

SPSS: DESCRIPTIVOS

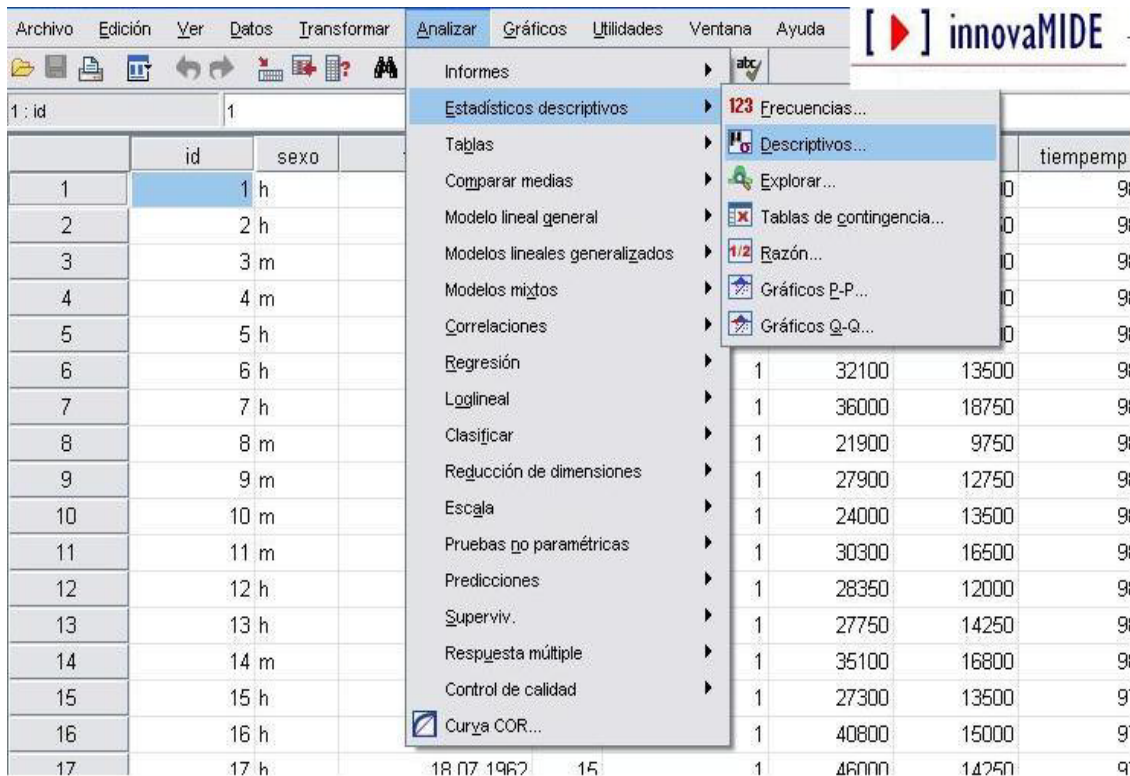
PROCEDIMIENTO DE ANÁLISIS INICIAL DE DATOS: DESCRIPTIVOS

A diferencia 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. Contiene algunos estadísticos descriptivos (tendencia central, dispersión y forma de la distribución) que también incluye el procedimiento Frecuencias, pero añade una opción: la posibilidad de obtener puntuaciones típicas. Para obtener estadísticos descriptivos y puntuaciones típicas:

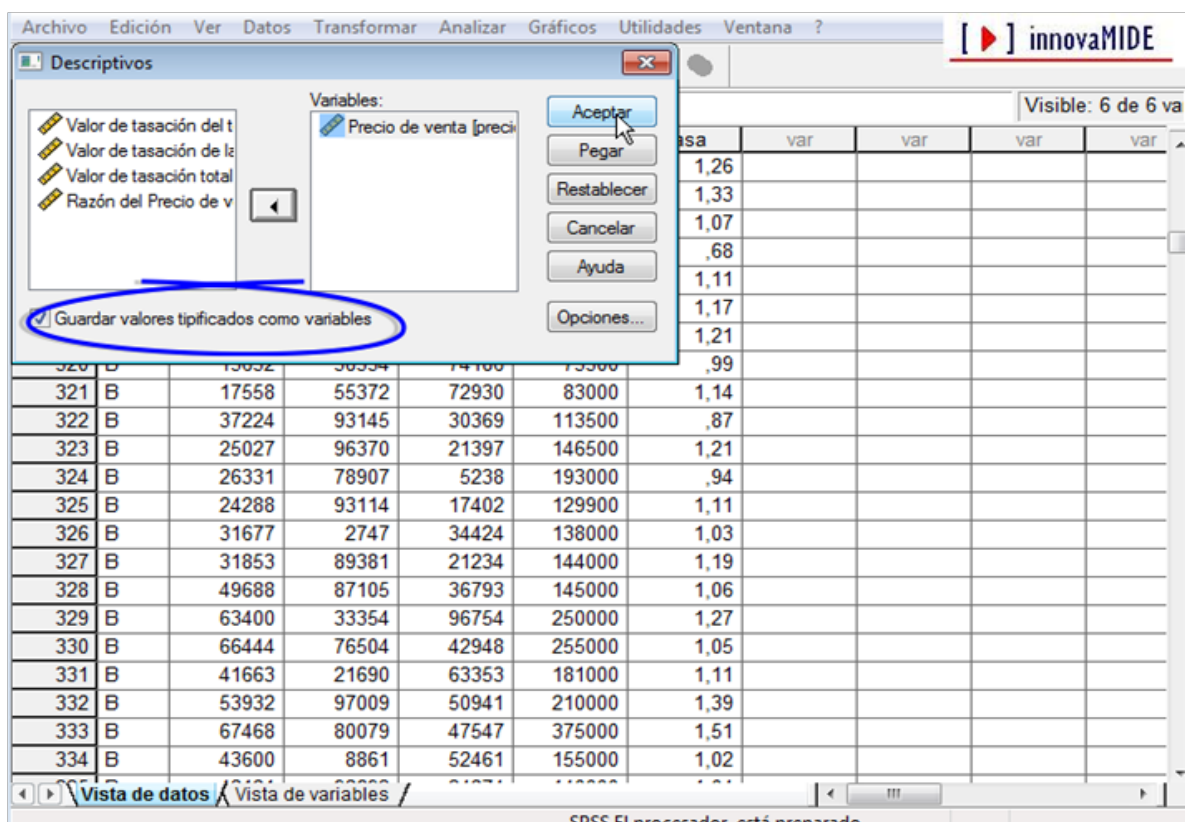
Menú:

- Analizar
 - Estadísticos Descriptivos

Y accedemos al cuadro de diálogo Descriptivos...



Seleccionamos variable/s que queremos observar, es necesario que la variable sea numérica, las variables de cadena no están disponibles en la lista de variables... Aceptar.



Opción: **Guardar valores tipificados como variables**. Los valores tipificados, llamados puntuaciones típicas, expresan el número de desviaciones típicas que cada valor se aleja de su media. Marcando la opción, el SPSS crea en el archivo de datos del Editor de datos una nueva variable con las puntuaciones típicas correspondientes a cada caso. La nueva variable recibe, por defecto, el nombre de la variable original con el prefijo z. Vemos el ejemplo a continuación:...

Vista de datos:

	barrio	valterr	valmejor	valtot	precio	tasa	Zprecio	var	var
2383	G	8316	13967	22283	35000	1,57	-,80670		
2384	G	8148	9384	17532	49000	2,79	-,52250		
2385	G	8148	11959	20107	15000	,75	-1,21270		
2386	G	9506	24079	33585	52000	1,55	-,46160		
2387	G	8148	17711	25859	32500	1,26	-,85745		
2388	G	8736	16551	25287	35000	1,38	-,80670		
2389	G	8165	15022	23187	39800	1,72	-,70926		
2390	G	5865	13678	19543	35550	1,82	-,79553		
2391	G	5865	17210	23075	44000	1,91	-,62400		
2392	G	7337	27612	34949	42000	1,20	-,66460		
2393	G	6937	9555	16492	11000	,67	-1,29390		
2394	G	5452	9705	15157	32000	2,11	-,86760		
2395	G	5452	7268	12720	28000	2,20	-,94880		
2396	G	6232	11757	17989	44000	2,45	-,62400		
2397	G	12953	21880	34833	54500	1,56	-,41085		
2398	G	6232	12780	19012	20000	1,05	-1,11120		
2399	G	11242	25812	37054	47000	1,27	-,56310		
2400	G	6232	11575	17807	20000	1,12	-1,11120		
2401	G	6232	5591	11823	21000	1,78	-1,09090		
2402	G	13229	22388	35617	32300	,91	-,86151		
2403	G	6365	16545	22910	20000	,87	-1,11120		
2404	G	6232	11100	17332	23200	1,34	-1,04624		

Vista de variables:

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas
1	barrio	Cadena	1	0	Barrio	{A, A}...	Ninguno	5
2	valterr	Numérico	8	0	Valor de tasac	Ninguno	Ninguno	8
3	valmejor	Numérico	9	0	Valor de tasac	Ninguno	Ninguno	8
4	valtot	Numérico	9	0	Valor de tasac	Ninguno	Ninguno	8
5	precio	Numérico	9	0	Precio de vent	Ninguno	Ninguno	8
6	tasa	Numérico	4	2	Razón del Pre	Ninguno	Ninguno	8
7	Zprecio	Numérico	11	5	Puntua: Preci	Ninguno	Ninguno	13
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								

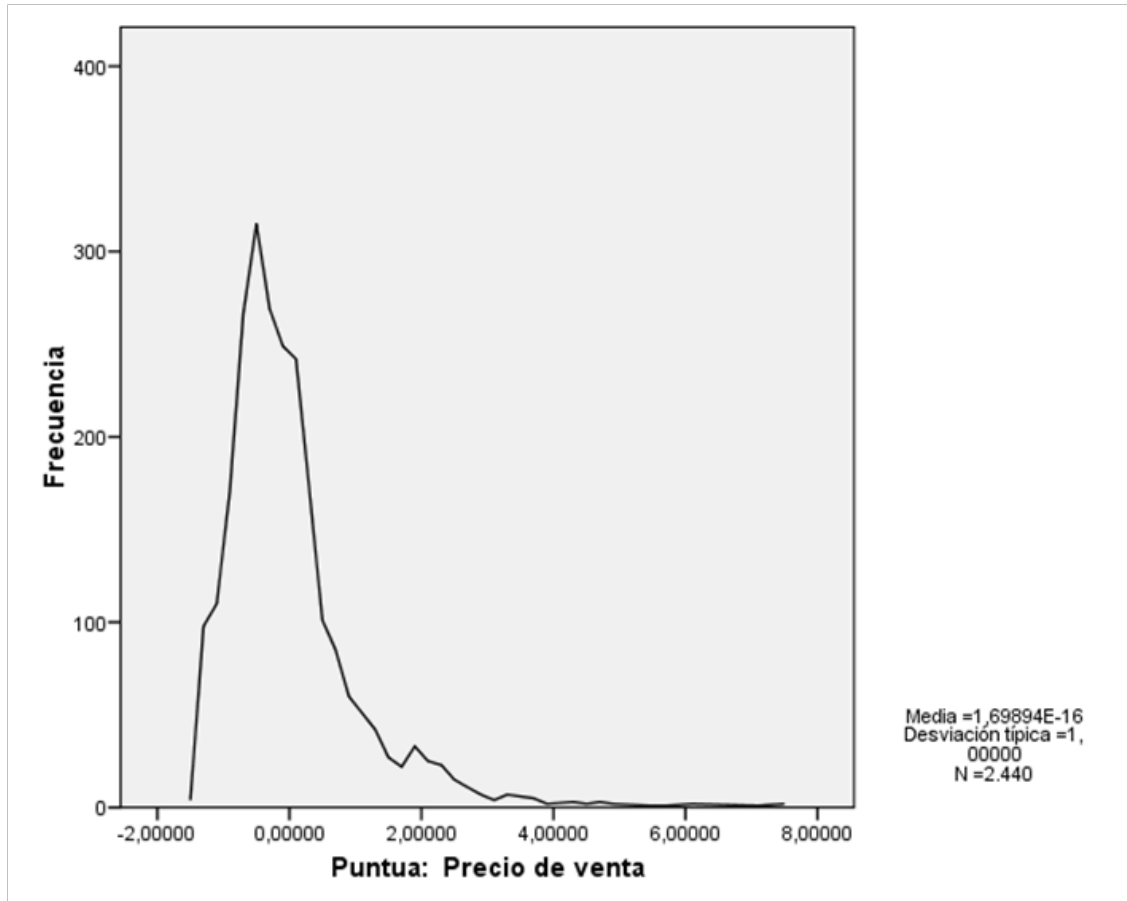
Descriptivos

El procedimiento Descriptivos muestra estadísticos de resumen univariados para varias variables en una única tabla y calcula valores tipificados (puntuaciones z). Las variables se pueden ordenar por el tamaño de sus medias (en orden ascendente o descendente), alfabéticamente o por el orden en el que se seleccionen las variables (el valor por defecto).

Cuando se guardan las puntuaciones z, éstas se añaden a los datos del Editor de datos y quedan disponibles para los gráficos, el listado de los datos y los análisis. Cuando las variables se registran en unidades diferentes (por ejemplo, producto interior bruto per cápita y porcentaje de alfabetización), una transformación de puntuación z pondrá las variables en una escala común para poder compararlas visualmente con más facilidad.

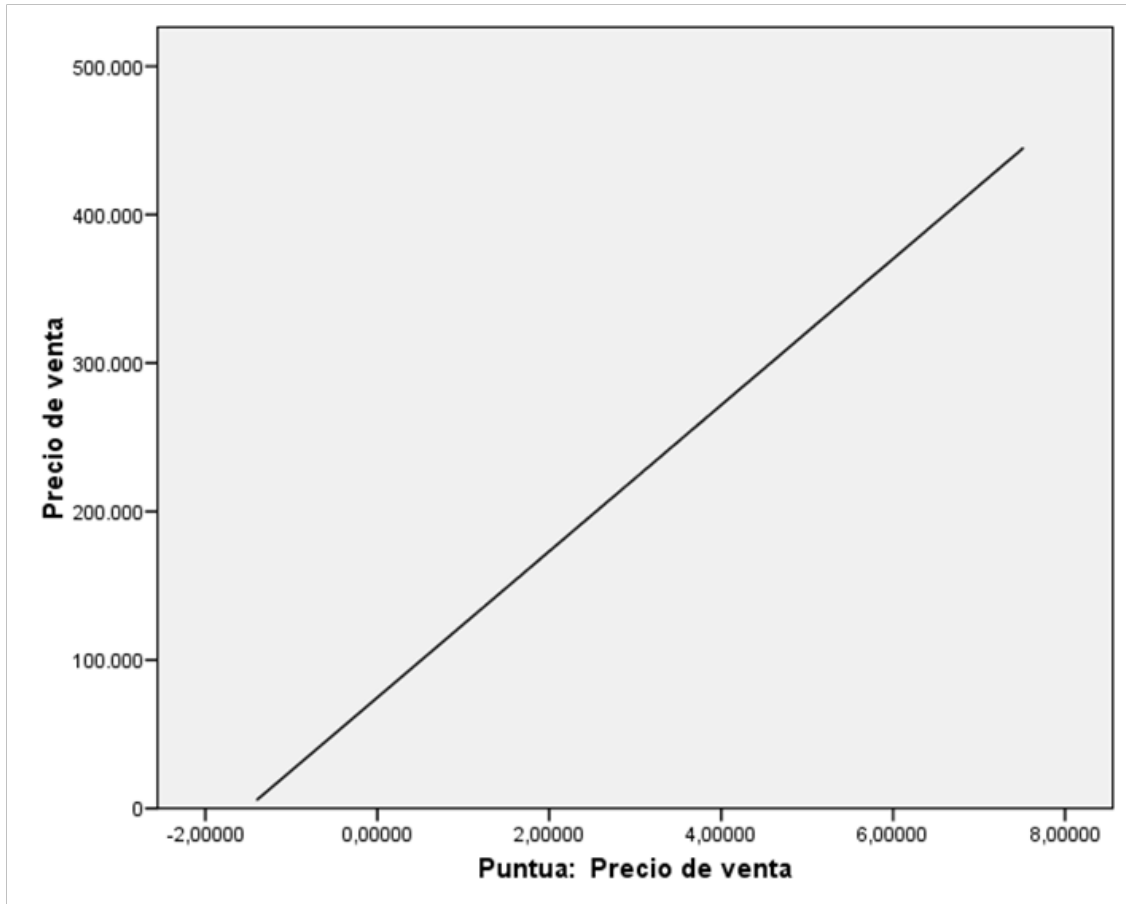
Siguiendo con el mismo ejemplo:

Puntuaciones tipificadas para el precio de venta de las viviendas en su representación gráfica tendría esta imagen:



Puntuá: Precio de venta en este caso son todas las puntuaciones tipificadas para cada una de los valores de precio de venta de casas, si hacemos el gráfico de las puntuaciones tipificadas, vemos que la frecuencia de las puntuaciones tipificadas cercanas al valor 0 es muy alta, sin embargo, las puntuaciones tipificadas altas tienen muy poca frecuencia;

Si creamos el gráfico que refleja la dependencia de las puntuaciones tipificadas del precio de venta de la variable precio de venta, observamos siguiente imagen:



El gráfico muestra que a mayor precio de venta la puntuación tipificada aumenta, es decir, a mayor precio de vivienda, la desviación típica aumenta.

Ejemplo.

Si cada caso de los datos contiene los totales de ventas diarias de cada vendedor (por ejemplo, una entrada para Javier, una para Carla y una para Francisco) recogidas cada día durante varios meses, el procedimiento Descriptivos puede calcular la media diaria de ventas para cada vendedor y ordenar los resultados del promedio de ventas de mayor a menor.

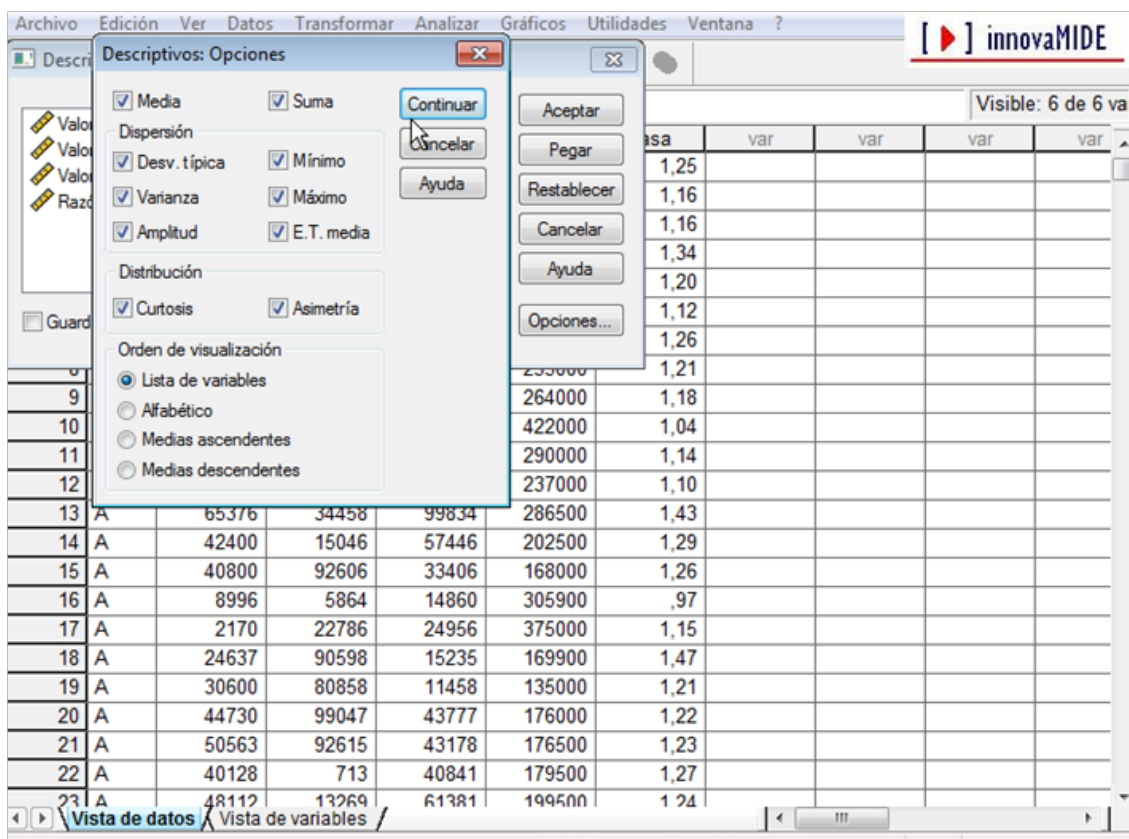
Estadísticos. Tamaño de muestra, media, mínimo, máximo, desviación típica, varianza, rango, suma, error típico de la media, curtosis y asimetría con sus errores típicos.

Datos. Utilice variables numéricas después de haberlas inspeccionado gráficamente para registrar errores, valores atípicos y anomalías de distribución. El procedimiento Descriptivos es muy eficaz para archivos grandes (de miles de casos).

Supuestos.

La mayoría de los estadísticos disponibles (incluyendo las puntuaciones z) se basan en la teoría normal y son adecuados para variables cuantitativas (medidas a nivel de razón o de intervalo) con distribuciones simétricas. Se deben evitar las variables con categorías no ordenadas o distribuciones asimétricas. La distribución de puntuaciones z tiene la misma forma que la de los datos originales; por tanto, el cálculo de puntuaciones z no es una solución para los datos con problemas.

Descriptivos: Opciones



Medidas de tendencia central.

- **Media y suma.** Se muestra por defecto la media o promedio aritmético.

Dispersión. Los estadísticos que miden la dispersión o variación en los datos incluyen la desviación típica, la varianza, el rango, el mínimo, el máximo y el error típico de la media.

- **Desviación típica.** Medida de dispersión en torno a la media. En una distribución normal, el 68% de los casos se encuentra dentro de una desviación típica respecto a la media y el 95% de los casos se encuentra dentro de dos desviaciones típicas respecto a la media. Por ejemplo, si la media de edad es 45, con una desviación típica de 10, el 95% de los casos estaría entre 25 y 65 en una distribución normal.
- **Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.
- **Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.
- **Mínimo.** Valor más pequeño de una variable numérica.
- **Máximo.** El mayor valor de una variable numérica.
- **E. T. media.** Medida de cuánto puede variar el valor de la media de una muestra a otra, extraídas éstas de la misma distribución. Puede utilizarse para comparar de forma aproximada la media observada con un valor hipotetizado (es decir, podremos concluir que dos valores son distintos si la razón de la diferencia respecto al error típico es menor que -2 o mayor que +2).

Distribución. La curtosis y la asimetría son los estadísticos que caracterizan la forma y simetría de la distribución. Estos estadísticos se muestran con sus errores típicos.

- **Curtosis.** Medida del grado en que las observaciones están agrupadas en torno al punto central. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.
- **Asimetría.** Medida de la asimetría de una distribución. La distribución normal es simétrica y tiene un valor de asimetría igual a 0. Una distribución que tenga una asimetría positiva significativa tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Como regla aproximada, un valor de la asimetría mayor que el doble de su error típico se asume que indica una desviación de la simetría.

Orden de presentación. Por defecto, las variables se muestran en el orden en que se hayan seleccionado. Si lo desea, se pueden mostrar las variables alfabéticamente, por medias ascendentes o por medias descendentes.

Elija en los menús:

- Analizar
 - Estadísticos descriptivos
 - Descriptivos...

En el cuadro de diálogo Descriptivos, pulse en Opciones.

- Seleccione uno o más estadísticos.

En el caso de nuestro ejemplo la tabla de resultados presenta siguientes datos para la variable Precio de Venta:

El Rango (la amplitud) es de 439500 puntos

El mínimo: 5500

El máximo: 445000

Suma: 182362033

La Media se sitúa en 74738,54 puntos

La Desviación Típica es de 49260,688 puntos

Varianza es de 2,4E+0009

La Asimetría es de 2,114 (con su Error típico de 0,05)

La Curtosis es de 7,64 (con su Error típico de 0.099)

Descriptivos: Procedimientos relacionados

Hay que utilizar el procedimiento Frecuencias para obtener la mediana, la moda, los cuartiles, los percentiles y un histograma. Para calcular estadísticos de resumen para cada uno de los distintos grupos de casos (por ejemplo, si quiere separar los estadísticos para hombres y mujeres o personas que viven en cuatro regiones diferentes de un país), utilice el procedimiento Explorar o Medias. También se puede utilizar Segmentar archivo en el menú Datos.

El material elaborado fue adaptado de Tutorial SPSS15.