

SPSS: UN INSTRUMENTO DE ANÁLISIS DE DATOS CUANTITATIVOS

Esperanza Bausela Herreras¹

Departamento de Filosofía y Ciencias de la Educación
Áreas M.I.D.E. y P.E.T.R.A.
Universidad de León (España)
dfcebh@unileon.es

Resumen: En este artículo presentamos un software informático, SPSS, como instrumento de análisis cualitativo de datos, presentamos las posibilidades básicas y más importantes del mismo. El SPSS, junto con el BMDP son los más utilizados en investigación aplicada a las Ciencias Sociales, hemos de señalar que aunque destacamos su utilidad, su fácil manejo y su fácil comprensión, no obstante, su uso se ve supeditada a un periodo determinado, condicionado por una licencia, lo cual en ocasiones dificulta y entorpece poder sacar su máximo rendimiento. Esta dirigido a alumnado que cursa materias relacionadas con la investigación educativa, en titulaciones como educación especial, audición y lenguaje y psicopedagogía, así, como alumnos doctorandos, siendo uno de los software que actualmente estamos impartiendo en la asignatura Métodos de Investigación en Educación de la licenciatura de psicopedagogía

Palabras Clave: SPSS, análisis de datos cuantitativos, matriz de datos.

Summary: In this paper we present a computer software, SPSS, like instrument of qualitative analysis of data, we present the basic and more important possibilities of this. The SPSS, together with the BMDP are the more used in research applied to Social Sciences, we must point out that although we highlight its utility, its easy handling and its easy understanding, nevertheless, its use is subordinated to a certain period, conditioned by a license, that which hinders in occasions and he/she hinders to be able to take out its maximum yield. This directed to pupil that studies matters related with the educational research, like special education, audition and language and education psychology and students doctorandos, being one of the software that at the moment are teaching in the subject Methods of Research Education of the education psychology degree.

Words Key: SPSS, analysis of quantitative data, womb of data,

¹ Becaria de investigación de la Universidad de León bajo la codirección de los profesores: Dr. Delio del Rincón Igea y Dr. Dionisio Manga Rodríguez.

1. Programas informáticos utilizados en el análisis de datos

Desde los años ochenta venimos experimentando una invasión del mundo de la informática en el campo de la investigación educativa a todos los niveles; bases de datos, bases bibliográficas... pero muy especialmente en el análisis de los datos. Los aplicaciones de la informática a la investigación son importantes, teniendo en cuenta que cada vez más la cantidad de datos que se manejan y la variedad de análisis que se realizan rebasan la capacidad del cálculo manual (Buendía, Colas y Hernández, 1997).

Una relación completa de programas de Estadística Aplicada, aparte de ser difícil de elaborar, no es el fin de esta tesis, por lo que únicamente presentamos algunos programas para el análisis de datos cuantitativos que nos pueden servir de ejemplo: BMDP, LISREL, SAS, SCA, SPAD, CLUSTAN, BOX – JENKIS, MDS, INDSICAL, OSIRIS, STAP – PACK. Todos estos programas son dirigidos a profesionales con un coste económico alto. A nivel más simple destacamos SYSTAT y MINITAB, asequibles a cualquier estudiante.

Las últimas versiones de todos estos programas, operando bajo Windows, comparten una característica básica como es la disponer de una hoja, que sin llegar a ser como una hoja de cálculo, permiten la definición de variables actuando en las cabeceras de columnas, para posteriormente registrar los datos en las filas. Esta capacidad, en cierta medida anula la necesidad del empleo de una de cálculo tipo LOTUS o EXCEL como paso previo, aunque en cualquier caso es posible, posteriormente la importación de los datos desde las mismas (González – Conde, 2000).

El SPSS, junto con el BMDP son los más utilizados en investigación aplicada a las Ciencias Sociales (Bisquerra, 1989), hemos de señalar que aunque destacamos su utilidad, su fácil manejo y su fácil comprensión, no obstante, su uso se ve supeditada a un periodo determinado, condicionado por una licencia, lo cual en ocasiones dificulta y entorpece poder sacar su máximo rendimiento.

Dada la relevancia de la informática en el análisis estadístico de datos, finalizamos enumerando las ventajas e inconvenientes que podrían derivarse de ello (Álvarez, García, Gil, Martínez, Romero y Rodríguez, 2002) (ver figura 1)

USO DE LA INFORMÁTICA EN EL ANÁLISIS	
VENTAJAS	INCONVENIENTES
Permite un importantísimo ahorro de tiempo y esfuerzo realizando un trabajo en horas e incluso días.	El aprendizaje del manejo de paquetes de programas estadísticos requiere un cierto esfuerzo.
Hace posible cálculos más exactos, evitando los redondeos y aproximaciones del cálculo manual.	A veces, la capacidad de cálculo del evaluador supera la capacidad para comprender el análisis realizado e interpretar los resultados.
Permite trabajar con grandes cantidades de datos, utilizando muestras mayores e incluyendo más variables.	Lleva a veces a una sofisticación innecesaria, al permitir el empleo de técnicas complejas para responder a cuestiones simples
Permite trasladar la atención desde las tareas mecánicas de cálculo alas tareas conceptuales: decisiones sobre el proceso, interpretación de resultados, análisis crítico.	

Figura 1 Ventajas e inconvenientes del uso de la informática en el análisis de datos (tomado de Álvarez, García, Gil, Martínez, Romero y Rodríguez, 2002)

2. Procedimiento y métodos estadísticos

El punto de partida para el análisis de datos comienza con una matriz de datos $n \times p$. Esta matriz se registro de forma natural en las filas y columnas de una hoja del programa SPSS. Una matriz es un conjunto de valores representadas en n filas y n columnas. Nuestra matriz de datos figura en el anexo, ésta esta configurada por xxx columnas, que

representa las variables investigadas y las variables investigadoras, e yyy filas que representan los casos objeto de estudio. En el anexo se incluyen las matrices de esta investigación, esta por razones de espacio ha tenido que ser seccionada en diversos fragmentos.

Una vez recogidos los datos, procedimos a describirlos y a resumirlos. Esta descripción se efectuó mediante *descripciones gráficas* (polígonos de frecuencias, curva normal..) y *descripciones numéricas* (promedios, medidas de variabilidad, forma de la distribución, medida de la relación entre variables...).

Las *técnicas multivariantes*, son según Bisquerra (1989) aquellas técnicas que se aplican al análisis de muchas variables, siendo el tratamiento de los datos, por tanto, multidimensional. Recogiendo las aportaciones de diversos autores el profesor Bisquerra (1989) propone una clasificación de los métodos multivariantes. Los *criterios clasificatorios* son: (i) métodos descriptivos o explicativos, (ii) número de variables dependientes, (iii) escalas de medida-. En cierta forma, esta clasificación es considerada por el autor como un intento de algoritmo conducente a la determinación del método más adecuado para las características del problema objeto de estudio. En él se formulan una serie de preguntas que facilitan la toma de decisión.

3. Descripción del SPSS

El SPSS (*Statistical Product and Service Solutions*) es una potente herramienta de tratamiento de datos y análisis estadístico.

1.1. Estructura general

⇒ Al iniciar una sesión aparece una ventana *Editor de datos*, es la ventana principal, similar a una hoja de cálculo. Esta ventana muestra dos contenidos diferentes:

- a. Los *datos* propiamente dichos.
- b. *Variables* del archivo acompañadas del conjunto de características que las definen.

⇒ La *barras de menús* contienen una serie de submenús desplegables que

permiten controlar la mayoría de las acciones que el SPSS puede llevar a cabo (ver imagen 1):

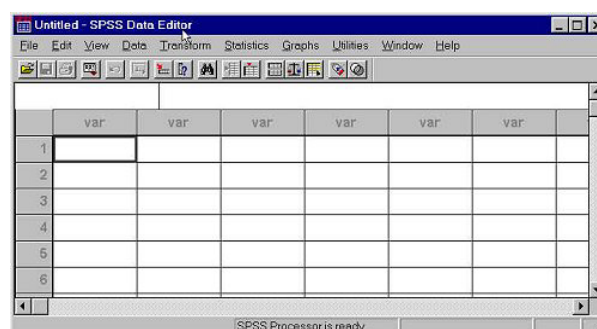


Imagen 1 Ventana del editor de datos

- *Archivo* → Abrir y guardar todo tipo de archivos
- *Edición* → Cortar, copiar, buscar...
- *Ver* → Permite controlar el aspecto de las celdas y en el visor de resultados, sirve para mostrar/ocultar resultados concretos.
- *Datos* → Contiene aspectos relacionados con: fusionar archivos, trasponer filas y columnas...
- *Transformar* → Permite crear y cambiar nuevas variables.
- *Insertar* → Contienen opciones para insertar textos, gráficos..
- *Galería* → Permite seleccionar diferentes tipos de gráficos para unos mismos datos.
- *Diseño* → Permite controlar características de un gráfico
- *Serie* → Permite mostrar /ocultar las categorías en el eje de abscisas.
- *Proceso* → Sirve para crear nuevas funciones y subgrupos de procesamiento, controlar color y tipo de letra de letra de las palabras clave.
- *Depurar* → Permite depurar procesos básicos y acceder a un editor de objetos.
- *Formato* → Permite cambiar la apariencia de textos, gráficos...
- *Analizar* → Contiene todos los procedimientos estadísticos.
- *Gráficos* → Permite generar todo tipo de gráficos.

- *Utilidades* → Permite obtener información sobre las variables y sobre el archivo de datos.
- *Ventana* → Permite cambiar de una ventana a otra dentro del SPSS y minimizar todas las ventanas abiertas.

⇒ *Barra de herramientas*, es un conjunto de botones-íconos que permiten ejecutar algunas de las funciones anteriores del SPSS de forma rápida y sencilla.

- a. Abrir archivo
- b. Guardar archivo
- c. Imprimir
- d. Recuperar cuadros de diálogo
- e. Deshacer/rehacer
- f. Ir a gráfico
- g. Ir a caso
- h. Variables
- i. Buscar
- j. Insertar caso
- k. Mostrar etiquetas de valor
- l. Usar conjuntamente
- m. Insertar variable
- n. Segmentar archivo
- o. Ponderar casos
- p. Seleccionar casos

1.2. Una sesión con el SPSS

Esta sesión es estándar, y en ella desarrollamos las siguientes tareas básicas:

- Abrir un archivo de datos:
 - ⇒ Seleccionar la opción *Abrir...* del menú *Archivo* para acceder al cuadro de diálogo *Abrir archivo*.
 - ⇒ Buscar en el cuadro de diálogo *Abrir archivo* el archivo titulado *Análisis de Necesidades* → seleccionarlo, → pulsar dos veces el botón derecho del ratón (o, alternativamente, seleccionarlo y pulsar el botón *Abrir*).
- Ejecutar un procedimiento estadístico:
 - ⇒ *Analizar* → seleccionar *Estadísticos descriptivos* → *Frecuencias* (igual forma de proceder con los otros estadísticos).
 - ⇒ *Gráficos* → *Frecuencias* → Seleccionar las opciones *Histograma* y *curva normal* → *Continuar* → *Aceptar*.

- Examinar los resultados:

La ventana del *Visor de resultados* se encuentra dividida verticalmente en dos paneles:

- a. *Esquema del Visor* (panel izquierdo), que contienen un índice de los resultados generados por el SPSS.
 - b. *Contenido del Visor* (panel derecho), que contienen los resultados (texto, tablas y gráficos) generados por el SPSS.
- *Imprimir*: Seleccionar la opción *Imprimir* del menú *Archivo*.
 - *Salir del SPSS*: La opción *Salir* cierra el programa.

1.3. El editor de datos

Un archivo de datos puede crearse de dos formas:

I.1. *Introduciendo* datos directamente en el Editor de datos.

I.2. *Importando* la información ya existente de alguna fuente externa (archivo de texto, hoja de cálculo o una base de datos).

La estructura del Editor de datos del SPSS es siempre la misma:

- Las filas representan casos.
- Las columnas representan variables.
- Cada casilla contiene un valor.
- El *archivo del Editor* de datos es siempre rectangular.
- Definir variables

Pulsar la solapa *Vista de Variables*. Nos permite llevar a cabo todas las tareas relacionadas con la definición de una variable:

- Asignar nombre a una variable

Situar el cursor, dentro de la columna *Nombre*, en la casilla correspondiente a la variable cuyo nombre se desea crear o editar y escribir el nombre.

- Definir el tipo de variable: Numérica, Coma, Punto, Notación científica, Fecha Dólar, Moneda personalizada y Cadena
 - ✓ Una variable con formato numérico, coma o punto, acepta cualquier número de decimales → *Cifras decimales*.
 - ✓ Una variable de cadena se almacena y procesa respetando la anchura establecida en la columna → *Anchura*.
 - Asignar etiquetas
- ⇒ Para Asignar etiqueta a una variable → Cursor en columna *Etiquetas*.
- ⇒ Para asignar etiquetas a los valores de una variable → situar el cursor en *valor* → escribir el valor → escribir etiqueta (por ejemplo, hombre) → pulsar añadir → los botones *Cambiar* y *Borrar* permiten modificarlos y eliminarlos.
- Definir valores perdidos → Situar el cursor en la columna *Perdidos*.
 - Definir el formato de columna → Situar el cursor en la columna *Columnas*.
 - Alinear texto → Situar el cursor en la columna *Alineación* → Por defecto del sistema variables numéricas derecha y variables de cadena izquierda.
 - Asignar un nivel de medida → Situar el cursor en la columna *Medida* → Niveles de medida:
 - *Escala* → Para variables cuantitativas obtenidas con una escala de intervalo o razón: edad, salario, temperatura...
 - *Ordinal* → Para variables cuantitativas obtenidas con una escala ordinal; nivel educativo, clase social...
 - *Nominal* → Para variables categóricas medidas con una

escala nominal; sexo, lugar de procedencia...

⇒ Introducir datos

El *editor de datos* nos ofrece la posibilidad de modificar el archivo de datos de múltiples maneras:

- Deshacer /rehacer
- Seleccionar datos
- Mover y copiar datos
- Borrar datos
- Buscar datos
- Buscar casos
- Buscar variables
- Insertar
- Insertar variables nuevas
- Insertar casos nuevos

1.4. Transformar datos

La preparación del archivo de datos incluye desde la simple detección y corrección de los posibles errores cometidos al introducir datos, hasta transformaciones, pasando por la recodificación de los códigos para los valores de alguna variable, o la creación de nuevas variables a partir de otras ya existentes. Todas estas opciones se encuentran en el menú **Transformar**.

- Calcular → seleccionar opción *Calcular*
- Categorizar condicionales → permite seleccionar dos casos: Incluir todos los casos e incluir si el caso satisface la condición
- Recodificar → Seleccionar la opción recodificar en distintas variables → seleccionar variable a recodificar y llevarla a la lista *Var. De entrada* → *Var. De resultado* → en el recuadro Variable de resultado escribir el nombre elegido de la nueva variable → pulsar *Cambiar* y situarlo en la *var. De entrada* → *Var. De resultado* → pulsar el botón *Si...* → pulsar el botón *valores antiguos y nuevos*.
- Categorizar variables → Seleccionar la opción categorizar variables → crear categorías y número de categorías.

1.5. Modificar archivos de datos

Los archivos de datos no siempre están organizados de forma idónea. En ocasiones puede ser interesante cambiar el orden de los casos, trasponer filas y las columnas. Todas estas opciones se encuentran en el menú **Datos**.

- Ordenar casos → Seleccionar la opción *Ordenar casos* del menú *Datos* → *ordenar por* u *ordenar de clasificación* (ascendente o descendente).
- Fundir archivos → Opción *Fundir archivo*, cuando se trabajan con dos archivos → Dos opciones: Añadir casos y añadir variables .
- Segmentar archivo → Opción *Segmentar archivo*.
- Seleccionar variables → Opción *Seleccionar casos, diversas opciones*:
 - Todos los casos
 - Si se satisface la condición
 - Muestra aleatoria de casos
 - Basándose en el rango del tiempo o de los casos
 - Usar variable de filtro → Los casos no seleccionados pueden ser: Filtrados y Eliminados.

II.1. Características descriptivas de la muestra

1.6. Análisis descriptivo: Frecuencias y descriptivos

En este punto describimos dos procedimientos que permiten obtener la información necesaria para caracterizar apropiadamente tanto variables categóricas como cuantitativas: el procedimiento **Frecuencias** y el procedimiento **Descriptivos**.

- *Analizar* → *Frecuencias* → *Estadísticos descriptivos* → trasladar variables → pulsar mostrar tablas de frecuencias → pulsar el botón *Estadísticos* → subcuadro de diálogo Frecuencias: *Estadísticos* disponibles:
 - *Valores percentiles*: Cuartiles, puntos de corte para K grupos iguales, percentiles.
 - *Tendencia central*: Media, mediana, moda y suma.

- *Dispersión*: Desviación típica, varianza, amplitud, mínimo, máximo y E.T. media.
- *Distribución*: Asimetría y curtosis.
- Los valores son puntos medios de grupos.

- *Analizar* → *Frecuencias* → *Estadísticos descriptivos* → trasladar variables → pulsar *mostrar tablas de frecuencias* → Pulsar el botón Gráficos. Tipos de gráficos que pueden elegirse: Ninguno, Gráfico de barras, Gráfico de sectores e Histograma. Con los siguientes valores: Frecuencias y Porcentaje.

- *Descriptivos*: A diferencia de lo que ocurre con el procedimiento Frecuencias, que contiene opciones para describir tanto variables categóricas como cuantitativas continuas. El procedimiento descriptivos está diseñado únicamente para *variables cuantitativas continuas*.

Analizar → seleccionar *Estadísticos descriptivos* → Descriptivos del menú *Descriptivos* → Trasladar una o más variables → *Opciones*:

- *Media*
- *Suma*
- *Dispersión*: desviación típica, varianza, amplitud, mínimo, máximo y E.T. media.
- *Distribución*: Curtosis y Asimetría.
- *Orden de visualización*: Lista de variables, alfabético, medias ascendentes, medias descendentes.

1.7. Análisis de variables categóricas: Tablas de contingencia

- Tabla de contingencia → *Analizar* → Seleccionar la opción *Tablas de contingencia* → Trasladar una variable categórica a la lista *Filas* y otra a la lista *Columna* → *Aceptar* (se puede seleccionar: Mostrar los gráficos de barras agrupadas o suprimir las tablas).
- *Estadísticos* → del cuadro de diálogo Tablas de contingencia, permite las opciones:
 - Chi-cuadrado
 - Correlaciones

- Datos nominales: Coeficiente de contingencia, Phi y V de Cramer
- Lambda y Coeficiente de incertidumbre
- Datos ordinales: Gamma, d de Somers, Tau-b de Kendall y Tau-c de Kendall
- Nominal por intervalo: Eta
- Kappa
- Riesgo
- McNemar
- Estadísticos de Cochran y de Mantel-Haenszel

llevarlas a la lista de *factores fijos*. ← **ANOVA Factorial**.

- *Analizar* → *Modelo lineal general* → *Univariante* → Seleccionar una variable cuantitativa (intervalo y /o razón) y llevarla al cuadro *dependiente* → Seleccionar dos o más variables y llevarlas a la lista de *factores fijos* → seleccionar la variables o variable al apartado *covariable*. ← **Análisis de Covarianza**.

II. 2. Procedimiento de análisis

II.2.1. Métodos explicativos

1.8. Contrastes sobre medias: Procedimientos Medias y Prueba t

- *Analizar* → *Comparar medias* → *Prueba t para dos muestras independientes* → seleccionar la variable independiente → definir la variable de agrupación (por ejemplo género) → definir grupos.

1.9. Análisis de varianza de un factor: ANOVA de un factor

- *Analizar* → *Comparar medias* → *ANOVA de un factor* → Seleccionamos las variables dependientes → Seleccionamos la variable categórica (nominal u ordinal).

1.10. Análisis de correlación lineal (bivariadas)

- *Analizar* → *Correlaciones* → *Bivariadas* → Seleccionar variables cuantitativas y llevar a la lista de variables (al menos dos) → seleccionar coeficiente de correlación (Pearson) → Marcar correlaciones significativas.

1.11. Análisis de varianza factorial y Análisis de Covarianza: El procedimiento Modelo Lineal general Univariante

- *Analizar* → *Modelo lineal general* → *Univariante* → Seleccionar una variable cuantitativa (intervalo y /o razón) y llevarla al cuadro *dependiente* → Seleccionar dos o más variables y

II.2.2. Métodos descriptivos

1.12. Análisis de conglomerados: Procedimiento K-medias

- *Analizar* → Clasificar → Conglomerados de K medias → Seleccionar Variables numéricas y trasladarlas a la lista de variables → Opcionalmente seleccionar una variable para identificar los casos en las tablas de resultados y en los gráficos y trasladarla a la lista de Etiquetar casos → Determinar N° de conglomerados → Método (iterar y clasificar y solo clasificar)

1.13. Análisis factorial

- *Analizar* → Reducción de datos → Análisis factorial → Seleccionar el conjunto de variables que se desea analizar y trasladarlas a la lista de variables → Pulsar *Descriptivos*: Estadísticos (descriptivos univariados) y Matriz de correlaciones (coeficientes, nivel de significación, determinante, inversa, reproducida, anti – imagen, KMO y prueba de esfericidad de Barlett) → *Extracción*: Seleccionar Método (Componentes principales, Mínimos cuadrados ponderados, Máxima verosimilitud, Ejes principales, Alfa e Imagen), *Analizar* (Matriz de correlaciones o Matriz de covarianza), *Extraer* (Autovalores mayores que, Número de factores) y *Mostrar* (solución factorial sin rotar y gráfico de sedimentación), N° de iteraciones por convergencia → Rotación: Seleccionar método (Ninguno, Varimax, Equamax, Oblimin directo y Promax) y *mostrar*

(solución rotada, gráfico de saturaciones)
 → *Opciones*: Valores perdidos y (Excluir casos según lista, Excluir casos según pareja o reemplazar por la media)
 Formato de visualización de los coeficientes (ordenados por tamaño y suprimir valores absolutos menores que).

II.2.3. Otros análisis

1.14. Análisis de variables de respuesta múltiple

La expresión variables de respuesta múltiple se utiliza para identificar variables en las que los sujetos pueden dar más de una respuesta, es decir, variables en las que un mismo sujeto puede tener valores distintos.

- Crear tantas variables dicotómicas como alternativas de respuesta tiene la pregunta → En el *editor de datos* se recoge en primer lugar las variables dicotómicas en las que el valor 1 indica que sí es necesario ese tipo de orientación y el valor 0 que no es necesario (método de dicotomías múltiples).

→ Las últimas tres columnas nos ofrecen la misma información que las cuatro variables dicotómicas, pero en formato de categorías múltiples. Así, el primer sujeto ha marcado dos categorías: académica y profesional, luego en la variable *Resp 1* tiene un 1 (código correspondiente a académica) y en la variable *Resp 2* tiene un 3 (código correspondiente a profesional); y como ya no ha marcado ninguna respuesta más, en la variable *Resp3* tiene un 0 (que funciona como código de valor perdido).

- *Cálculo de frecuencias* → Analizar → Estadísticos descriptivos → Frecuencias → Seleccionar las tres variables dicotómicas.
- *Definir conjuntos de respuestas múltiples* → Analizar → Respuestas múltiples → Definir conjuntos de respuestas múltiples → seleccionamos las variables y las trasladamos a la lista de *Variables del conjunto* → Seleccionamos opción: *dicotómicas* con el valor contado o categorías, seleccionamos el rango → Asignamos *nombre y etiqueta* al nuevo conjunto → *Añadir* al conjunto de

respuestas múltiples. A partir de este momento el conjunto creado podrá utilizarse para obtener tablas de frecuencias y tablas de contingencia.

- Tablas de frecuencia → Analizar → Respuestas múltiples → Frecuencias de respuestas múltiples → Seleccionamos el conjunto de respuestas múltiples que deseamos describir y lo trasladamos a la lista *Tablas para* → Seleccionamos los valores perdidos, excluir casos según lista dentro de las dicotomías.
- Tablas de contingencia → Analizar → Respuestas múltiples → Tablas de contingencia de respuesta múltiple → seleccionamos variable o variables y la trasladamos a la lista de *filas* → seleccionamos un segundo conjunto de variables o variable y la trasladamos a la lista de *columnas* → pulsamos *opciones* → Porcentaje de casilla (total) y porcentaje basado en *casos*.

4. Referencias bibliografía y bibliografía consultada

- Álvarez, V., García, E., Gil, J., Martín, P., Romero, S. & Rodríguez, J. (2002). *Diseño y evaluación de programas*. Madrid: EOS.
- Bisquerra Alzina, R. (1989). *Introducción conceptual al análisis multivariable. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD*. Barcelona: PPU.
- Buendía, L., Colás, P. & Hernández, F. (1997). *Métodos de investigación en psicopedagogía*. Madrid: Mac Graw Hill.
- Gardner, C. R. (2003). *Estadística para psicología usando SPSS para Windows*. México: Prentice Hall.
- González – Conde Llopis, C. (2000). *Estadística aplicada con Excel 97*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Pardo, A. & Ruíz, M.A. (2002). *SPSS 11. Guía para el análisis de datos*. Madrid: McGraw Hill.