

## PRÁCTICA 4

# ESTADÍSTICA

## Practica 4

JAVIER RIVILLA ARREDONDO

53247378D



# ESTADÍSTICA

## EJERCICIOS

1. La sección de nóminas de una empresa dispone de la siguiente información sobre salarios mensuales en euros de sus trabajadores:

Salarios mensuales € - Nº trabajadores

1000 - 1500 80  
1500 - 2000 20  
2000 - 2500 10  
2500 - 3000 10  
3000 - 3500 6

Calcular:

- a) El salario mensual medio de los trabajadores de la empresa.  
b) La desviación típica y la mediana de la distribución de salarios.  
c) ¿Qué salario mensual es superado por el 25% de los trabajadores?
- a) El salario mensual medio de los trabajadores de la empresa.  
[1000-1500[-> valor medio = 1250  
[1500-2000[->valor medio = 1750  
[2000-2500[-> valor medio = 2250  
[2500-3000[-> valor medio = 2750  
[3000-3500[->valor medio = 3250

En la siguiente imagen se mostrará los datos anteriores en el programa:

|   | salarios | num_trabajadores |
|---|----------|------------------|
| 1 | 1250     | 80               |
| 2 | 1750     | 20               |
| 3 | 2250     | 10               |
| 4 | 2750     | 10               |
| 5 | 3250     | 6                |

Ponderamos, “Datos” -> “Ponderar Casos”.

Una vez ponderado, vamos los estadísticos descriptivos.

Analizar->Estadísticos descriptivos->Frecuencias->Estadísticos->Media

### Estadísticos

salarios

|       |          |         |
|-------|----------|---------|
| N     | Válido   | 126     |
|       | Perdidos | 0       |
| Media |          | 1623,02 |

### salarios

|             | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|-------------|------------|------------|-------------------|----------------------|
| Válido 1250 | 80         | 63,5       | 63,5              | 63,5                 |
| 1750        | 20         | 15,9       | 15,9              | 79,4                 |
| 2250        | 10         | 7,9        | 7,9               | 87,3                 |
| 2750        | 10         | 7,9        | 7,9               | 95,2                 |
| 3250        | 6          | 4,8        | 4,8               | 100,0                |
| Total       | 126        | 100,0      | 100,0             |                      |

Al ponderar por número de trabajadores vemos que tenemos 126 valores y la media es de 1623,02 trabajadores.

- b) La desviación típica y la mediana de la distribución de salarios.

Nos vamos a “Analizar”->”Estadísticos descriptivos”->”Frecuencias”->”Estadísticos” y seleccionamos Mediana y desviación estándar.

### Estadísticos

salarios

|                     |          |         |
|---------------------|----------|---------|
| N                   | Válido   | 126     |
|                     | Perdidos | 0       |
| Mediana             |          | 1250,00 |
| Desviación estándar |          | 593,082 |



- c) ¿Qué salario mensual es superado por el 25% de los trabajadores?

Para calcularlo iremos a:

“Analizar”->”Estadísticos descriptivos”->”Frecuencias”->”Estadísticos”  
Introducimos en percentiles 75.

**Frecuencias: Estadísticos**

**Valores percentiles**

☐ Cuartiles

☐ Puntos de corte para: 10 grupos iguales

☒ Percentiles:

**Tendencia central**

☐ Media

☐ Mediana

☐ Moda

☐ Suma

☐ Los valores son puntos medios de grupos

**Dispersión**

☐ Desviación estándar

☐ Varianza

☐ Rango

☐ Mínimo

☐ Máximo

☐ Error estándar media

**Distribución**

☐ Asimetría

☐ Curtosis

### Estadísticos

salarios

|             |          |         |
|-------------|----------|---------|
| N           | Válido   | 126     |
|             | Perdidos | 0       |
| Percentiles | 75       | 1750,00 |

2. Se ha estudiado el número de accesos recibidos al día por un servidor FTP en un cuatrimestre. (El fichero datos-pr4-ejer2\_2015.sav con el número de accesos se puede encontrar en campus virtual).

Calcular:

- Agrupar los datos en intervalos de la misma amplitud.
- Número medio de accesos diarios.
- Desviación típica, moda y mediana.
- Primer cuartil.

- Agrupar los datos en intervalos de la misma amplitud.

Nos vamos a "Analizar" -> "Estadísticos descriptivos" -> "Descriptivos".

### Estadísticos descriptivos

|                      | N   | Mínimo | Máximo | Media | Desviación estándar |
|----------------------|-----|--------|--------|-------|---------------------|
| Accesos              | 122 | 1      | 25     | 14,12 | 4,040               |
| N válido (por lista) | 122 |        |        |       |                     |

Agrupamos los intervalos de longitud 5.

[1-5[->valor medio = 2,5  
[5-10[->valor medio = 7,5  
[10-15[->valor medio = 12,5  
[15-20[->valor medio = 17,5  
[20-25[->valor medio = 22,5

El último intervalo será cerrado.

Una vez tenemos los intervalos recodificamos en una variable distinta llamada cod\_accesos.

**Recodificar en distintas variables**

Variable numérica -> Variable de resultado:  
Accesos -> cod\_accesos

Variable de resultado  
Nombre: cod\_accesos  
Etiqueta:  
Cambiar

Valores antiguos y nuevos...

Si la opción... (condición de selección de casos opcional)

Aceptar Pegar Restablecer Cancelar Ayuda

**Recodificar en distintas variables: Valores antiguos y nuevos**

Valor antiguo

- ☐ Valor:
- ☐ Perdido del sistema
- ☐ Perdido del sistema o perdido del usuario
- ☒ Rango:  
hasta  
Rango, INFERIOR hasta valor:  
Rango, valor hasta SUPERIOR:  
☐ Todos los demás valores

Valor nuevo

- ☒ Valor:
- ☐ Perdido del sistema
- ☐ Copiar valores antiguos

Antiguo -> Nuevo:

|               |         |
|---------------|---------|
| 1 thru 4,9    | -> 2,5  |
| 5 thru 9,9    | -> 7,5  |
| 10 thru 14,99 | -> 12,5 |
| 15 thru 19,9  | -> 17,5 |
| 20 thru 25    | -> 22,5 |

Añadir Cambiar Eliminar

☐ Las variables de resultado son cadenas Anchura: 8

☐ Convertir cadenas numéricas en números ('5' -> 5)

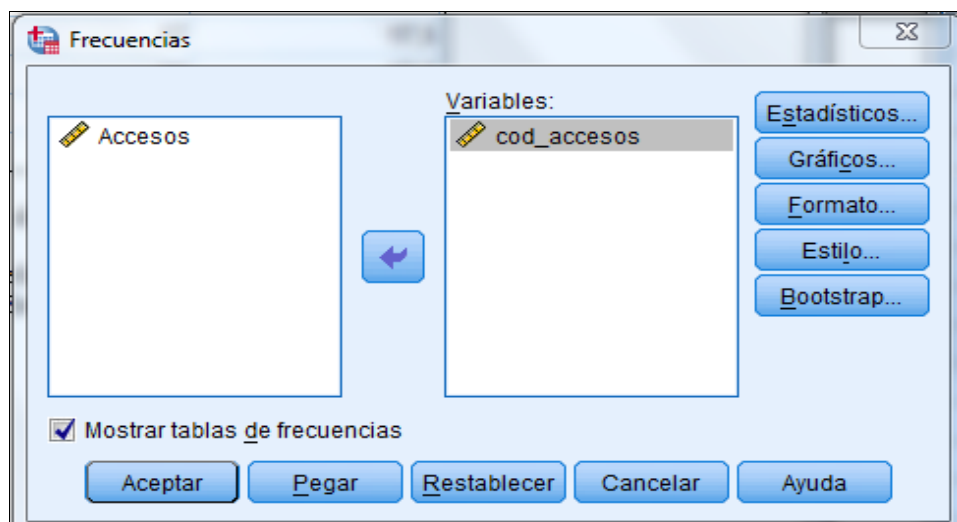
Continuar Cancelar Ayuda

Tras esto tendremos los siguientes datos:

|    | Accesos | cod_accesos |
|----|---------|-------------|
| 1  | 10      | 12,5        |
| 2  | 17      | 17,5        |
| 3  | 15      | 17,5        |
| 4  | 8       | 7,5         |
| 5  | 1       | 2,5         |
| 6  | 16      | 17,5        |
| 7  | 15      | 17,5        |
| 8  | 18      | 17,5        |
| 9  | 15      | 17,5        |
| 10 | 11      | 12,5        |
| 11 | 7       | 7,5         |
| 12 | 12      | 12,5        |
| 13 | 12      | 12,5        |
| 14 | 16      | 17,5        |
| 15 | 19      | 17,5        |
| 16 | 17      | 17,5        |
| 17 | 12      | 12,5        |
| 18 | 12      | 12,5        |
| 19 | 2       | 2,5         |
| 20 | 15      | 17,5        |
| 21 | 17      | 17,5        |
| 22 | 11      | 12,5        |
| 23 | 17      | 17,5        |
| 24 | 4       | 2,5         |

b) Número medio de accesos diarios.

La variable que hemos recodificado (cod\_accesos) haremos lo siguiente:  
 “Analizar”->”Estadísticos descriptivos”->”Frecuencias”->”Estadísticos”->”Media”.



### Estadísticos

cod\_accesos

|       |          |        |
|-------|----------|--------|
| N     | Válido   | 122    |
|       | Perdidos | 0      |
| Media |          | 14,549 |

- c) Desviación típica, moda y mediana.

Haremos lo mismo que antes, solo que marcaremos lo que nos piden:

**Frecuencias: Estadísticos**

**Valores percentiles**

- ☐ Cuartiles
- ☒ Puntos de corte para: 10 grupos iguales
- ☐ Percentiles:
- Añadir
- Cambiar
- Eliminar

**Tendencia central**

- ☐ Media
- ☒ Mediana
- ☒ Moda
- ☐ Suma
- ☐ Los valores son puntos medios de grupos

**Dispersión**

- ☒ Desviación estándar
- ☐ Varianza
- ☐ Rango
- ☐ Mínimo
- ☐ Máximo
- ☐ Error estándar media

**Distribución**

- ☐ Asimetría
- ☐ Curtosis

Continuar Cancelar Ayuda

### Estadísticos

cod\_accesos

|                     |          |        |
|---------------------|----------|--------|
| N                   | Válido   | 122    |
|                     | Perdidos | 0      |
| Mediana             |          | 15,000 |
| Moda                |          | 17,5   |
| Desviación estándar |          | 4,3482 |

- d) Primer cuartil.

Los tramos se dividen en 25, siendo el primer cuartil 25% por lo tanto tendremos que poner el percentil 25.

“Analizar”->”Estadísticos descriptivos”->”Frecuencias”->”Estadísticos”->”Percentiles”.

**Frecuencias: Estadísticos**

**Valores percentiles**

☐ Cuartiles

☐ Puntos de corte para: 10 grupos iguales

☒ Percentiles:

Añadir 25,0

Cambiar

Eliminar

**Tendencia central**

☐ Media

☐ Mediana

☐ Moda

☐ Suma

☐ Los valores son puntos medios de grupos

**Dispersión**

☐ Desviación estándar ☐ Mínimo

☐ Varianza ☐ Máximo

☐ Rango ☐ Error estándar media

**Distribución**

☐ Asimetría

☐ Curtosis

Continuar Cancelar Ayuda

### Estadísticos

cod\_accesos

|             |          |        |
|-------------|----------|--------|
| N           | Válido   | 122    |
|             | Perdidos | 0      |
| Percentiles | 25       | 12,500 |

3. Durante una semana se registra el número de accesos (peticiones http) a los 1530 sitios almacenados en un servidor web. Es necesario analizar las diferencias de carga entre unos sitios y otros para mejorar el rendimiento general del servidor. (El fichero pr4-ejer3\_2015 con el número de accesos se puede encontrar en campus virtual). Contesta a las siguientes preguntas basándote en el cálculo de percentiles.

- Se estima que el 20 por ciento de sitios recibe menos de .....visitas al día.
- Se estima que el 40 por ciento de sitios recibe al menos .....visitas al día.
- Explica la resolución del ejercicio.

- Se estima que el 20 por ciento de sitios recibe menos de ....visitas al día.

Haremos el percentil de 20 porque es el que cumple que deja un 80% de valores por encima de este.



**Frecuencias: Estadísticos**

**Valores percentiles**

☐ Cuartiles

☐ Puntos de corte para: 10 grupos iguales

☒ Percentiles: 20,0

Añadir Cambiar Eliminar

**Tendencia central**

☐ Media

☐ Mediana

☐ Moda

☐ Suma

☐ Los valores son puntos medios de grupos

**Dispersión**

☐ Desviación estándar ☐ Mínimo

☐ Varianza ☐ Máximo

☐ Rango ☐ Error estándar media

**Distribución**

☐ Asimetría

☐ Curtosis

Continuar Cancelar Ayuda

| Estadísticos |          |        |
|--------------|----------|--------|
| Accesos      |          |        |
| N            | Válido   | 1530   |
|              | Perdidos | 0      |
| Media        |          | 546,21 |
| Percentiles  | 20       | 135,20 |

Ahora para saber el número de visitas multiplicaremos percentil por media:  
 Esto nos dará: 73847,592 visitas.

Lo que deduciremos que se estima un 20 por ciento de sitios, que recibe menos del 135,20 visitas al día.

- b) Se estima que el 40 por ciento de sitios recibe al menos ....visitas al día.

Realizaremos el mismo procedimiento que el anterior, solo que en este caso es el percentil 60.

**Frecuencias: Estadísticos**

**Valores percentiles**

☐ Cuartiles

☐ Puntos de corte para: 10 grupos iguales

☒ Percentiles:

**Tendencia central**

☒ Media

☐ Mediana

☐ Moda

☒ Suma

☐ Los valores son puntos medios de grupos

**Dispersión**

☐ Desviación estándar

☐ Varianza

☐ Rango

☐ Mínimo

☐ Máximo

☐ Error estándar media

**Distribución**

☐ Asimetría

☐ Curtosis

### Estadísticos

#### Accesos

|             |          |        |
|-------------|----------|--------|
| N           | Válido   | 1530   |
|             | Perdidos | 0      |
| Media       |          | 546,21 |
| Suma        |          | 835705 |
| Percentiles | 60       | 618,00 |

Como vemos el percentil de 40 será 366. El número de visitas será SUMA-(percentil\*MEDIA) = 498147,22

Por lo que se estima que el 40 por ciento de sitios recibe al menos 498147,22 visitas al día.

- c) Explica la resolución del ejercicio.

En el apartado a usamos el percentil 20 obteniendo 135,20 también tenemos que hallar la media para aplicar la fórmula, algo claro. Y ya sacamos las visitas.

En cuanto al apartado b, sacamos el percentil de 20 obtenemos 366, hallaremos suma y media , también para aplicar determinada fórmula.

4. Se ha realizado un estudio de seguridad sobre 230 redes inalámbricas. El número de paquetes ARP necesarios para averiguar la clave de cada red viene dado por la siguiente tabla:

| Paquetes(miles) | 0-50 | 50-100 | 100-150 | 150-200 | 200-250 | 250-300 | 300-350 | 350-400 |
|-----------------|------|--------|---------|---------|---------|---------|---------|---------|
| Num Redes       | 14   | 28     | 42      | 57      | 70      | 12      | 5       | 2       |

- a) Construir la tabla de frecuencias.  
b) Representar la distribución mediante un diagrama de barras y el polígono de frecuencias.

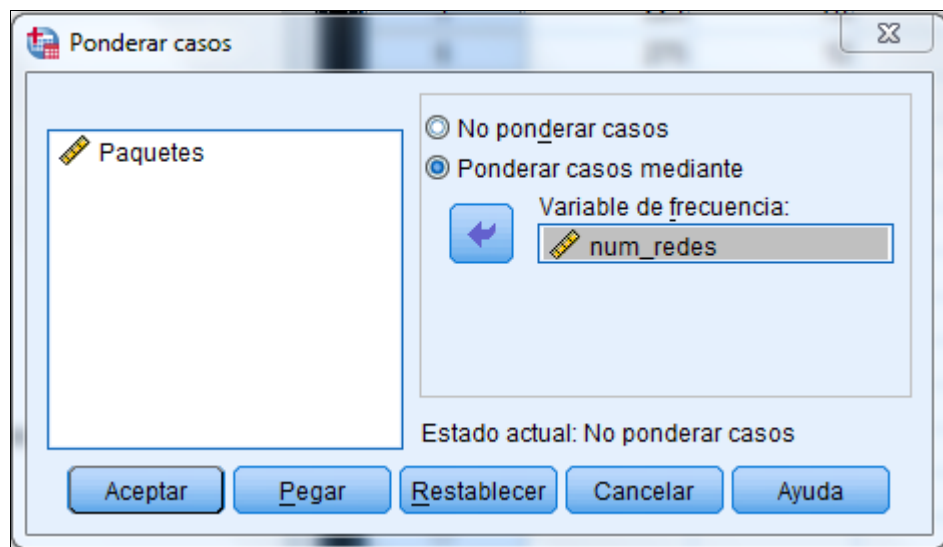
- c) Calcular la media, mediana, desviación típica, moda y percentil 20.  
d) ¿Por debajo de qué número de paquetes se encuentran el 75% de la muestra?

- a) Construir la tabla de frecuencias.

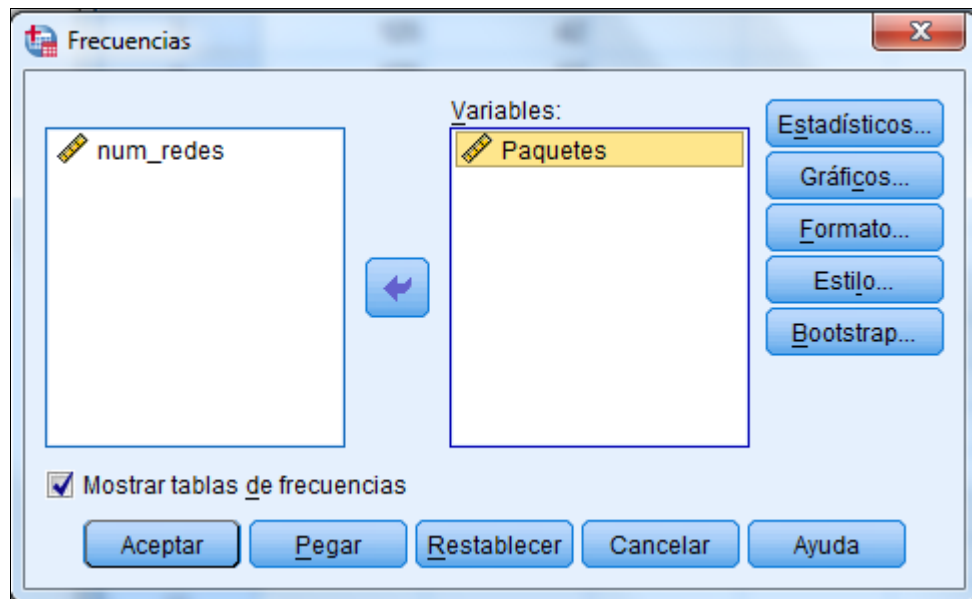
Para empezar:

[0-50[->valor medio=25.  
[50-100[->valor medio=75.  
[100-150[->valor medio=125.  
[150-200[->valor medio=175.  
[200-250[->valor medio=225.  
[250-300[->valor medio=275.  
[300-350[->valor medio=325.  
[350-400 [->valor medio=375.

Una vez tengamos definidos los intervalos con los valores medios, introducimos los datos y ponderamos.



Una vez ponderamos nos vamos a “Analizar”->”Estadísticos descriptivos”->”Frecuencias”.



#### Estadísticos

Paquetes

|   |          |     |
|---|----------|-----|
| N | Válido   | 230 |
|   | Perdidos | 0   |

#### Paquetes

|        |       | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|-------|------------|------------|-------------------|----------------------|
| Válido | 25    | 14         | 6,1        | 6,1               | 6,1                  |
|        | 75    | 28         | 12,2       | 12,2              | 18,3                 |
|        | 125   | 42         | 18,3       | 18,3              | 36,5                 |
|        | 175   | 57         | 24,8       | 24,8              | 61,3                 |
|        | 225   | 70         | 30,4       | 30,4              | 91,7                 |
|        | 275   | 12         | 5,2        | 5,2               | 97,0                 |
|        | 325   | 5          | 2,2        | 2,2               | 99,1                 |
|        | 375   | 2          | ,9         | ,9                | 100,0                |
|        | Total | 230        | 100,0      | 100,0             |                      |

- b) Representar la distribución mediante un diagrama de barras y el polígono de frecuencias.

“Gráficos”->”cuadro de dialogo antiguo”->”barras”->”Simple”:

Definir barras simples: Resúmenes para grupos de casos

num\_redes

Las barras representan

☒ N.º de casos ☐ % de casos

☐ Número acumulado ☐ % acumulado

☐ Otro estadístico (ej., la mediana)

Variable:

Cambiar estadístico...

Eje de categorías:

Paquetes

Panel mediante

Filas:

Anidar variables (sin filas vacías)

Columnas:

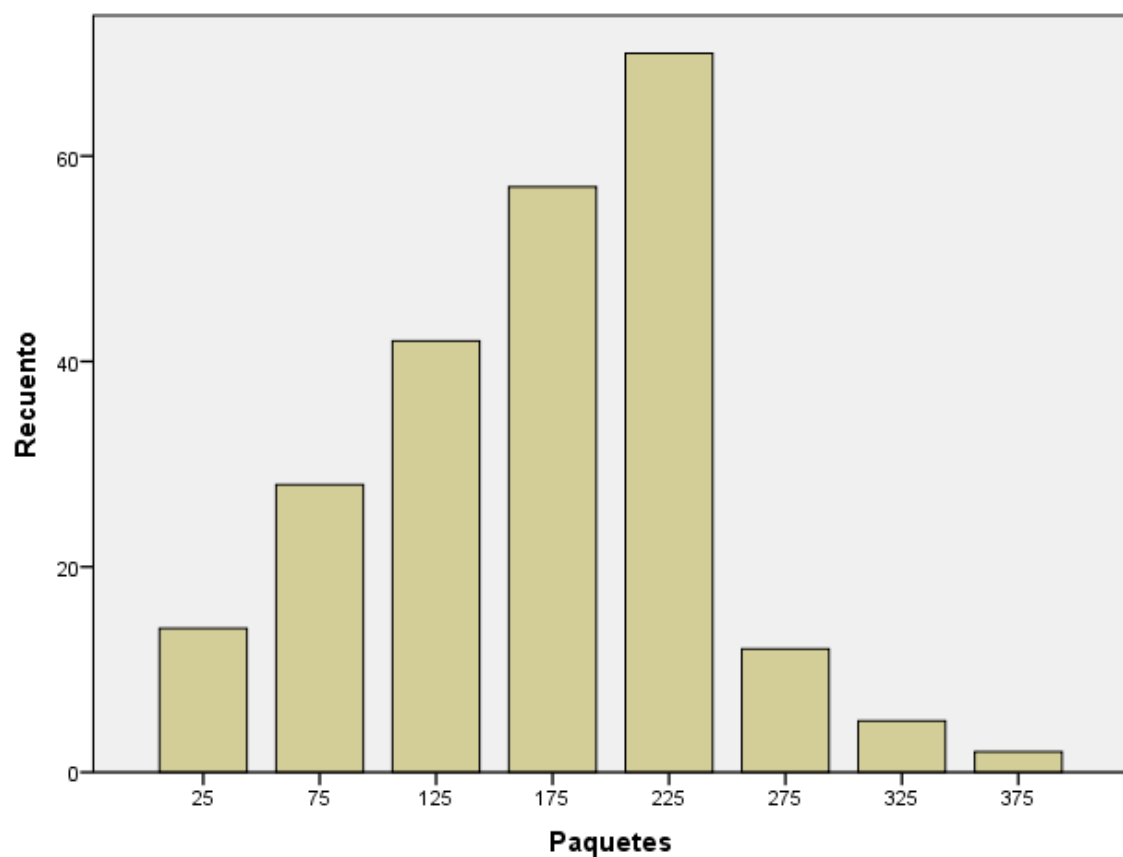
Anidar variables (sin columnas vacías)

Plantilla

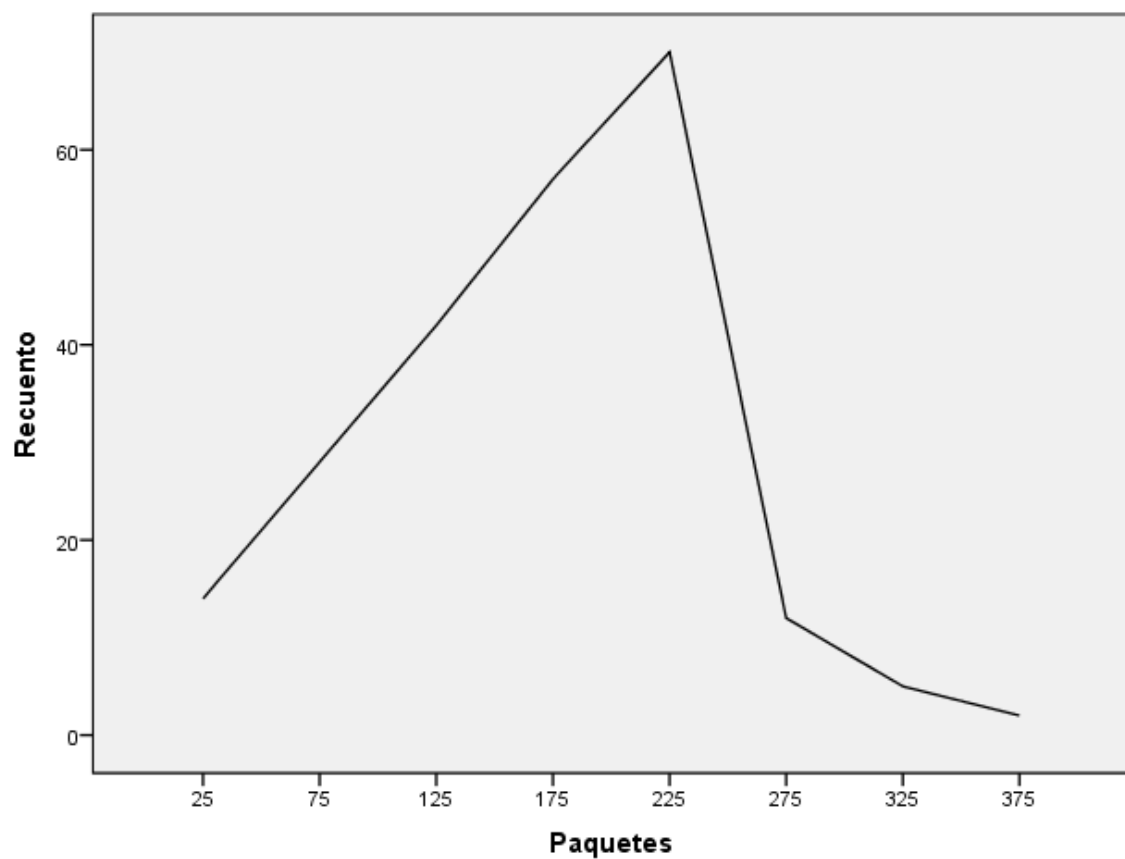
☐ Usar las especificaciones gráficas de:

Archivo...

Aceptar Pegar Restablecer Cancelar Ayuda



Para obtener el polígono de frecuencias iremos a gráficos>cuadro de diálogo antiguo->”Líneas”->”Simple”.



- c) Calcular la media, mediana, desviación típica, moda y percentil 20.

Este apartado es como todos los anteriores, tendremos que hacer lo siguiente:

“Analizar”->”Estadísticos Descriptivos”->”Frecuencias”->”Estadísticos”

Después de esto seleccionamos lo que nos piden.

**Estadísticos**

Paquetes

|                     |          |        |
|---------------------|----------|--------|
| N                   | Válido   | 230    |
|                     | Perdidos | 0      |
| Media               |          | 170,00 |
| Mediana             |          | 175,00 |
| Desviación estándar |          | 71,836 |
| Percentiles         | 20       | 125,00 |

- d) ¿Por debajo de qué número de paquetes se encuentran el 75% de la muestra?

Este caso es como los anteriores, puesto que en ello tendremos que obtener el percentil 75.

**Frecuencias: Estadísticos**

**Valores percentiles**

☐ Cuartiles

☐ Puntos de corte para: 10 grupos iguales

☒ Percentiles:

**Tendencia central**

☐ Media

☐ Mediana

☐ Moda

☐ Suma

☐ Los valores son puntos medios de grupos

**Dispersión**

☐ Desviación estándar ☐ Mínimo

☐ Varianza ☐ Máximo

☐ Rango ☐ Error estándar media

**Distribución**

☐ Asimetría

☐ Curtosis

Y obtenemos lo siguiente:

### Estadísticos

Paquetes

|             |          |        |
|-------------|----------|--------|
| N           | Válido   | 230    |
|             | Perdidos | 0      |
| Percentiles | 75       | 225,00 |