

TEXTOS DOCENTS

300

ANÁLISIS ESTADÍSTICO MEDIANTE APLICACIONES INFORMÁTICAS. SPSS, STATGRAPHICS, MINITAB Y EXCEL

Lluís Salafranca Cosialls
Vicenta Sierra Olivera
M. Isabel Núñez Peña
Antonio Solanas Pérez
David Leiva Ureña

Departament de Metodologia de les Ciències del Comportament
Facultat de Psicologia

TEXTOS DOCENTS

300

ANÁLISIS ESTADÍSTICO MEDIANTE APLICACIONES INFORMÁTICAS. SPSS, STATGRAPHICS, MINITAB Y EXCEL

Lluís Salafranca Cosialls
Vicenta Sierra Olivera*
M. Isabel Núñez Peña
Antonio Solanas Pérez
David Leiva Ureña

Departament de Metodologia de les Ciències del Comportament
Facultat de Psicologia

* Departament de Mètodes Quantitatius i de Gestió
ESADE. Universitat Ramon Llull

Publicacions i Edicions



UNIVERSITAT DE BARCELONA



INTRODUCCIÓN

El presente texto se ha realizado para favorecer el aprendizaje de la Estadística y de algunas aplicaciones informáticas desarrolladas para el análisis de datos. El objetivo principal consiste en facilitar esta tarea aprovechando las posibilidades que ofrecen los recursos informáticos. No es un texto sólo de Estadística ni tampoco está exclusivamente ideado para aprender algunas aplicaciones informáticas, pues existen varios libros ya publicados con este objetivo. Es un texto con el cual se pretende que los alumnos aprendan Estadística omitiendo cualquier tipo de cálculo y, en cambio, se habitúen a la utilización de aplicaciones informáticas estadísticas.

A partir de unos datos correspondientes a una supuesta investigación, se presentan diferentes ejercicios resueltos de Estadística. El lector encontrará que se le guía continuamente sobre cómo realizar el análisis estadístico mediante la aplicación informática y, posteriormente, se le ayuda en la interpretación de los resultados que él o ella misma va alcanzando. Por tanto, a la vez que adquiere conocimientos sobre el manejo de los diferentes procedimientos de las aplicaciones estadísticas informatizadas, aprende cómo abordar algunos problemas de investigación mediante las técnicas estadísticas y la interpretación de los resultados obtenidos. También existen unos ejercicios propuestos, no resueltos, que intentan sugerir otros problemas y animar al lector a resolverlos con la formación que ya ha adquirido mediante los ejercicios resueltos. En la resolución de estos problemas se pretende, en mayor medida, que el lector ejercite un enfoque reflexivo e interpretativo sobre los resultados obtenidos, en lugar de una mera exposición de los resultados gráficos o numéricos proporcionados por la aplicación informática.

Durante el curso 1999-2000 se sometió a prueba la viabilidad, aportación, operatividad y resultados de este modelo de formación en el ámbito de la Estadística, concretándose en un texto precursor. Sus contenidos se utilizaron como parte del *Treball propi* que los alumnos debían realizar y enmarcado en la asignatura de *Anàlisi de Dades en Psicologia*. En el mencionado curso académico se planificaron una serie de asesorías para que los alumnos tuvieran un apoyo en las tareas que implicaban el seguimiento del texto. En el año 2000 se editó finalmente el texto revisado. La experiencia paulatina a lo largo de los años confirmó el carácter autoexplicativo del texto, favoreciendo su trabajo sin la necesidad de asesorías, con unos mínimos conocimientos informáticos y estadísticos correspondientes al nivel de primer ciclo de la licenciatura, por lo que estas asesorías fueron remplazadas por consultas puntuales. Una mención de agradecimiento debe ser realizada al conjunto de alumnos que, a lo largo de diferentes cursos, ejercieron la tarea de monitores, y que con sus aportaciones han contribuido a mejorar los contenidos y propuesta que se presenta en este nuevo libro.

La evolución de los contenidos y enfoques docentes de la asignatura *Anàlisi de Dades en Psicologia*, ante el reto que representa la convergencia europea, así como la necesaria revisión de las versiones de los programas utilizadas inicialmente ha derivado en la edición de este nuevo texto. En éste se ha considerado ampliar el número de aplicaciones informáticas, incorporando a las iniciales *SPSS* y *StatGraphics*, las aplicaciones *Minitab* y *Excel*, aumentando igualmente la cantidad de técnicas revisadas, así como los ejercicios, tanto resueltos como propuestos. El presente texto se propone también como *Treball propi* del alumno de la asignatura *Anàlisi de Dades en Psicologia*, sin que ello obste para que consideremos este libro adecuado para cualquier estudiante, profesional o investigador básico o aplicado que tenga interés en mejorar y adquirir conocimientos en el campo de la Estadística Informatizada. Aquellos lectores que estuviesen interesados en el seguimiento del presente texto pueden solicitar a los autores, vía correo electrónico, los ficheros de datos originales en formato *SPSS*, *StatGraphics*, *Minitab* o *Excel*. Evidentemente, los alumnos de la licenciatura de Psicología de la Universitat de Barcelona dispondrán de acceso a estos ficheros a través del *Dossier electrònic* de la asignatura.

Aunque esperamos que el proceso realizado en los años anteriores haya minimizado los problemas y errores que se hallen en el texto, es posible que muchos aspectos sean todavía mejorables. Por este motivo, rogamos a los lectores nos hagan partícipes de aquellas propuestas/comentarios que consideren interesantes a fin de seguir mejorando este material docente en el futuro.

Todos los autores del presente texto deseamos que este material resulte útil en el aprendizaje de la Estadística y, con este objetivo, lo presentamos a todos los potenciales lectores.

Lluís Salafranca¹, Vicenta Sierra², María Isabel Núñez¹, Antoni Solanas¹ y David Leiva¹

¹Departament de Metodologia de les Ciències del Comportament de la Universitat de Barcelona

²Departament de Mètodes Quantitatius i de Gestió. ESADE. Universitat Ramon Llull.

lsalafranca@ub.edu; vicenta.sierra@esade.edu; inunez@ub.edu; antonio.solanas@ub.edu y dleivaur@ub.edu.

Barcelona, Mayo 2005

ÍNDICE

INTRODUCCIÓN	3
ÍNDICE	5
PRESENTACIÓN DEL EJEMPLO	11
CAPÍTULO I	
1. SPSS para Windows (Versión 12.0.1)	13
1.1. Introducción	13
1.2. Ejecución del programa y elementos principales	14
1.3. Lectura de un fichero de datos en formato SPSS (*.SAV)	17
1.4. Gestión y Modificación de variables	22
1.4.1. <i>Generar nuevas variables</i>	22
1.4.2. <i>Recodificación de Variables</i>	24
1.4.3. <i>Otras opciones</i>	26
1.5. Gestión de los datos	27
1.6. Guardar información en un archivo externo e impresión de listados	29
CAPÍTULO II	
2. Análisis estadístico descriptivo univariante	31
2.1. Análisis estadístico descriptivo de la variable cuantitativa ELECTINI	31
2.2. Ejercicio propuesto	37
CAPÍTULO III	
3. Análisis estadístico descriptivo univariante por niveles de una variable de agrupación	39
3.1. Descripción de la variable EDAD en función de los grupos formados por la variable SEXO.	39
3.2. Comentarios respecto a la información de tipo inferencial	46
3.3. Ejercicio propuesto	47
3.3.1. <i>Cuestiones sobre estadística descriptiva</i>	47
3.3.2. <i>Cuestiones sobre estadística inferencial</i>	48
CAPÍTULO IV	
4. Asociación bivalente en tablas de contingencia	49
4.1. Asociación entre variables cualitativas	49
4.1.1. <i>Estudio descriptivo de la asociación entre el SEXO y el SUBTIPO</i>	50
4.1.2. <i>Estudio inferencial de la asociación entre el SEXO y el SUBTIPO</i>	56
4.2. Asociación entre variables ordinales	56
4.3. Ejercicios propuestos	57
4.3.1. <i>Ejercicio sobre asociación de variables cualitativas</i>	57
4.3.2. <i>Ejercicio sobre asociación de variables ordinales</i>	58

CAPÍTULO V

5. Análisis bivariante de variables cuantitativas: índices de asociación y modelos de regresión	59
5.1. Estudio de la asociación bivariante entre las variables cuantitativas medidas al inicio de la investigación, para los sujetos con fobia específica, mediante índices y gráficos	59
5.1.1. <i>Estudio descriptivo</i>	60
5.1.2. <i>Estudio inferencial</i>	65
5.2. Modelo de regresión simple entre las variables ANSIEDAD y PENSINI	65
5.2.1. <i>Modelo de regresión lineal</i>	66
5.2.2. <i>Modelo de regresión no lineal</i>	68
5.3. Ejercicio propuesto	69

CAPÍTULO VI

6. Asociación de una variable cuantitativa y una cualitativa	71
6.1. Comparación de medias para muestras independientes.	72
6.1.1. <i>Comparación de medias de la variable actividad eléctrica dérmica antes del tratamiento (ELECTINI) en las poblaciones control y tratamiento (GRUPO)</i>	74
6.1.2. <i>Comparación de medias de la variable actividad eléctrica dérmica después de la intervención (ELECTFIN) en las poblaciones control y tratamiento (GRUPO)</i>	78
6.2. Comparación de medias para muestras relacionadas	80
6.2.1. <i>Comparación de medias para la variable actividad eléctrica de la piel antes (ELECTINI) y después (ELECTFIN) del tratamiento contra el trastorno fóbico en la población a la cual no se administra intervención (población control)</i>	82
6.2.2. <i>Comparación de medias para la variable actividad eléctrica de la piel antes (ELECTINI) y después (ELECTFIN) del tratamiento contra el trastorno fóbico en la población a la cual se administra intervención (población tratamiento)</i>	82
6.2.3. <i>Condiciones de aplicación</i>	83
6.3. Determinación de la intensidad de asociación.	83
6.4. Ejercicio propuesto	85

CAPÍTULO VII

7. Análisis de la Variancia	87
7.1. Análisis de la variancia para grupos independientes.	87
7.2. Ejercicio propuesto	94

CAPÍTULO VIII

8. StatGraphics 5.1.	95
8.1. Introducción	95
8.2. Tipos de archivos	95
8.3. La ventana principal del sistema	96
8.4. Modificación de las variables	98
8.5. Generación de datos	100
8.6. Codificación de datos	103
8.7. Ordenación de la matriz de datos	105
8.8. Selección de casos.	106
8.9. Ventana de análisis	108

CAPÍTULO IX

9. Distribuciones de probabilidad.	111
9.1. Distribución binomial	111
9.2. Distribución binomial negativa	114
9.3. Distribución normal	116
9.4. Distribución exponencial negativa de un parámetro	118
9.5. Ejercicio propuesto	121

CAPÍTULO X	
10. Descripción de distribuciones empíricas de datos	123
10.1. Análisis descriptivo univariante: variable categórica	123
10.2. Análisis descriptivo univariante: variable numérica	124
10.3. Ejercicio propuesto	130
CAPÍTULO XI	
11. Análisis de componentes principales	131
11.1. Análisis de la dependencia estadística entre las variables originales	131
11.2. Reducción de la dimensión mediante el análisis de componentes principales	133
11.3. Ejercicio propuesto	137
CAPÍTULO XII	
12. Pruebas estadísticas de conformidad	139
12.1. Prueba de conformidad para el valor del parámetro media en una población	139
12.2. Prueba de conformidad para un valor del parámetro proporción de una población	143
12.3. Prueba de conformidad en distribución para una variable aleatoria en la población	146
12.4. Tamaño de la muestra para estimar el valor del parámetro proporción	148
12.5. Ejercicio propuesto	150
CAPÍTULO XIII	
13. Tablas de contingencia	151
13.1. Tablas de contingencia: cuantificación de la asociación	151
13.2. Tablas de contingencia: decisión estadística	153
13.3. Ejercicio propuesto	156
CAPÍTULO XIV	
14. Comparación entre dos poblaciones	157
14.1. Decisión estadística sobre el parámetro diferencia de medias: Dos grupos de datos independientes	157
14.2. Decisión estadística sobre el parámetro diferencia de medias: Dos medidas repetidas	161
14.3. Comparación entre las formas de dos distribuciones de variables aleatorias: Dos grupos de datos independientes	163
14.4. Comparación entre las formas de dos distribuciones de variables aleatorias: Dos medidas repetidas	166
14.5. Ejercicio propuesto	167
CAPÍTULO XV	
15. Minitab versión 14	169
15.1. Introducción	169
15.2. Inicio de la sesión y entorno de trabajo de Minitab	170
15.2.1. Barra de menú principal y herramientas	171
15.2.2. Ventana de sesión	171
15.2.3. Ventana de datos	172
15.2.4. Ventana de gestión de proyectos	172
15.2.5. Ventana de gráficos	173
15.3. Barra de menús principal	174
15.3.1. Instrucciones de manipulación de ficheros	174
15.3.2. Instrucciones de manipulación de datos	175
15.3.3. Instrucciones de cálculo y generación de datos	176
15.3.4. Herramientas estadísticas	177
15.3.5. Herramientas gráficas	178
15.3.6. Uso de la ayuda en Minitab	181
15.4. Trabajo con el proyecto FOBIA	182

CAPÍTULO XVI	
16. Estadística descriptiva univariante	187
16.1. Análisis descriptivo de variables cuantitativas	187
16.2. Análisis descriptivo de variables cualitativas	192
16.3. Ejercicio propuesto	196
CAPÍTULO XVII	
17. Análisis estadístico descriptivo bivariante	197
17.1. Descripción bivariante con datos cuantitativos	197
17.2. Descripción bivariante con datos cualitativos	201
17.3. Descripción bivariante para variables con diferente escala de medida	204
17.4. Ejercicio propuesto	208
CAPÍTULO XVIII	
18. Pruebas de conformidad en distribución	209
18.1. Exploración del grado de ajuste de la distribución empírica a la distribución normal	209
18.2. Pruebas de conformidad para contrastar normalidad	213
18.3. Procedimientos de transformación	217
18.4. Ejercicio propuesto	220
CAPÍTULO XIX	
19. Modelo de regresión lineal simple	221
19.1. Selección de un subgrupo de datos de la base de datos Fobia.MPJ	221
19.2. Análisis descriptivo	224
19.3. Determinación del modelo de regresión simple	228
19.4. Utilidades del modelo de regresión lineal simple: Predicción	232
19.5. Ejercicio propuesto	234
CAPÍTULO XX	
20. Microsoft Excel 2002	235
20.1. Introducción	235
20.2. Entrada al programa y descripción del entorno de trabajo	235
20.3. Funciones estadísticas	237
20.4. Herramientas de análisis de datos	241
20.5. Elaboración de tablas y gráficos dinámicos	242
20.6. Análisis de subpoblaciones: utilización de las opciones Filtro y Subtotales	244
20.7. Guardar la información	246
CAPÍTULO XXI	
21. Análisis estadístico descriptivo para variables cuantitativas	247
21.1. Análisis descriptivo de la variable PULSOINI en el grupo de hombres	247
21.2. Análisis descriptivo de la variable PULSOINI en el grupo de mujeres	251
21.3. Ejercicio propuesto	252
CAPÍTULO XXII	
22. Análisis estadístico descriptivo mediante tablas y gráficos dinámicos	253
22.1. Análisis descriptivo de las variables SUBTIPO, SEXO y EDAD mediante gráficos dinámicos	253
22.2. Análisis descriptivo de las variables SUBTIPO, SEXO y EDAD mediante tablas dinámicas	259
22.3. Ejercicio propuesto	261
CAPÍTULO XXIII	
23. Prueba de relación entre una variable cualitativa dicotómica y una variable cuantitativa	263
23.1. Análisis de la variable ELECTINI en los grupos tratamiento y control	263

23.1.1. Prueba de homogeneidad de variancias para la variable ELECTINI en las poblaciones tratamiento y control	263
23.1.2. Prueba de comparación de medias para la variable ELECTINI en las poblaciones tratamiento y control	266
23.2. Análisis de la variable ELECTFIN en los grupos tratamiento y control	267
23.2.1. Prueba de homogeneidad de variancias para la variable ELECTFIN en las poblaciones tratamiento y control	267
23.2.2. Prueba de comparación de medias para la variable ELECTFIN en las poblaciones tratamiento y control	268
23.3. Ejercicio propuesto	269
Bibliografía	271

PRESENTACIÓN DEL EJEMPLO

La totalidad de los análisis estadísticos que se realizan en el presente libro, ya sean los ejercicios resueltos mediante *SPSS*, *StatGraphics*, *Minitab* o *Excel*, se fundamentan en una hipotética investigación desarrollada en un ámbito profesional de la Psicología. La investigación ha sido ideada con la finalidad de que se puedan plantear distintas preguntas, ya sea en los ejercicios resueltos como en los propuestos, propias de un curso de introducción a la Estadística. Los valores de las variables que se hallan en la matriz de datos fueron simulados para facilitar que se pudieran proponer distintas cuestiones referidas tanto a problemas estadísticos como metodológicos.

Como punto de partida, situémonos en una investigación cuyo principal objetivo es someter a prueba la eficacia de un tratamiento psicológico, siendo la finalidad de éste disminuir los efectos perjudiciales de los pensamientos no adaptados y el exceso de activación fisiológica. En concreto, la investigación considera aquellas personas que padecen algún tipo de fobia. Sobre esta problemática, se dispone de evidencia empírica que sustenta la opinión según la cual los pacientes fóbicos, y sin entrar en distinciones entre distintas tipologías, presentan una conjunción de pensamientos no adaptados que les impide afrontar la situación temida y, además, una elevada y excesiva activación fisiológica. Esta última, al exponer a las personas fóbicas ante la situación que les desencadena la respuesta de temor, actúa como un factor de retroalimentación de los pensamientos no adaptados y provocan conductas que revelan una evidente ansiedad.

La investigación que tomaremos como referencia a lo largo del presente libro se realizó en distintos centros. Colaboraron varias unidades terapéuticas, todas ellas en la misma provincia, especializadas en el tratamiento de la conducta fóbica, estando constituida la muestra definitiva por 160 personas, cada una de las cuales padecía uno u otro tipo de fobia. La totalidad de los individuos pertenecientes a la muestra fueron seleccionados aleatoriamente entre aquellos que contactaban con los centros colaboradores. El proceso de selección de las personas que formarían parte de la muestra se realizó durante el primer semestre del año 2005. Como uno de los criterios para incluir a una persona en la muestra, además de su aceptación mediante la firma de un documento en el cual manifestaba su acuerdo para participar en la investigación, se exigía que todavía no se le hubiese iniciado intervención psicológica o farmacológica alguna y sólo se dispusiera del diagnóstico de algún tipo de fobia. También se estableció que, para pertenecer a la muestra, la problemática fóbica implicara un considerable inconveniente para desenvolverse en la vida diaria del paciente. Por ejemplo, no sería adecuada para el estudio una persona que tuviera temor a viajar en avión pero que rara vez se hubiera hallado ante la necesidad de desplazarse en ese medio de transporte. En el caso contrario, si una persona padeciera temor a hablar en público y su trabajo requiriese de esa actividad con elevada frecuencia, sería un individuo idóneo para incluirlo en la muestra. Con estos requerimientos para incorporar a una persona a la muestra se pretendía evitar posibles factores de confusión, como pueden ser la interacción entre tratamientos psicológicos o farmacológicos anteriores y los diversos grados de severidad de la problemática fóbica, que incidieran en los resultados de la investigación. En este mismo sentido, los profesionales de cada centro que colaboraban en la investigación fueron entrenados en todos los aspectos relevantes y relacionados con ésta, ya fuera la administración de pruebas psicológicas, la aplicación del tratamiento psicológico y la recogida de los datos, procurándose así que todos los profesionales siguiesen patrones idénticos y evitar, de esta manera, sesgos en el estudio.

En una fase inicial de la investigación, los colaboradores de cada centro registraron para cada individuo, previa confirmación de las condiciones requeridas para incluir a las personas en la muestra, información referente a su sexo, edad, nivel educativo, subtipo de fobia padecida, ansiedad e intensidad de los pensamientos fóbicos. Si era necesario administraron pruebas psicológicas, como así fue para medir el grado de ansiedad y la intensidad de los pensamientos fóbicos. También se hicieron registros fisiológicos, mediante un ordenador portátil, en esta fase inicial de la investigación. Estas medidas, que se obtuvieron exponiendo a los sujetos de la muestra ante la situación temida, fueron la frecuencia cardiaca y la actividad electrodérmica de la piel, mediciones que requirieron la implantación de electrodos. Una vez se hubo obtenido la información inicial se asignaron al azar los sujetos de la muestra a uno de los dos grupos, control o tratamiento, balanceándose ambos grupos respecto al sexo de sus componentes; o sea, se garantizó que en los dos grupos existiera una misma cantidad de mujeres y hombres. Cabe mencionar que se aseguró que los sujetos experimentales no supieran en qué grupo se hallaban, estrategia que se denomina *simple ciego*, para evitar que los resultados de la investigación estuvieran afectados por efectos derivados del ajuste de los sujetos experimentales a las expectativas del investigador; es decir, que los sujetos

participantes en la investigación mostraran un comportamiento acorde con lo que ellos suponían que esperaría el investigador, pudiéndose confirmar así de forma artificiosa las hipótesis de la investigación.

En el grupo de control se planificaron encuentros con la misma periodicidad que las sesiones terapéuticas en el grupo de tratamiento. En las sesiones rutinarias para los sujetos incluidos en el grupo de control éstos podían expresar preguntas referidas a su problemática y el terapeuta respondía de una forma informal, evitando este último en todo momento cualquier tipo de acción que pudiera considerarse una intervención o tratamiento psicológico. De esta manera podría evaluarse el *efecto placebo*, que en este caso sería la mejora resultante del hecho de que los pacientes se supieran atendidos. En cuanto a los sujetos experimentales asignados al grupo de tratamiento, éstos asistieron durante tres meses a sesiones semanales de una hora de duración, interviniéndoseles mediante técnicas de reestructuración cognitiva —este procedimiento incide en la sustitución de los pensamientos no adaptados y referidos al estímulo que desencadena la respuesta fóbica por otros más adecuados—, técnicas de relajación y control de la activación fisiológica mediante retroalimentación —esta estrategia consiste en que se aprenda a reconocer, primero, y controlar, después, las señales que indican una elevada y excesiva activación ante el estímulo que provoca la respuesta fóbica— y, por último, la exposición directa al estímulo temido. Una vez finalizada la fase experimental, se administraron de nuevo las pruebas o tests psicométricos y se registraron otra vez las variables fisiológicas, que no son otras que la frecuencia cardíaca y la actividad electrodérmica de la piel, en las mismas condiciones que en la medición anterior. Es importante mencionar que todas las medidas, en la fase inicial de la investigación y al finalizar la misma, se obtuvieron individualmente, procurándose evitar en lo posible los efectos derivados de la interacción entre los sujetos experimentales.

A modo de resumen, seguidamente especificaremos y describiremos las distintas variables incluidas en la investigación, haciendo mención a sus posibles valores y, en algunos casos, a las unidades de medida:

1. **SEXO**, que puede tomar los valores *mujer y hombre*;
2. **SUBTIPO**, cuyos posibles valores son *social, específica y agorafobia*;
3. **EDAD**, variable que fue medida en años cumplidos;
4. **EDUCA**, que se corresponde al nivel educativo de los sujetos experimentales y cuyos valores son *bajo, medio y alto*;
5. **DETECCI**, que tiene a *precoz, intermedio y tardío* por valores factibles en el presente estudio;
6. **GRUPO**, que puede tomar el valor *control o tratamiento*, dependiendo del grupo al cual fue asignado el sujeto;
7. **PENSINI**, medición que se corresponde con la puntuación total obtenida en una prueba psicométrica constituida por 90 reactivos, cada uno de los cuales admitía la respuesta *sí o no* según indicara o no un elemento de respuesta fóbica;
8. **PULSOINI**, cuantificada como el número de pulsaciones por minuto —recuérdese que la frecuencia considerada normal oscila entre 60 y 90 pulsaciones por minuto o *ppm*—;
9. **ELECTINI**, mensurada como la resistencia eléctrica que ofrece la piel a la transmisión de una señal eléctrica y cuya unidad de medida fue establecida en *microhmios*;
10. **ANSIEDAD**, medida por medio de una prueba psicológica que cuantifica el nivel o grado de ansiedad rasgo;
11. **PENSFIN**, que fue medida tal como se hizo para **PENSINI**, pero tras la finalizar el periodo de intervención psicológica;
12. **PULSOFIN**, cuya medición fue idéntica a **PULSOINI**, salvo que se realizó después de que concluyera la fase de la investigación en la cual se llevó a cabo el tratamiento psicológico;
13. y **ELECTFIN**, cuyas medidas se obtuvieron igual que las obtenidas para **ELECTINI**, pero al concluir la fase experimental del estudio.

1. SPSS PARA WINDOWS (VERSIÓN 12.0.1)

1.1 Introducción

El programa *SPSS (Statistical Package for Social Science)* constituye una potente aplicación estadística de la que, desde su versión inicial hace más de 35 años, se han desarrollado diferentes versiones. Inicialmente fue planteado para grandes sistemas informáticos, si bien, en la actualidad, se halla operativo en una gran cantidad de entornos (Unix, Vax, MsDos, Mac, Windows, etc.). De hecho, las normas generales de sintaxis y programación son las mismas, lo que facilita que datos y programas sean fácilmente transportables de un entorno a otro; aun así la implementación del programa para cada sistema operativo presenta algunas particularidades que deben ser conocidas y tenidas en cuenta. Así mismo, aunque nos hallemos en el mismo entorno, las diferentes versiones del programa estadístico incorporan, sucesivamente, nuevas capacidades respecto a las anteriores versiones convirtiéndolo en más versátil y potente.

En el presente capítulo presentaremos una breve introducción (sin ánimo de sustituir en ningún momento la función de los manuales del programa) de la manipulación del programa en versiones desarrolladas en entorno de ordenador personal (PC) y para la interfaz gráfica tipo *Windows*. Como consideraciones previas tengamos en cuenta que:

- En primer lugar, la versión utilizada para el siguiente texto es la 12.0.1 en castellano. Esta versión requiere, para ser soportada, disponer al menos de plataforma *Windows 98*, *Windows ME*, *Windows NT 4.0.*, *Windows 2000* o *Windows XP* (versión esta última utilizada en el presente texto) con los correspondientes “*Service Pack*” de actualización. Por tanto, desde otros entornos *Windows* anteriores no será posible trabajar con la presente versión (por ejemplo, para la versión de *Windows 3.1* la versión más actualizada de *SPSS* que puede ejecutarse es la 6.1.3). Otros requisitos del sistema son: un procesador tipo Pentium, un mínimo de 128 MB de memoria RAM, 220 MB libres en la unidad de disco fijo, una lectora de CD-ROM y un adaptador gráfico SVGA con una resolución mínima de 800x600.
- En segundo lugar, debemos también mencionar que el lenguaje básico del programa se desarrolló, inicialmente, para una estructura de comandos. Por ejemplo, la versión 5.0 y anteriores para PC requerían, para la ejecución de los diferentes procedimientos, la escritura del comando adecuado ya fuera *on-line* o bien en un fichero de procesamiento por lotes (fichero que contenía un grupo de comandos para su ejecución secuencial). De hecho, estos comandos se encuentran subyacentes aún en la actualidad en la interfaz gráfica y es posible la ejecución de los procedimientos implementados en el programa mediante ellos, además de los cuadros de diálogo propios del entorno *Windows*. Cabe mencionar que, algunos subcomandos y opciones sólo son accesibles para su ejecución desde la ventana de escritura de comandos.
- Finalmente, mencionemos que, entre otras mejoras, a partir de la versión 7.5 y siguientes se ha realizado una mayor integración del programa en el entorno gráfico, incorporándose el tratamiento de muchos elementos como objetos independientes, gráficos y tablas de tipo interactivo, un navegador para el fichero de resultados, exportación de ficheros de resultados en formato HTML, etc. Así mismo, se han incorporado módulos de nuevos procedimientos, tanto propiamente estadísticos como de elaboración y recogida de información de encuestas, implementándose la posibilidad de utilizar un servidor remoto para llevar a cabo análisis complejos con un gran volumen de datos sin bloquear el ordenador personal (lo que se ha dado en llamar análisis distribuido).

1.2. Ejecución del programa y elementos principales.

Para la ejecución del programa¹, localizaremos, en la barra de tareas situada al pie del escritorio, el botón de **Inicio**, que nos permite acceder al menú desplegable. En este menú se halla la opción de **Todos los Programas**, dentro de ésta, la opción **SPSS para Windows** en la cual se puede hallar la opción **SPSS 12.0 para Windows** (o la correspondiente a la versión del programa que vayamos a ejecutar), tal y como se muestra en la Figura 1.1.

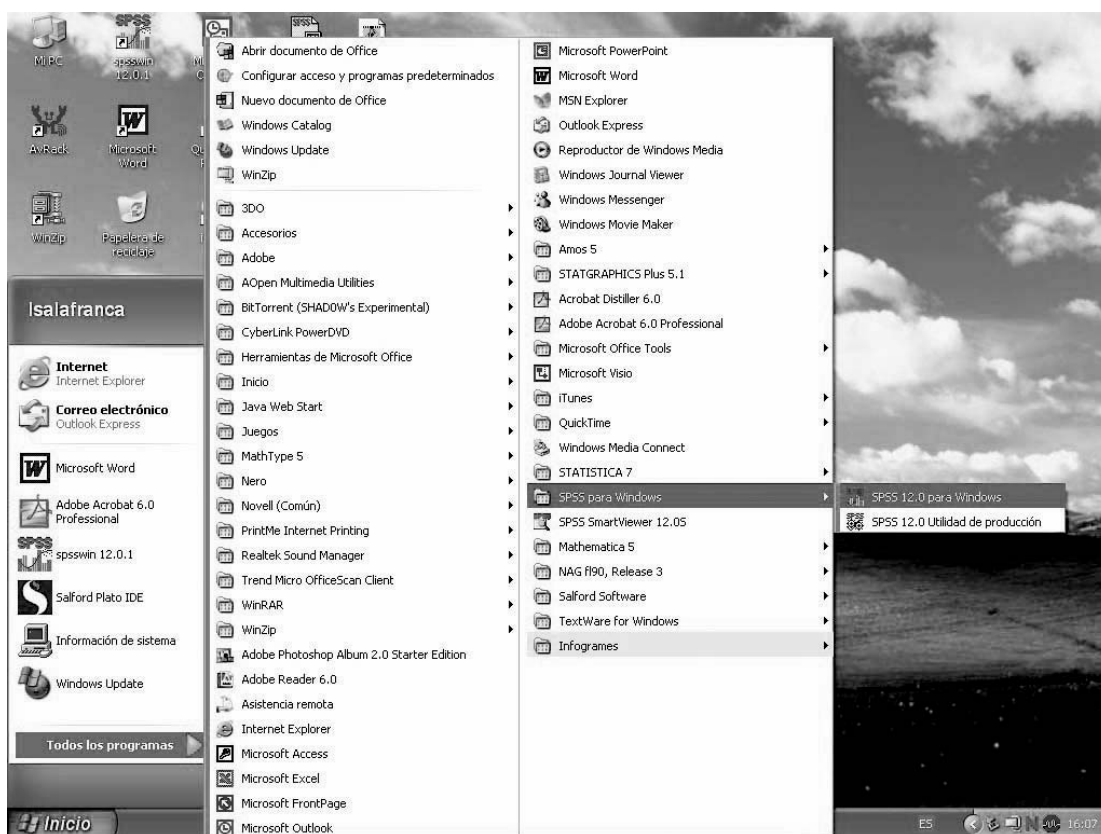


Figura 1.1

Una vez activada la opción indicada en el párrafo anterior mediante una pulsación del botón izquierdo del ratón, esta acción nos permitirá acceder al programa. Al efectuar esta operación, se nos abrirá en el escritorio de forma automática al inicio de cada sesión, la ventana del **Editor de datos SPSS**, que podemos observar en la Figura 1.2. En dicha ventana hallaremos:

1. **Barra de título:** donde se identifica cuál es la ventana abierta. Al inicio de la sesión se identifica como *Sin título*.
2. **Barra principal de menús:** Se sitúa en la parte superior de la ventana del editor de datos y contiene los menús desplegables que nos permiten acceder a la mayoría de funcionalidades del SPSS:

¹ Alternativamente a lo aquí indicado también puede ejecutarse mediante un acceso directo al programa, caso de que se disponga de él, ubicado en el escritorio.

- a. **Archivo:** Permite acceder y crear diferentes tipos de ficheros, así como, entre otras operaciones, el control de la impresión.
 - b. **Edición:** Opciones habituales del entorno *Windows* para copiar, cortar, eliminar, etc., así como las opciones de configuración que por defecto tiene implementadas el programa.
 - c. **Ver:** Permite controlar algunas opciones de configuración de la ventana abierta.
 - d. **Datos:** Permite crear el diccionario de etiquetas de un archivo de datos, además de manipular ficheros de datos y los casos de éstos.
 - e. **Transformar:** Contiene las opciones encaminadas a la manipulación de variables.
 - f. **Analizar:** Accede a las técnicas estadísticas de análisis numérico. La tipología y número de técnicas disponibles dependerá de los módulos del programa para los que se disponga de licencia de ejecución. Para obtener una descripción de las técnicas estadísticas se recomienda abrir el menú de **Ayuda** disponible dentro de cada técnica y la consulta de cualquier manual de estadística.
 - g. **Gráficos:** Menú de acceso a las técnicas gráficas. Para estas técnicas es vigente el comentario realizado en el apartado anterior.
 - h. **Utilidades:** Contiene los controles de algunas utilidades del programa, como por ejemplo la introducción de comentarios en el fichero de datos o la definición y uso de conjuntos de variables.
 - i. **Ventana:** Permite controlar atributos de las ventanas del programa.
 - j. **Ayuda(?):** Se identifica con el símbolo del interrogante y da acceso a los ficheros de ayuda.
3. **Barra de Herramientas:** Situada inmediatamente debajo de la anterior barra, contiene botones con iconos que permiten el acceso directo a algunas de las operaciones más habitualmente utilizadas de la barra principal de menús, como el acceso a ficheros, impresión de listados, etc. Así mismo, si situamos el puntero del ratón sobre cualquiera de los iconos de la barra de herramientas, nos indicará la funcionalidad de dicho icono.
 4. **Línea de Edición de Datos:** En la que es posible introducir y modificar la información de la matriz de datos. En general, identifica la información de la celda activa (aquella en la que nos hallamos actualmente ubicados) de la matriz de datos. Esta opción sólo está disponible en la **Vista de Datos**.
 5. **Matriz de Datos:** Donde nos aparecerá la información estructurada en forma de matriz *sujetos x variables* (si se trata de la **Vista de Datos**) o bien *variables x características* (si se trata de **Vista de Variables**).
 6. **Pestaña de Visualización de Datos/Variables:** El acceso de intercambio de ambas matrices se encuentra en la parte inferior izquierda. Aparece en forma de dos pestañas (**Vista de datos**, véase Figura 1.2, y *Vista de variables* véase Figura 1.3), en los que, mediante una pulsación del botón izquierdo del ratón, podemos visualizar los datos (*variables x sujeto*) o bien las variables (cada una de las variables con sus características incluidas en el diccionario de datos).
 7. **Barra de Estado:** Situada en la parte inferior izquierda nos indica el proceso u operación que está realizando el programa. En el caso de no estar realizando ninguna operación, se nos muestra el mensaje **SPSS El procesador está preparado**.

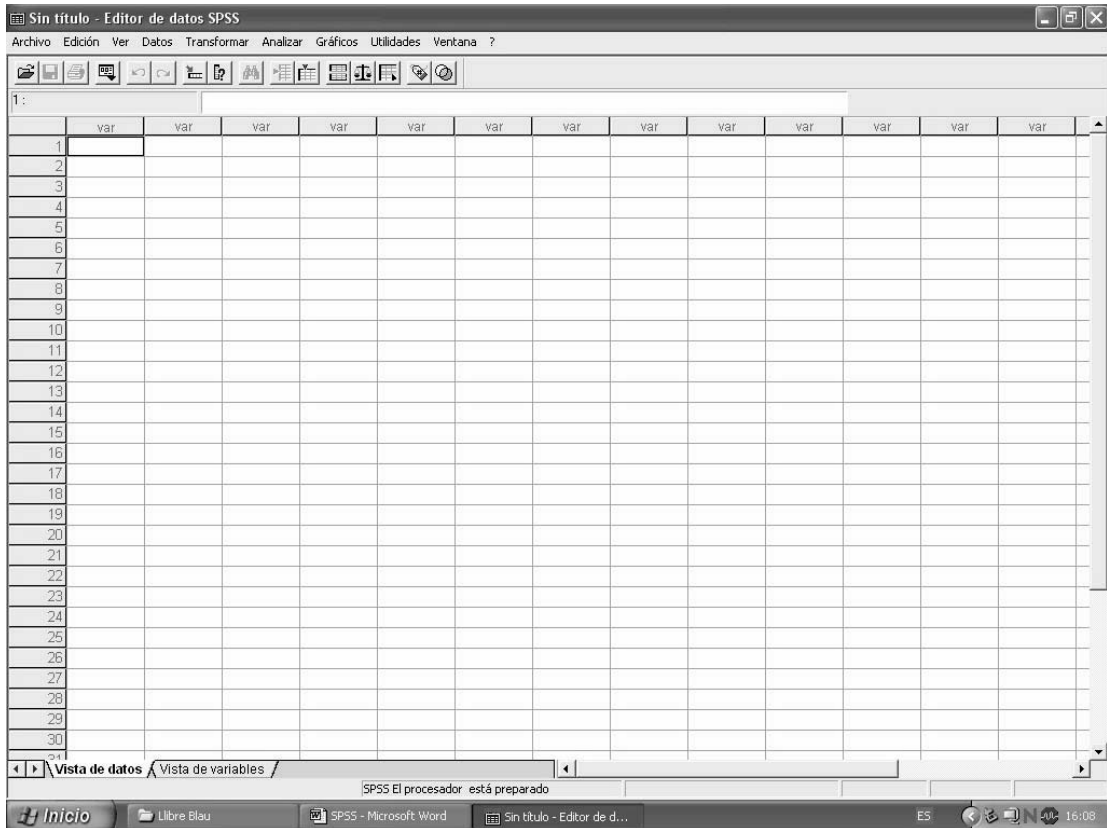


Figura 1.2

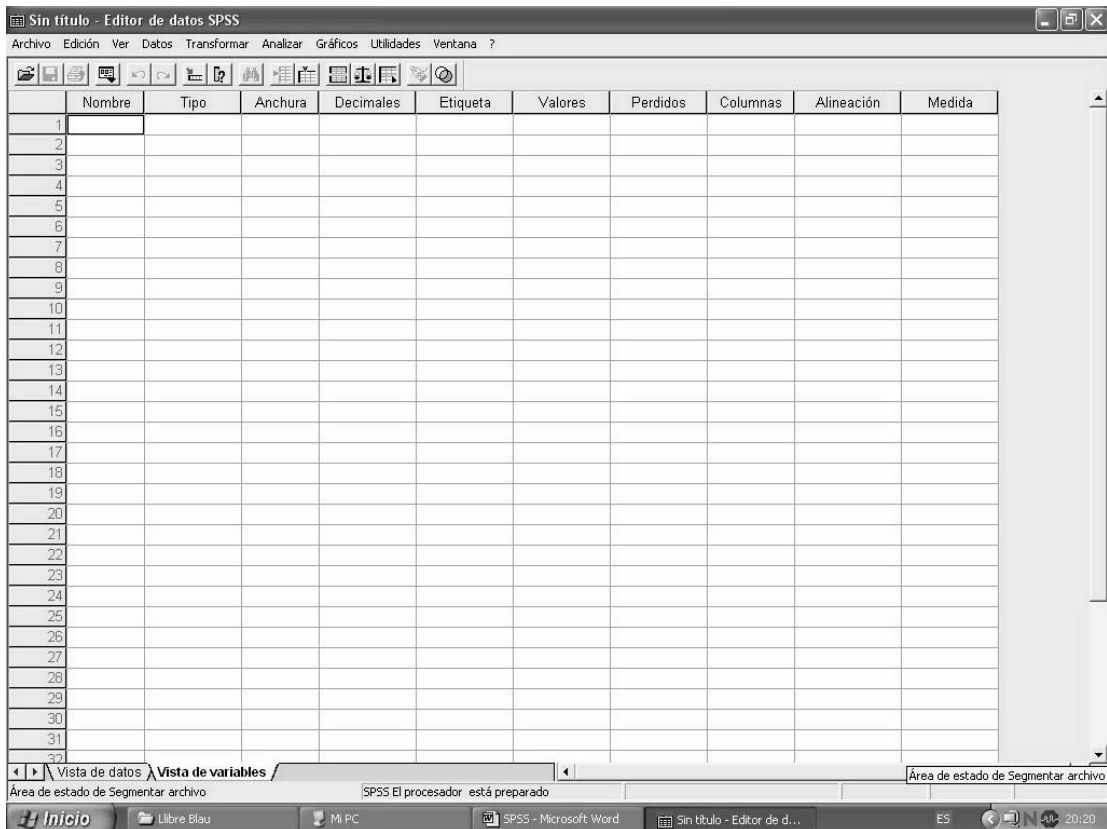


Figura 1.3

Durante la ejecución del programa, dependiendo de las operaciones que realicemos, podremos tener activadas otras ventanas, que configuran objetos independientes en el escritorio de Windows. Entre otras, tenemos los siguientes tipos de ventanas adicionales:

- **Ventana del visor de Resultados:** Se inicia automáticamente al ejecutar cualquier procedimiento de análisis o al abrir un fichero de resultados previamente existente. Se identifica, en caso de que no se corresponda con un fichero generado anteriormente o en caso de que contenga las primeras ejecuciones realizadas mediante el paquete estadístico, como **Resultados 1- Visor SPSS**. En esta ventana se ubicarán los resultados de los análisis realizados. Su apariencia y estructura, en general, es similar a la ventana del editor de datos, substituyendo la zona de la matriz de datos por una pantalla dividida en dos compartimentos: a la izquierda se sitúa un navegador y a la derecha la zona donde se ubicarían propiamente los resultados. Mediante el navegador, que incluye un guión de los elementos que componen el listado de resultados, podemos desplazarnos de forma rápida a cualquier punto del listado. Es posible mantener abiertas varias ventanas de resultados, pero solamente una de ellas será la ventana designada para contener los resultados de un análisis. Dentro del menú desplegable de **Utilidades** hallaremos la opción **Designar ventana** para proceder a determinar la ventana destino de los análisis. No debe confundirse el término ventana designada con ventana activa, esta última es la que se tiene seleccionada en un momento dado. Debemos tener en cuenta que, como consecuencia de los cambios introducidos en el formato de los archivos de resultados *SPSS*, puede que los documentos creados con cualquiera de las versiones *SPSS* 12.0.X no sean legibles en las versiones anteriores.
- **Ventana del editor Sintaxis:** Podemos abrir una ventana de estas características siempre que no deseemos trabajar con los iconos y menús desplegables sino con un fichero de comandos ejecutado secuencialmente (en procesamiento por lotes o modo Batch). También es posible trabajar con varias ventanas de sintaxis, pero igual que ocurría con las ventanas de resultados, sólo una de ellas puede ser la designada. En muchos procedimientos observaremos el botón **Pegar** que nos permite insertar en esta ventana el conjunto de comandos especificados mediante el cuadro de diálogo. Alternativamente, si conocemos la sintaxis de los comandos que deseamos ejecutar, podemos escribirlos directamente dentro de la ventana. En la barra de menús desplegables nos aparecerá la opción **Ejecutar**, que nos permitirá la ejecución de todos o parte de los comandos incluidos en la ventana de sintaxis.
- **Ventana de Procesos:** Mediante esta ventana se activa un entorno de programación que utiliza el lenguaje *Sax BASIC* e incluye un editor de cuadros de diálogo, un visor de objetos, funciones de depuración y ayuda sensible al contexto. A través de esta ventana podemos automatizar o personalizar cualquier tipo de proceso.
- **Ventana del editor de gráficos, de resultados de texto y de tablas pivote:** Ventanas para la manipulación de las características y contenidos de los gráficos, los resultados de texto o las tablas pivote generados al ejecutar algún análisis mediante el programa.

1.3. Lectura de un fichero de datos en formato SPSS (*.SAV).

Podemos analizar mediante el programa cualquier matriz de datos que se haya cargado previamente en la ventana del editor de datos. Además, desde la versión número 10 del *SPSS* es posible realizar múltiples ejecuciones del programa, cada una con datos diferentes (por supuesto ubicadas en ventanas de datos distintas) y durante la sesión ir intercambiando la ventana de datos activa sobre la cual se realiza la ejecución. A continuación detallaremos el procedimiento para cargar una matriz de datos previamente construida, en la ventana del editor de datos. Si bien el programa *SPSS* puede importar y exportar ficheros de datos desde/a otros entornos (*ASCII*, *Excel*, *Dbase...*), por los intereses del presente texto expondremos solamente la manipulación de ficheros de datos propios, es decir, un fichero previamente generado desde la ventana de edición de datos del programa. Habitualmente, los ficheros de datos en formato *SPSS* se identifican con la extensión *sav*. Estos ficheros no sólo contienen los datos de las variables recogidas en el conjunto de los individuos, sino que, también, integran el denominado **diccionario de los datos**; es decir, información respecto al tipo de variable, etiquetas de los valores, existencia de valores *missings*,

etc. La información contenida en el diccionario de datos es relevante para la manipulación de la matriz por parte de los comandos del programa.

Referiremos a continuación el proceso para activar un fichero en este formato. En primer lugar, desde la barra principal de menús desplegados se activará la opción **Archivo** mediante una pulsación del botón izquierdo del ratón o, alternativamente, mediante la *hot key* correspondiente (pulsación simultánea de las teclas **Alt** y **c**)². Una vez tenemos el menú desplegado, seleccionaremos la opción **Abrir** (pulsación de la tecla **a**, puesto que ya no es necesario volver a pulsar **Alt**), que al ser activado nos permite seleccionar el tipo de fichero al que deseamos acceder. En este caso la opción deseada sería la de **Datos** (en la Figura 1.4 se ilustra el menú desplegable activado). Este último paso, al seleccionarlo, nos permitirá acceder al cuadro de diálogo donde podremos especificar la unidad, directorio o carpeta (mediante **Buscar en:**) y fichero correspondientes (véase Figura 1.5) que deseamos activar en la ventana de edición de datos. En nuestro caso localizaríamos el fichero *fobia.sav* en su ubicación y este nombre sin la extensión lo escribiríamos como **Nombre del archivo**. Una vez realizada la selección, pulsaremos con el ratón sobre el botón **Aceptar**, con lo que el programa efectuará la operación de carga del fichero. A partir de este punto podremos realizar cualquier operación con los datos que tenemos activos.

Cualquier tipo de cambio que se efectúe sobre la matriz de datos, ya sea sobre el diccionario o sobre los datos existentes, ya sea añadiendo nuevas variables, solamente tendrá repercusión sobre los datos contenidos en la ventana del editor de datos. Si queremos que estos cambios queden recogidos en el fichero inicialmente cargado o en un nuevo fichero externo, deberemos proceder a la grabación del contenido de la ventana. Para realizar tal operación nos remitimos al apartado 1.6 donde se expone la grabación del contenido de las diferentes ventanas en ficheros externos.

Hemos mencionado la existencia de un diccionario de datos incorporado al fichero de datos y la relevancia de su información para la manipulación de la matriz de datos. Para conocer su contenido únicamente debemos pulsar con el botón izquierdo del ratón con el puntero sobre la pestaña **Vista de Variables** de la parte inferior izquierda (encima de la barra de estado) o, si se prefiere, se puede acceder mediante una doble pulsación del botón izquierdo del ratón sobre el nombre de una variable cualquiera situada en el borde superior de la matriz de datos. Así, observaremos otra matriz donde se hallan representadas cada una de las variables con las características que se pueden identificar de ellas. Desde dicha matriz podremos especificar o modificar las características del nombre, tipo, anchura, decimales, etiquetas de variable y de valores, valores perdidos, formato de columna y escala de medida de dichas variables (véase Figura 1.6). Para seleccionar y modificar alguna de las características de las variables simplemente debemos ubicarnos en la casilla correspondiente con el puntero del ratón y pulsar el botón izquierdo. Una vez seleccionado, podremos o bien directamente escribir sobre la casilla, o alternativamente, en la parte derecha de la celda aparecerá una zona sensible en la que podremos actuar al posicionar encima el puntero del ratón y pulsar el botón izquierdo. Mencionemos algunas de las particularidades de las características que podemos definir:

- a) En lo que respecta al **Nombre**, en esta versión del programa se ha ampliado el máximo de caracteres posible de 8, que era el máximo en versiones anteriores, hasta 64. Además, el nombre no puede empezar por un carácter numérico y deben evitarse los espacios en blanco y los caracteres reservados (por ejemplo, los asteriscos) o aquellos nombres que indiquen una operación o mandato del programa (por ejemplo, **AND**). El nombre que le indiquemos será reconocido por el programa, de forma interna, para esta variable y es el que utilizaremos para realizar cualquier tipo de manipulación u operación con la misma.

² Para identificar las *hot key* de las distintas opciones únicamente debemos pulsar la tecla de **Alt**.

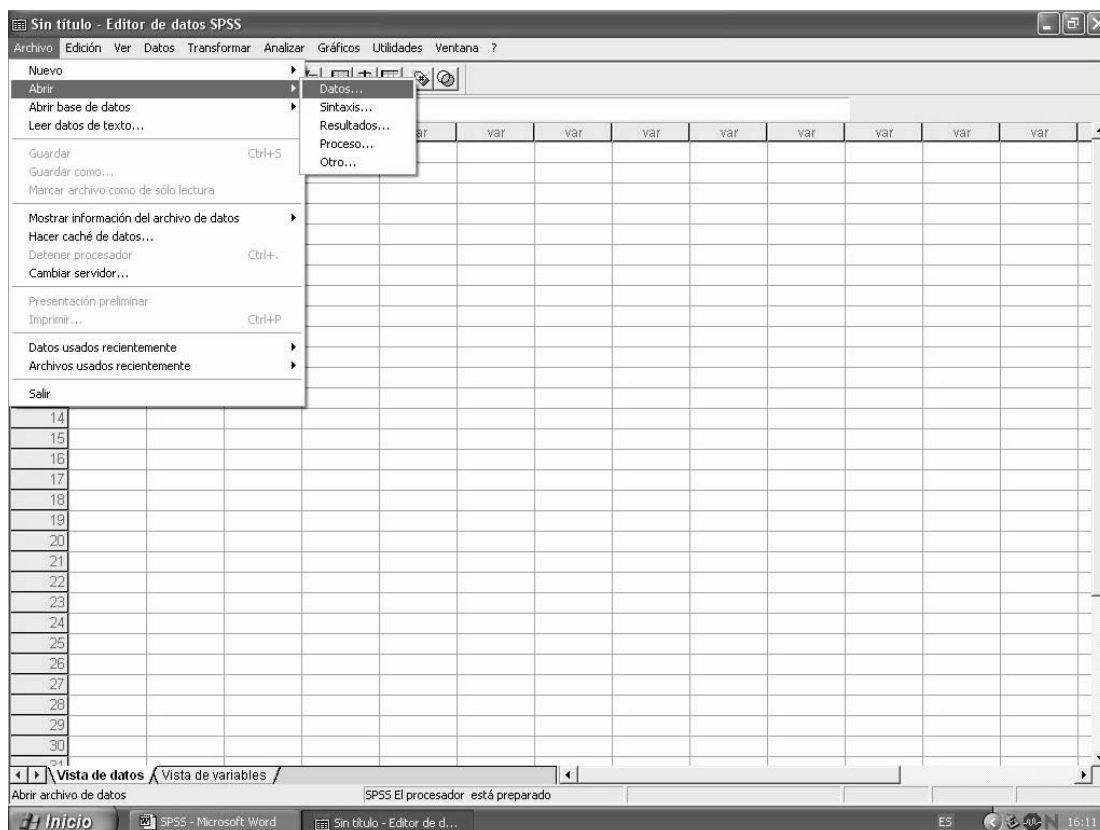


Figura 1.4

- b) Respecto al **Tipo**, nos recoge el código o formato con el cual se han introducido los datos de entre las opciones que se señalan en el cuadro de diálogo correspondiente. Mencionemos las más habituales: **numérico**, con indicación del ancho máximo y el número de decimales; **cadena** (valores alfanuméricos), con indicación del número máximo de caracteres³. Es recomendable utilizar, siempre que sea posible, códigos numéricos para la introducción de valores en la matriz de datos, puesto que algunas técnicas estadísticas se han implementado en el SPSS de tal forma que solamente reconocen las variables para su análisis si éstas contienen valores numéricos.
- c) La columna correspondiente a **Etiqueta** nos permite dar un identificador de hasta 256 caracteres a las variables. En este caso las etiquetas no presentan las restricciones que anteriormente hemos mencionado para el nombre de las variables.
- d) La columna correspondiente a **Valores** nos permite proporcionar etiquetas identificativas a los números utilizados para los valores de variables no cuantitativas.
- e) La casilla **Perdidos** permite definir valores omitidos o *missing* para la variable por parte del usuario. Estos *missings* de usuario se añaden a los que el propio sistema reconoce como tales sin ningún tipo de indicación.
- f) Las casillas de **Columnas** y **Alineación** establecen el ancho de columna de cada variable, así como la alineación de los valores dentro de la celda.
- g) Finalmente, debemos indicar el tipo de **Medida** de la variable (nominal, ordinal, o escala, ya sea ésta de intervalo o razón)

³ La anchura, y en su caso el número de decimales, quedan especificados en las dos columnas que siguen a la de **Tipo**.

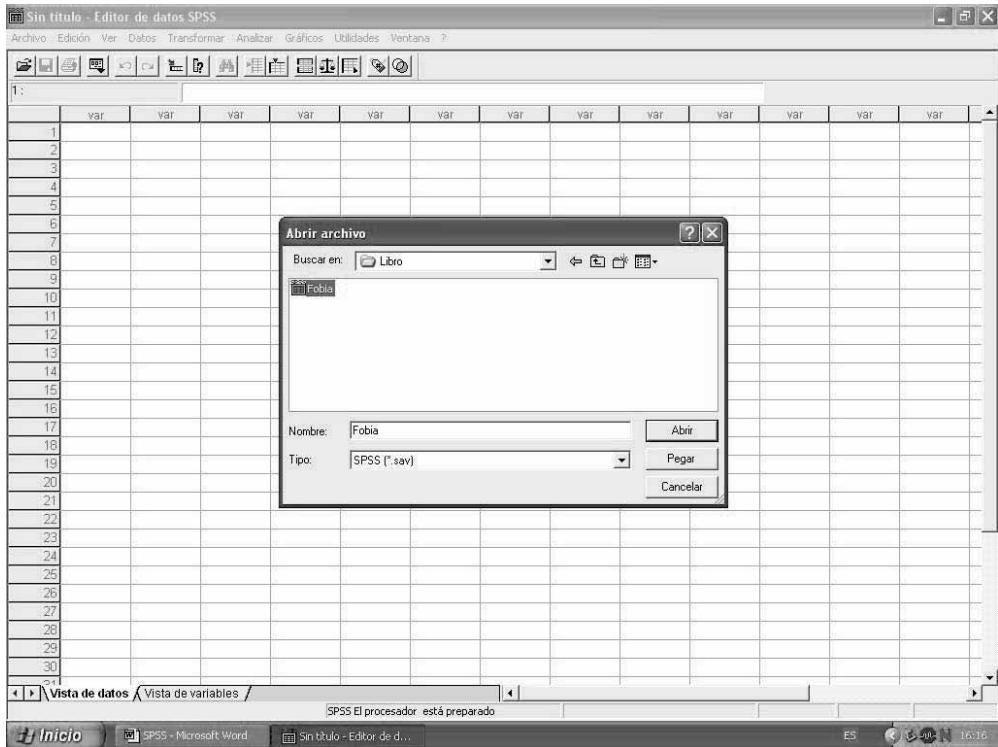


Figura 1.5

Si hemos terminado las modificaciones a realizar en los elementos anteriores, para salir de la matriz correspondiente, solamente es necesario volver a pulsar con el botón izquierdo en **Vista de Datos** y las modificaciones efectuadas se llevarán a cabo. Se ha de tener en cuenta que cualquier modificación que se efectúe en la matriz de variables será llevada a cabo por el programa sin pedir ningún tipo de confirmación, por lo que se recomienda tener mucha precaución en el uso de este recurso, ya que es fácil realizar modificaciones no deseadas sin percatarse de ello.

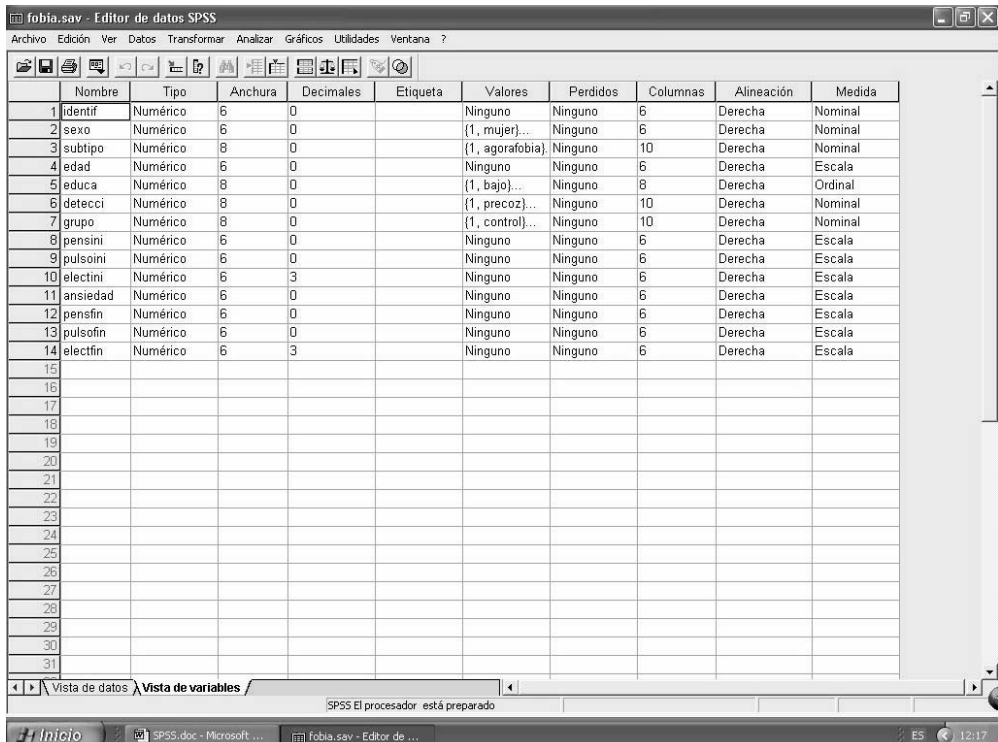


Figura 1.6

Al visualizar la pestaña de **Vista de datos**, los datos de las variables cuyos valores han sido etiquetados pueden ser visualizados en el editor de dos formas. Lo usual, por defecto, es que se nos muestren tal como han sido entradas originalmente (véase Figura 1.7). Pero también es posible visualizar las etiquetas para cada valor. Para ello en la opción **Ver** de la barra de principal de menús desplegables deberemos activar la opción **Etiquetas de valor**, tal como se muestra en la Figura 1.8. Una vez activada, todos aquellos valores de las variables que hayan sido etiquetados se mostrarán en la **Vista de datos** mediante la etiqueta correspondiente que los identifique. Visualizar el fichero de datos de esta forma acostumbra a ser más ilustrativo del contenido, tal y como se observa en la Figura 1.9. El cambio de visualización no afecta a la forma en que están archivados internamente los datos, y puede volverse a la visualización inicial simplemente desmarcando la opción **Etiquetas de valor**.

identifi	sexo	subtipo	edad	educa	detecci	grupo	pensini	pulsoini	electini	ansiedad	pensfn	pulsofn	electfn	var
1	1	1	2	28	2	1	1	77	129	268	36	78	145	353
2	2	1	2	25	1	3	1	77	137	375	21	74	140	263
3	3	1	1	24	2	2	1	67	127	207	17	76	123	278
4	4	1	1	29	3	2	1	76	130	261	19	79	134	246
5	5	1	2	21	3	1	1	80	142	394	38	81	126	267
6	6	1	2	32	3	1	1	74	141	349	40	73	131	407
7	7	1	2	25	1	3	1	81	143	265	42	80	140	225
8	8	1	1	25	2	2	1	79	133	351	25	72	142	261
9	9	1	3	26	2	3	1	69	134	254	25	70	132	411
10	10	1	2	27	1	3	1	79	138	295	22	65	137	354
11	11	1	1	28	1	3	1	76	132	323	21	77	142	262
12	12	1	2	23	1	3	1	75	141	312	26	80	146	258
13	13	1	2	19	3	1	1	74	132	362	21	79	136	256
14	14	1	2	15	3	1	1	81	132	243	31	78	142	498
15	15	1	2	23	3	1	1	77	135	314	33	82	139	217
16	16	1	3	28	2	2	1	73	147	445	35	76	134	330
17	17	1	3	24	2	2	1	75	134	292	35	78	134	346
18	18	1	1	26	2	2	1	74	138	261	26	81	139	394
19	19	1	3	20	2	2	1	72	135	220	23	75	132	344
20	20	1	1	14	2	2	1	75	132	265	26	81	131	341
21	21	1	2	32	1	2	2	74	138	307	19	47	86	124
22	22	1	1	22	1	2	2	76	144	265	35	33	68	009
23	23	1	3	17	1	3	2	81	129	253	34	43	93	171
24	24	1	2	30	2	1	2	83	136	290	37	39	81	009
25	25	1	2	23	2	1	2	78	137	346	39	41	87	110
26	26	1	3	23	2	2	2	77	131	230	21	33	95	072
27	27	1	2	31	2	1	2	72	133	273	25	35	92	028
28	28	1	1	30	2	2	2	80	135	269	36	35	93	071
29	29	1	2	23	1	2	2	80	144	331	31	30	77	028
30	30	1	2	21	1	3	2	72	137	301	24	43	101	038

Figura 1.7

identifi	sexo	subtipo	edad	educa	detecci	grupo	pensini	pulsoini	electini	ansiedad	pensfn	pulsofn	electfn	var
1	1	1	2	28	2	1	1	77	129	268	36	78	145	353
2	2	1	2	25	1	3	1	77	137	375	21	74	140	263
3	3	1	1	24	2	2	1	67	127	207	17	76	123	278
4	4	1	1	29	3	2	1	76	130	261	19	79	134	246
5	5	1	2	21	3	1	1	80	142	394	38	81	126	267
6	6	1	2	32	3	1	1	74	141	349	40	73	131	407
7	7	1	2	25	1	3	1	81	143	265	42	80	140	225
8	8	1	1	25	2	2	1	79	133	351	25	72	142	261
9	9	1	3	26	2	3	1	69	134	254	25	70	132	411
10	10	1	2	27	1	3	1	79	138	295	22	65	137	354
11	11	1	1	28	1	3	1	76	132	323	21	77	142	262
12	12	1	2	23	1	3	1	75	141	312	26	80	146	258
13	13	1	2	19	3	1	1	74	132	362	21	79	136	256
14	14	1	2	15	3	1	1	81	132	243	31	78	142	498
15	15	1	2	23	3	1	1	77	135	314	33	82	139	217
16	16	1	3	28	2	2	1	73	147	445	35	76	134	330
17	17	1	3	24	2	2	1	75	134	292	35	78	134	346
18	18	1	1	26	2	2	1	74	138	261	26	81	139	394
19	19	1	3	20	2	2	1	72	135	220	23	75	132	344
20	20	1	1	14	2	2	1	75	132	265	26	81	131	341
21	21	1	2	32	1	2	2	74	138	307	19	47	86	124
22	22	1	1	22	1	2	2	76	144	265	35	33	68	009
23	23	1	3	17	1	3	2	81	129	253	34	43	93	171
24	24	1	2	30	2	1	2	83	136	290	37	39	81	009
25	25	1	2	23	2	1	2	78	137	346	39	41	87	110
26	26	1	3	23	2	2	2	77	131	230	21	33	95	072
27	27	1	2	31	2	1	2	72	133	273	25	35	92	028
28	28	1	1	30	2	2	2	80	135	269	36	35	93	071
29	29	1	2	23	1	2	2	80	144	331	31	30	77	028
30	30	1	2	21	1	3	2	72	137	301	24	43	101	038

Figura 1.8

identif	sexo	subtipo	edad	educa	detecci	grupo	pensini	pulsoini	electini	ansiedad	pensfin	pulsofin	electfin	var
1	1	mujer	social	28	medio	precoz	control	77	129	,268	36	78	145	,353
2	2	mujer	social	25	bajo	tardío	control	77	137	,375	21	74	140	,283
3	3	mujer	agorafobia	24	medio	intermedio	control	67	127	,207	17	76	123	,278
4	4	mujer	agorafobia	29	alto	intermedio	control	76	130	,281	19	79	134	,246
5	5	mujer	social	21	alto	precoz	control	80	142	,394	38	81	126	,267
6	6	mujer	social	32	alto	precoz	control	74	141	,349	40	73	131	,407
7	7	mujer	social	25	bajo	tardío	control	81	143	,265	42	80	140	,225
8	8	mujer	agorafobia	25	medio	intermedio	control	79	133	,351	25	72	142	,261
9	9	mujer	específica	26	medio	tardío	control	69	134	,254	25	70	132	,411
10	10	mujer	social	27	bajo	tardío	control	79	138	,295	22	85	137	,354
11	11	mujer	agorafobia	28	bajo	tardío	control	76	132	,323	21	77	142	,282
12	12	mujer	social	23	bajo	tardío	control	75	141	,312	26	80	146	,258
13	13	mujer	social	19	alto	precoz	control	74	132	,362	21	79	136	,256
14	14	mujer	social	15	alto	precoz	control	81	132	,243	31	78	142	,498
15	15	mujer	social	23	alto	precoz	control	77	135	,314	33	82	139	,217
16	16	mujer	específica	28	medio	intermedio	control	73	147	,445	35	76	134	,330
17	17	mujer	específica	24	medio	intermedio	control	75	134	,292	35	78	134	,346
18	18	mujer	agorafobia	26	medio	intermedio	control	74	138	,281	26	81	139	,394
19	19	mujer	específica	20	medio	intermedio	control	72	135	,220	23	75	132	,344
20	20	mujer	agorafobia	14	medio	intermedio	control	75	132	,285	26	81	131	,341
21	21	mujer	social	32	bajo	intermedio	tratamiento	74	138	,307	19	47	86	,124
22	22	mujer	agorafobia	22	bajo	intermedio	tratamiento	76	144	,285	35	33	88	,009
23	23	mujer	específica	17	bajo	tardío	tratamiento	81	129	,253	34	43	93	,171
24	24	mujer	social	30	medio	precoz	tratamiento	83	136	,290	37	39	81	,009
25	25	mujer	social	23	medio	precoz	tratamiento	78	137	,346	39	41	87	,110
26	26	mujer	específica	23	medio	intermedio	tratamiento	77	131	,230	21	33	95	,072
27	27	mujer	social	31	medio	precoz	tratamiento	72	133	,273	25	35	92	,028
28	28	mujer	agorafobia	30	medio	intermedio	tratamiento	80	135	,289	36	35	93	,071
29	29	mujer	social	23	bajo	intermedio	tratamiento	80	144	,331	31	30	77	,028
30	30	mujer	social	21	bajo	tardío	tratamiento	72	137	,301	24	43	101	,038
31	31	mujer	agorafobia	20	bajo	tardío	tratamiento	81	136	,312	26	44	96	,000

Figura 1.9

1.4. Gestión y Modificación de variables.

Para realizar la mayoría de operaciones que corresponden a este subsistema deberemos acceder mediante la opción **Transformar** de la barra principal de menús desplegables. A continuación comentamos algunas de las opciones de transformación que permite el *SPSS*.

1.4.1. Generar nuevas variables

Seleccionaremos la opción **Calcular**, que nos activará el cuadro de diálogo que permite el cálculo de una nueva variable (véase Figura 1.10). En dicho cuadro, en la parte dedicada a variable de destino, deberemos especificar el nombre de la nueva variable a generar (con las mismas restricciones que anteriormente hemos indicado para los nombres de variables). Pulsando con el ratón en el botón **Tipo y Etiqueta** se nos abre otro cuadro de diálogo donde deberemos seleccionar el tipo de variable que generaremos (Numérico o Cadena) y, si lo deseamos, escribir una etiqueta más larga que nos identifique el contenido de la variable que vamos a generar. Para activar los cambios presionamos el botón **Continuar**. En el cuadro **Expresión Numérica** construiremos la expresión que nos permitirá generar la nueva variable. Dicha expresión puede estar formada por:

- a) **Variables:** Debajo del botón **Tipo y Etiqueta** observamos la presencia de un listado con los nombres de las variables que se hallan disponibles en la matriz de datos activa.
- b) **Operadores:** Tanto de carácter aritmético, relacional o lógicos. En la calculadora disponible se encuentran los símbolos que representan a estos operadores: “+”, suma; “-”, resta; “*”, multiplicación; “/”, división; “**”, potencia; “<”, menor que; “>”, mayor que; “<=”, menor o igual que; “>=”, mayor o igual que, “=”, igualdad lógica, “≠”, desigualdad lógica; “&”, operador lógico “y”; “|”, operador lógico “o”; “~”, negación lógica; “()”, agrupación.
- c) **Funciones:** Bajo el cuadro de expresión numérica podemos visualizar un cuadro contenedor de las funciones matemáticas que tiene disponibles el programa, así como, entre paréntesis, los argumentos necesarios para cada una de las funciones.
- d) **Constantes:** Cuando en la expresión numérica es preciso incluir algún valor entero o real.

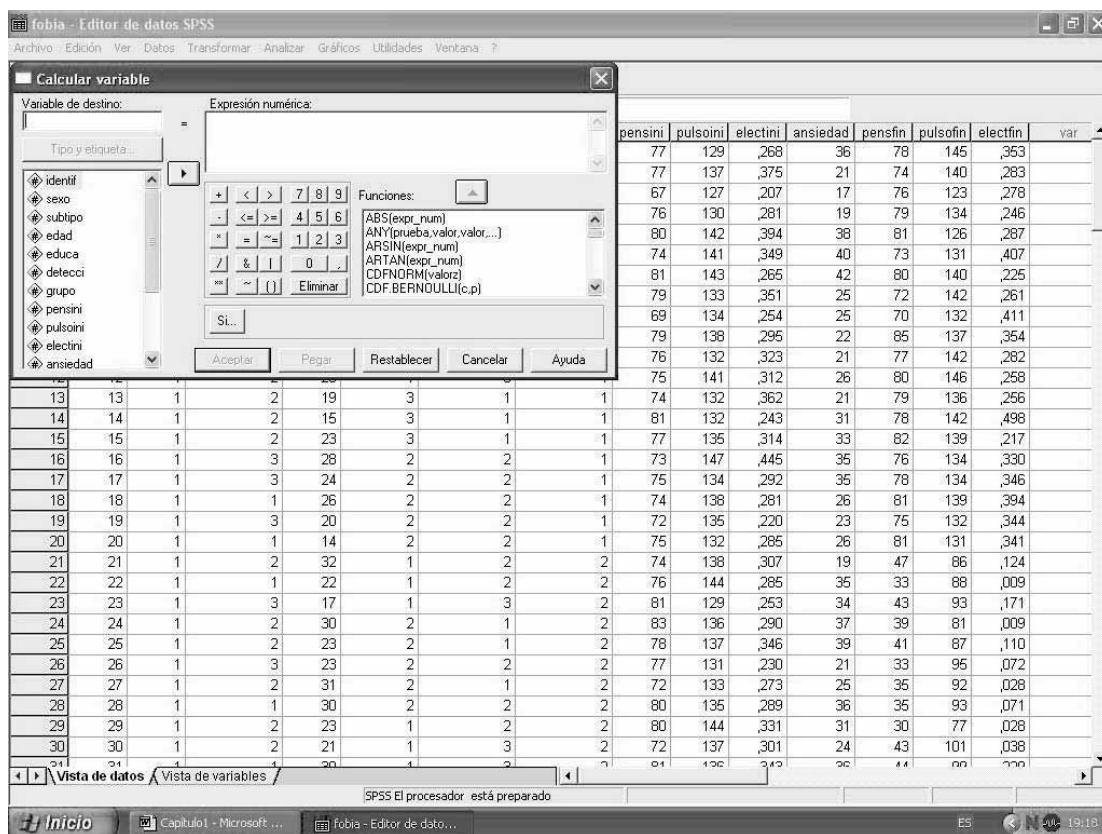


Figura 1.10

Observamos también la presencia de un botón identificado como **Si...**, cuya activación nos permitirá acceder a otro cuadro de diálogo mediante el que podremos construir una expresión condicional que provocará que la nueva variable sólo sea generada si la expresión es cierta. Para activar esta expresión deberemos marcar la opción **Incluir si el caso satisface la condición**, puesto que, por defecto, está señalada la opción **Incluir todos los casos**. Una vez especificada la expresión pulsaremos el botón **Continuar** para volver al cuadro de diálogo anterior. A continuación, presionamos el **Aceptar**. Automáticamente el programa nos enviará a la Ventana del editor de datos, apareciendo en la última columna de la matriz la nueva variable generada. Si, por el contrario, pulsamos el botón **Pegar**, el conjunto de comandos especificado a través de los cuadros de diálogo se nos escribirá automáticamente en la **Ventana de Sintaxis** designada, y la generación de la nueva variable quedará pendiente hasta que ejecutemos los comandos desde dicha ventana.

1.4.2. Recodificación de variables

En ocasiones nos puede interesar proceder a recodificar alguna de las variables preexistentes, por ejemplo, mediante la aplicación de otra escala de medida diferente de la original. En este caso damos por supuesto que la transformación a aplicar es permisible sobre los datos originales. Para llevar a cabo el proceso, escogeremos del menú desplegable **Transformar** la opción **Recodificar**. Al activar esta opción se nos ofrecerá la posibilidad tanto de recodificar sobre la misma variable (véase Figura 1.11) como generar una nueva variable con los valores recodificados (véase Figura 1.13). La decisión de escoger una u otra opción vendrá determinada por el hecho de si nos interesa conservar o no en la matriz de datos activa los valores originales de la variable. Recordemos que esta recodificación, tanto si genera una nueva variable como si no lo hace, únicamente afectará a los datos ubicados en el editor de datos. Para hacer permanentes las modificaciones en un fichero externo deberemos proceder a guardar el contenido de la ventana del editor de datos (véase apartado 1.6).

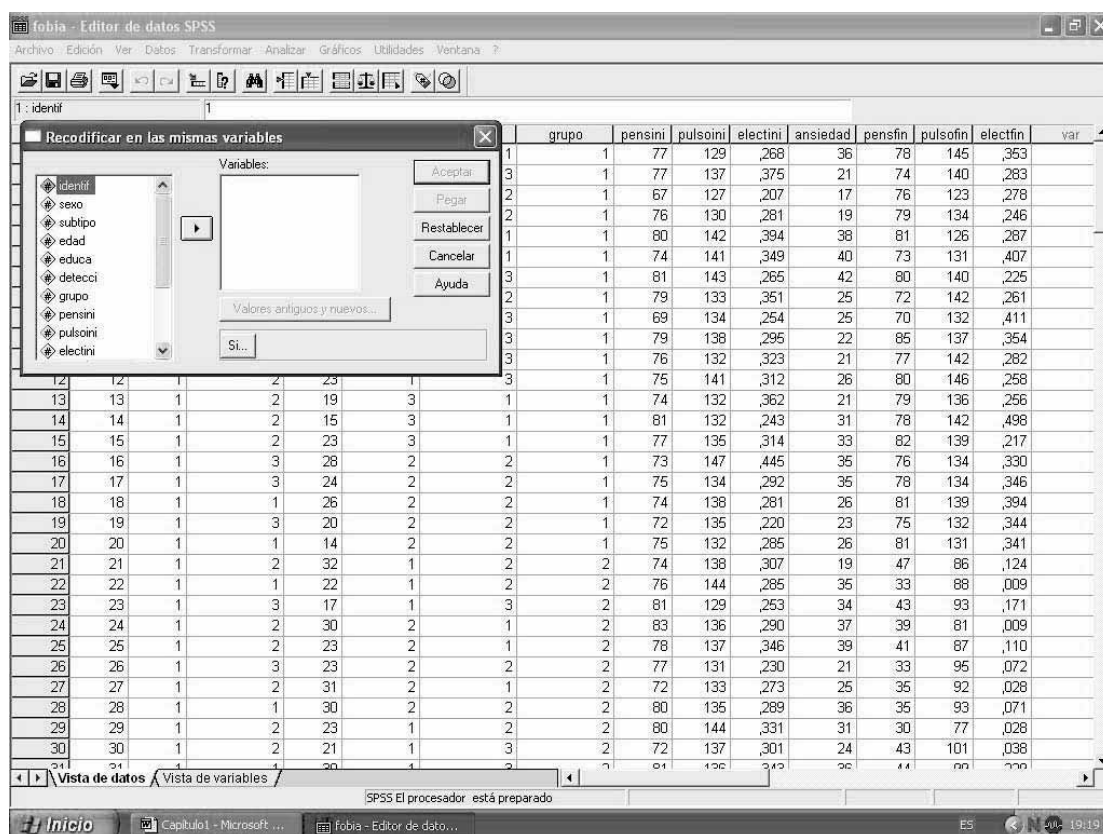


Figura 1.11

Si escogemos la primera de las opciones, es decir, recodificar sobre los datos de la misma variable, simplemente tendremos que pasar, en el cuadro de diálogo adecuado, la o las variables a recodificar del listado de la izquierda al cuadro en blanco de la derecha. Resaltaremos la variable escogida señalando el nombre de la misma con el puntero del ratón y, posteriormente, pulsaremos el botón de entrada (►). Así conseguiremos ubicar la variable en el cuadro de la derecha. Es posible deshacer el proceso resaltando de igual forma la variable a eliminar del cuadro de la derecha y pulsando el mismo botón anterior, que en este caso se nos presenta en posición invertida (◄).

Cuando únicamente nos interese realizar el proceso de recodificación para aquellos sujetos que cumplan una regla determinada, procederemos a activar la recodificación condicional pulsando el botón **Si...** Mediante éste accederemos al cuadro de diálogo donde podremos especificar y generar la expresión cuyo cumplimiento marcará la recodificación efectiva del valor del sujeto. También será necesario marcar la opción **Incluir si el caso satisface la condición**, en lugar de la opción **Incluir todos los casos**. Volveremos al cuadro de diálogo inmediatamente anterior mediante la habitual opción **Continuar**.

Pulsando posteriormente el botón **Valores antiguos y nuevos**, entraremos en el cuadro de diálogo que nos permitirá concretar el criterio de recodificación (véase Figura 1.12). A tal fin, sustituimos un valor antiguo (el original de la matriz de datos) por uno nuevo (el valor recodificado). Finalmente, pulsando el botón **Añadir** nos aparecerá la codificación escogida en el cuadro **Antiguo**→ **Nuevo**. Para especificar la recodificación podemos escoger entre las siguientes opciones:

- Recodificar valor a valor.
- Recodificar los valores *missing* de sistema (perdido por el sistema).
- Recodificar los valores perdidos por el sistema o por el usuario.
- Recodificar un rango de valores (ambos inclusive).
- Recodificar desde el mínimo hasta un determinado valor.
- Recodificar desde un determinado valor hasta el máximo.
- Recodificar los restantes valores.

Para los valores resultantes podemos escoger entre dos opciones:

- Especificar un único valor.
- Determinar el valor de salida como perdido por el sistema.

En caso de que nos hayamos equivocado al realizar alguna de las recodificaciones siempre es posible modificarla o eliminarla antes de abandonar el presente cuadro de diálogo, resaltándola con el ratón en el cuadro **Antiguo**→ **Nuevo** y pulsando, alternativamente los botones **Cambiar** o **Borrar**, según sea la opción deseada.

También debemos tener en cuenta para el proceso de recodificación que cada valor de la matriz original es recodificado una única vez y que los valores no recodificados permanecerán inalterados en la nueva variable. Pulsando el botón **Continuar** volvemos al cuadro de diálogo anterior y, pulsando el botón **Aceptar**, conseguiremos efectuar el proceso de recodificación.

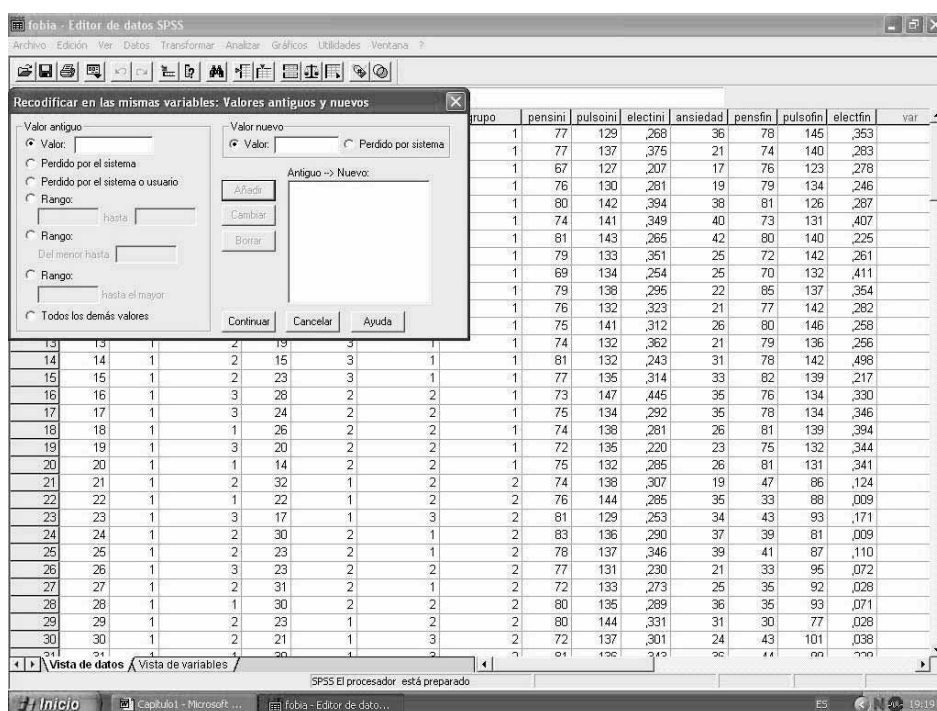


Figura 1.12

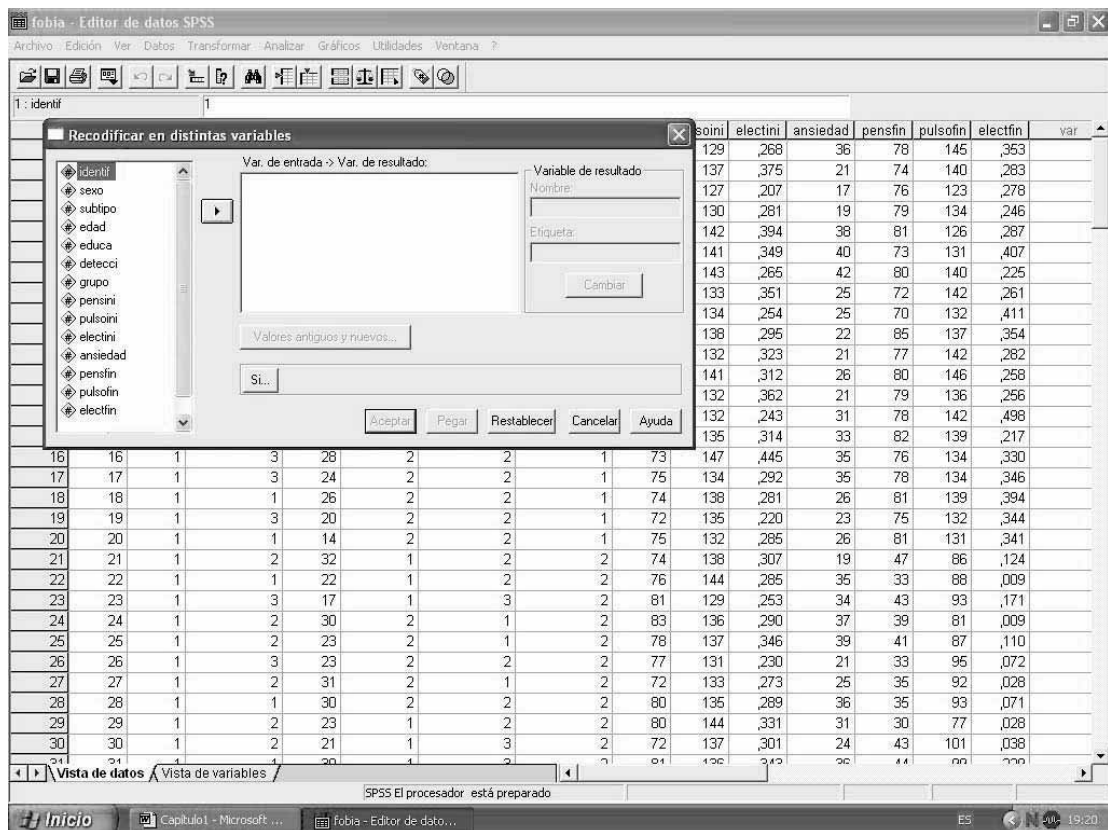


Figura 1.13

Si escogemos dentro de la opción **Recodificar** la segunda de las opciones (**En distintas variables**), el proceso de recodificación nos generará una nueva variable a partir de la original. En gran parte el proceso es similar al descrito anteriormente para la recodificación sobre la misma variable, pero existen algunas particularidades que conviene apuntar. Tal y como se puede observar en la Figura 1.13, deberemos especificar en el recuadro correspondiente un nombre para la **Variable de entrada** (variable original) y otro para la **Variable de resultado** (nombre de la nueva variable ya recodificada). La variable original la escogeremos de entre las variables existentes en la matriz de datos de la lista del cuadro de la izquierda, resaltándola con el ratón y ubicándola en el cuadro de la derecha mediante el botón de entrada (►). El nombre de la variable de resultado lo especificaremos donde pone **Nombre** en el cuadro de variable de resultado, pulsando posteriormente el botón **Cambiar**. Este proceso de recodificación también admite el uso de una expresión condicional que limite la recodificación a aquellos sujetos que la cumplan. Mediante el botón **Valores antiguos y nuevos** realizaremos el procesamiento de recodificación, de forma similar a la descrita para el proceso anterior, que finalizará con la pulsación de los botones **Continuar** y **Aceptar**. La nueva variable generada aparecerá en la primera columna libre de la matriz de datos del editor. Adicionalmente, este proceso de transformación permite, como opciones para los valores nuevos, copiar los antiguos, generar una cadena como variable de resultado y transformar una cadena numérica en un número.

1.4.3. Otras opciones

Mediante este subsistema de gestión y modificación de ficheros de datos es posible efectuar otras operaciones, tales como: la **Recodificación automática de valores**, la **Ordenación por rangos** y el **Recuento de valores**. A estos procedimientos también se accede desde la opción **Transformar** de la barra principal de menús. Además, podemos realizar las acciones **Definición** y **Uso de conjuntos de variables** desde la opción **Utilidades** de la barra principal de menús.

Para una visión más detallada y mejor comprensión de estas opciones volvemos a sugerir la utilización de los tutoriales y menús de ayuda incluidos en el programa.

1.5 Gestión de los Datos

El subsistema de gestión y transformación de los datos desde el editor nos ofrece diferentes posibilidades de gestión de los casos de la matriz de datos activa. A este subsistema se accede mediante la opción **Datos** de la barra principal de menús. Las tareas que permite este subsistema se pueden agrupar en dos tipos:

- a) **Gestión y manipulación de tipo genérico:** nos permite definir y copiar las propiedades de las variables mediante un proceso automatizado, insertar nuevas variables o casos en un punto determinado de la matriz y desplazar la visualización hasta un caso concreto de la matriz de datos. Al respecto de estas posibilidades, en muchos casos, su funcionamiento es similar a otros programas conocidos. Además las opciones que permiten son accesibles, en algunos casos, a través de otros subsistemas (por ejemplo la gestión de variables). Por ello remitimos al lector interesado en alguna de ellas al manual de ayuda del programa.
- b) **Gestión y transformación de archivos:** nos permite manipular los casos del archivo o de diversos archivos de datos de forma que su estructura se halle en consonancia con el procedimiento de *SPSS* que se desee utilizar. De entre la posibilidades existentes, podemos mencionar la ordenación de los sujetos (ascendente o descendente) según los valores de una determinada variable; trasponer la matriz de datos; reestructurar la matriz de datos mediante un asistente; agregar variables; identificar casos duplicados; seleccionar casos; ponderar casos en función de una variable; segmentar el fichero de trabajo para realizar un mismo análisis o procedimiento para todos los grupos de segmentación con una única ejecución; y fundir distintos archivos de datos para agregar nuevos sujetos con variables comunes o nuevas variables para los mismos sujetos.

En el presente texto nos centraremos, sobre todo, en exponer los procedimientos de selección específica de casos. Cuando nos interese seleccionar solamente una parte de los sujetos incluidos en la matriz de datos para realizar algún tipo de operación o análisis con algunas de sus variables, deberemos acceder al subsistema de manipulación de casos mediante la opción **Seleccionar casos** dentro del menú **Datos** de la barra principal de menús (como podemos ver en la Figura 1.14).

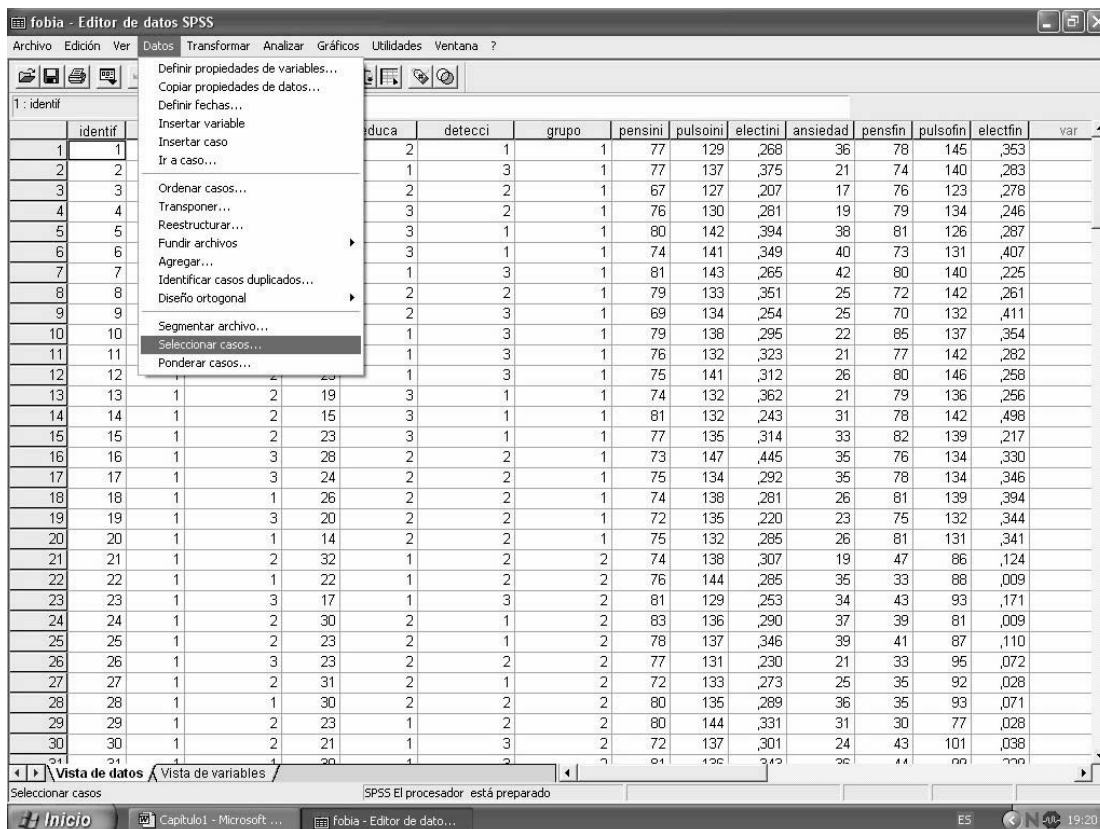


Figura 1.14

Al activar la opción **Seleccionar casos** accederemos al cuadro de diálogo que nos permite realizar dicha operación, tal como se muestra en la Figura 1.15. Observamos cómo tenemos activada la primera de todas las opciones, que indica que vamos a trabajar con todos los casos. Si deseamos seleccionar un grupo específico de sujetos deberemos activar alguna de las otras condiciones, detalladas a continuación:

- a) La segunda opción nos permite realizar la selección condicional según el criterio de la expresión que podemos generar a partir del cuadro de diálogo que se nos activa al pulsar el botón **Si...**. Este cuadro de diálogo es de características similares al que se había utilizado para el cálculo de nuevas variables.
- b) La tercera opción nos permitirá seleccionar una *muestra aleatoria* de los casos incluidos en la matriz de datos. Esta selección se puede realizar en base a un porcentaje aproximado o a un número exacto de casos. El muestreo se realiza sin reposición.

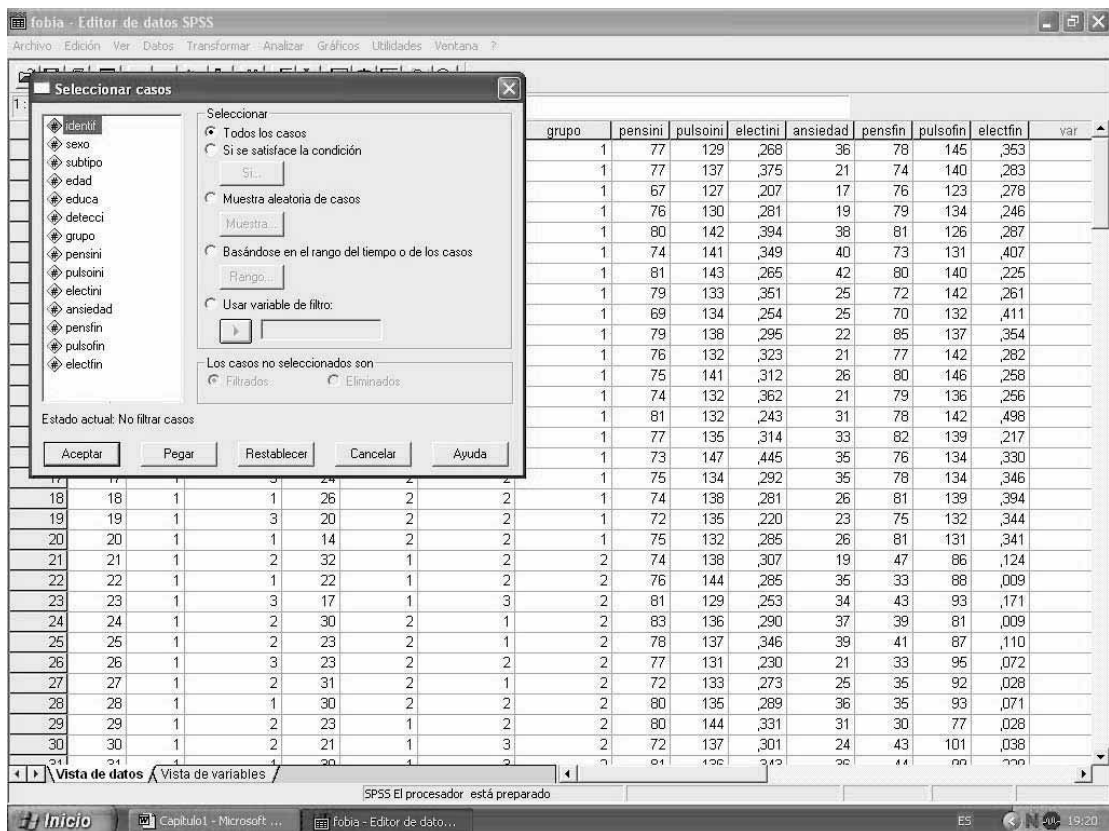


Figura 1.15

- c) La cuarta opción nos permite seleccionar los casos según un determinado segmento de la matriz de datos, basándose en un rango de números de caso o en un rango de fechas u horas.
- d) La última de las opciones nos permite seleccionar los sujetos utilizando como filtro los valores de una variable (se seleccionan todos los sujetos que tienen valores no nulos en aquella variable).

Nótese como, para completar el proceso de selección, es necesario indicar si éste es temporal (**Filtrados**), con lo que los sujetos no seleccionados permanecen en la matriz de datos activa y se puede volver a operar con ellos deshaciendo el proceso de selección o permanente (**Eliminados**), con lo que los sujetos no seleccionados desaparecen de esta matriz de datos y para recuperarlos deberíamos cargar de nuevo los datos originales desde el fichero externo. Para realizar de forma efectiva la selección escogida, deberemos pulsar el botón **Aceptar** del cuadro de diálogo. Para recuperar los objetos filtrados, simplemente debemos volver a marcar en el cuadro de diálogo correspondiente la opción **Todos los casos** y pulsar el botón **Aceptar**. En caso de que nos interese la última de las opciones y eliminar los sujetos que no hemos analizado tendremos la precaución, al finalizar la sesión, de salir sin grabar el fichero de datos. Esto es así porque, al haberlo modificado antes de salir, nos preguntará si deseamos guardar las modificaciones, lo que afectaría al fichero externo y perderíamos la información eliminada. En todo caso, si deseamos mantener un fichero con la información no eliminada, previamente deberemos guardar los datos en otro fichero.

1.6 Guardar información en un archivo externo e impresión de listados

Si deseamos archivar el contenido de cualquiera de las ventanas en un fichero externo, en primer lugar nos ubicaremos en dicha ventana para activarla mediante la opción **Ventana**. Al elegir esta opción, y al final del menú desplegable, identificaremos un listado de todas las ventanas abiertas en la presente sesión *SPSS*.

Con el puntero del ratón señalaremos la ventana escogida y accederemos a ella presionando el botón izquierdo. Situados en esta ventana, y estando ésta activa, escogeremos la opción **Archivo** del menú principal. Posteriormente, seleccionaremos la subopción **Guardar como**, tanto si es la primera vez que se va a guardar el contenido de la ventana como si queremos almacenarlo en un fichero con distinto nombre al que estábamos utilizando. Una vez escogida la opción correspondiente, accederemos al cuadro de diálogo que nos permitirá llevar a cabo la grabación. En la Figura 1.16 se muestra un ejemplo para guardar, en un fichero externo, el contenido de una ventana de resultados. Dependiendo del tipo de ventana desde que ejecutemos la opción de guardar el contenido, cada uno de los distintos tipos de ficheros que se crean llevará asociada una extensión por defecto:

- a) *sav*, en caso de archivos de datos
- b) *sps*, en caso de archivos de resultados (a diferencia de versiones anteriores en las que la extensión era *lst*)
- c) *sps*, en caso de archivos de sintaxis.

En el nombre del archivo sustituiremos el asterisco por una cadena que no supere los ocho caracteres. No es necesario especificar la extensión, pues ésta se asigna por defecto. Así mismo, seleccionaremos la unidad y el directorio de destino del fichero de salida. Al pulsar la tecla **Aceptar** se ejecutará la grabación en el fichero externo. Esta operación deberá repetirse para cada uno de los contenidos de las ventanas que deseemos grabar. Podemos acceder al contenido de estos ficheros en sesiones posteriores mediante las opciones **Archivo** y **Abrir**. En el cuadro de diálogo que se nos activa especificaremos la ruta de acceso hasta el fichero (directorio y subdirectorios) y la clase de archivo en **Tipos de archivos** (debe indicarse la extensión correspondiente al tipo de archivo que deseamos abrir).

Para imprimir el contenido de una ventana activa simplemente será necesario pulsar el icono correspondiente a **Imprimir** en la barra de herramientas, que ocupa la tercera posición comenzando por la izquierda. También podemos, desde la barra de menú principal, escoger la opción **Archivo** y activar la opción **Imprimir**. En el cuadro de diálogo que se abrirá podremos determinar si deseamos imprimir todo el contenido de la ventana o solamente una selección de ésta, así como el número de copias que vayamos a imprimir. Pulsando el botón **Aceptar** se procederá a la impresión efectiva del contenido de la ventana.

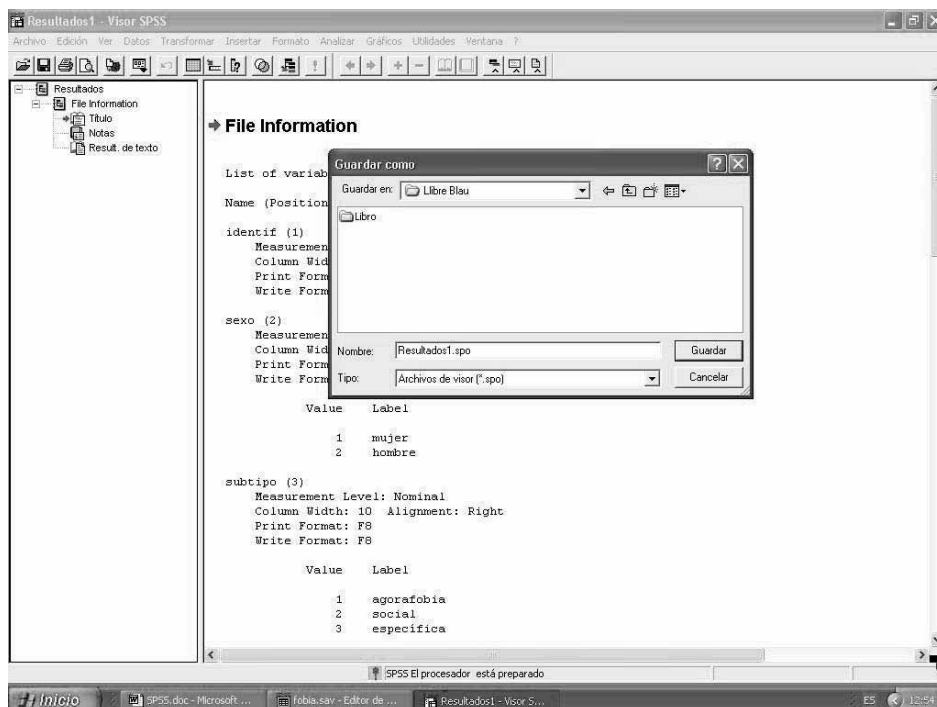


Figura 1.16