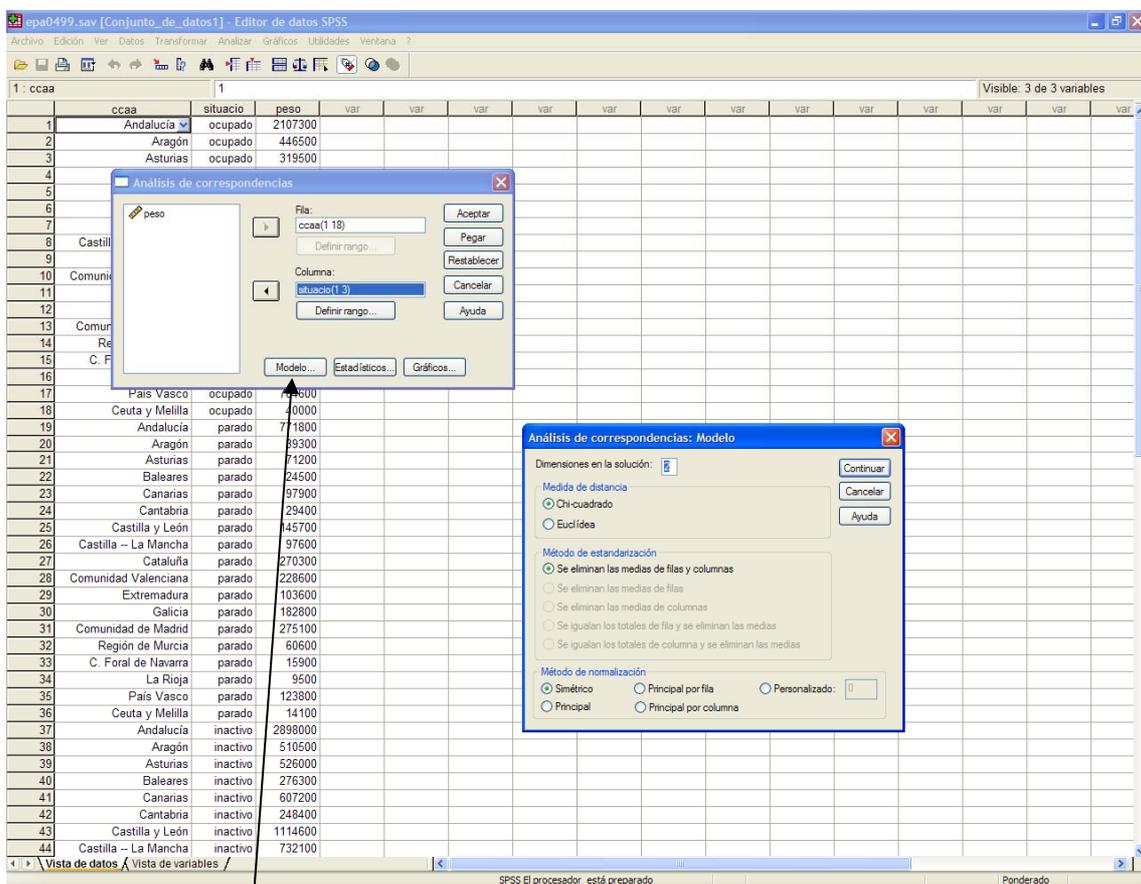
Analizar – Reducción de datos – Análisis de correspondencias ...



Indica que el fichero está ponderado

Tras ejecutar las orden anterior, se accede a un cuadro de diálogo en el que, como es costumbre en SPSS, las variables del fichero de datos figuran en un cuadro a la izquierda. Seleccionamos la variable que deseamos que figure en las filas de la tabla de contingencia (en este caso la comunidad autónoma), indicamos el rango de valores de la variables (de 1 a 18), y hacemos lo mismo con la variable de columna (la situación respecto del empleo).

Aparecen cinco botones a la derecha (Aceptar, Pegar, Restablecer, Cancelar, Ayuda), comunes a cualquier procedimiento de SPSS. En la parte inferior figuran tres botones específicos del procedimiento. En el caso del “Análisis de correspondencias” estos botones son Modelo, Estadísticos y Gráficos. Cada uno de ellos permite configurar cierto conjunto de características del método que se desea aplicar.



Al pinchar en “**Modelo...**” se abre un nuevo cuadro de diálogo, en el que vamos a dejar las opciones por defecto, que son:

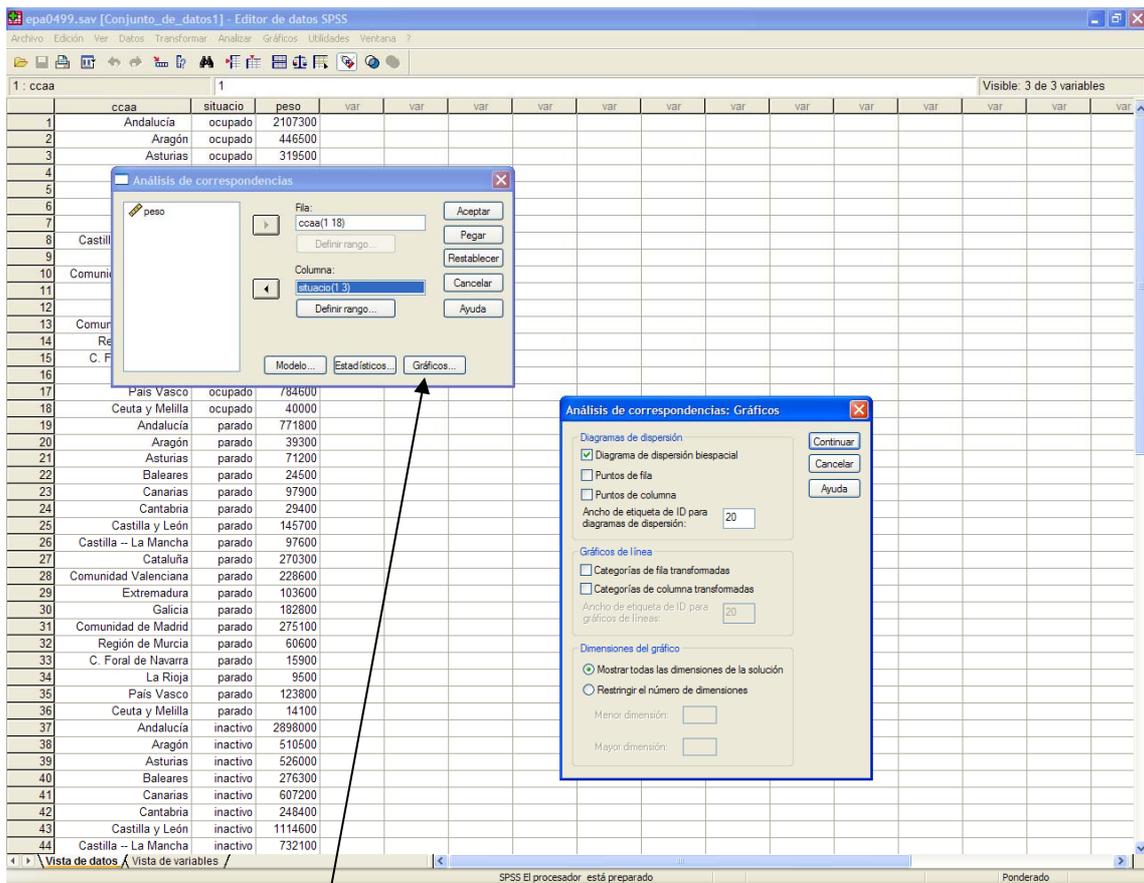
1. Dos dimensiones en la solución
2. Se centra restando las medias de fila y columna
3. El método de normalización es simétrico, que es el que permite la representación simultánea y se corresponde con los apuntes.

The screenshot shows the SPSS interface with a data editor window displaying a contingency table. The table has columns for 'ccaa', 'situacio', and 'peso'. The 'situacio' variable has categories 'ocupado' and 'parado'. The 'peso' variable has values ranging from 4000 to 1114600. Two dialog boxes are open: 'Análisis de correspondencias' and 'Análisis de correspondencias: Estadísticos'. The 'Análisis de correspondencias' dialog box shows 'peso' as the row variable and 'situacio(1,3)' as the column variable. The 'Análisis de correspondencias: Estadísticos' dialog box has several options checked, including 'Tabla de correspondencias', 'Inspección de los puntos de fila', 'Inspección de los puntos de columna', 'Perfiles de fila', and 'Perfiles de col.'. The 'Estadísticos...' button in the first dialog box is highlighted with a red arrow pointing to the second dialog box.

ccaa	situacio	peso	
1	Andalucía	ocupado	2107300
2	Aragón	ocupado	446500
3	Asturias	ocupado	319500
4			
5			
6			
7			
8	Castilla		
9			
10	Comuni		
11			
12			
13	Comun		
14	Re		
15	C. F.		
16			
17	Pais Vasco	ocupado	784600
18	Ceuta y Melilla	ocupado	4000
19	Andalucía	parado	771800
20	Aragón	parado	39300
21	Asturias	parado	71200
22	Baleares	parado	24500
23	Canarias	parado	97900
24	Cantabria	parado	29400
25	Castilla y León	parado	145700
26	Castilla -- La Mancha	parado	97600
27	Cataluña	parado	270300
28	Comunidad Valenciana	parado	228600
29	Extremadura	parado	103600
30	Galicia	parado	182800
31	Comunidad de Madrid	parado	275100
32	Región de Murcia	parado	60600
33	C. Foral de Navarra	parado	15900
34	La Rioja	parado	9500
35	Pais Vasco	parado	123800
36	Ceuta y Melilla	parado	14100
37	Andalucía	inactivo	2898000
38	Aragón	inactivo	510500
39	Asturias	inactivo	526000
40	Baleares	inactivo	276300
41	Canarias	inactivo	607200
42	Cantabria	inactivo	248400
43	Castilla y León	inactivo	1114600
44	Castilla -- La Mancha	inactivo	732100

En el submenú de “Estadísticos...” solicitamos:

1. Tabla de correspondencias
2. Inspección de los puntos de fila
3. Inspección de los puntos de columna
4. Perfiles de fila
5. Perfiles de columna



En el submenú de **“Gráficos...”** solicitamos las opciones por defecto, que son:

1. El diagrama de dispersión bieespacial
2. Que muestre todas las dimensiones de la solución

Una vez realizadas las selecciones que hemos indicado, se puede ejecutar el procedimiento. Se proporcionarán los resultados del análisis en la hoja de resultados de SPSS.

La mayor parte de los resultados proporcionados por SPSS se entienden a la luz de los contenidos expuestos en los apuntes, y coinciden con los que se obtienen con el programa R.

La principal diferencia consiste en que el SPSS aplica un enfoque basado en el análisis de correlación canónica, lo cual da lugar a un valor propio, que coincide con la raíz cuadrada del valor propio de los apuntes, que SPSS denomina Inercia (igual que en los apuntes).

Las puntuaciones en la dimensión se corresponden con las matrices A y B, en cada caso, con la diferencia de que mientras las puntuaciones de las matrices A y B están estandarizadas de modo que la suma ponderada de cuadrados sea el autovalor (en términos de inercia), las puntuaciones están estandarizadas para que dicha suma sea el autovalor (de la correlación canónica).

Las tablas de contribuciones consisten en proporciones de inercia por filas, por columnas o por componentes, donde la inercia bruta coincide con lo que figura en los apuntes.

Nótese que el enfoque de correlación canónica también es aplicado por el comando "corresp" de R, que por este motivo, da los mismos resultados que SPSS.