

## **Tema 2: Organización de datos**

1. Introducción.
2. Bases de datos.
3. Codificación, introducción de datos, depuración, transformaciones y tratamiento de ficheros.
  4. Valores atípicos
  5. Valores faltantes
6. Distribución de frecuencias.
7. Gráficas para variables cualitativas y cuantitativas.

## 2. Bases de datos.

1. Codificación, introducción de datos, depuración, transformaciones y tratamiento de ficheros.

2. Valores atípicos

3. Valores faltantes

### 3. Distribución de frecuencias.

Elementos

**VALORES DE LA VARIABLE EN LA MUESTRA ( $X_i$ )**

**FRECUENCIA ABSOLUTA ( $n_i$ ):** número de veces que se repite el valor  $X_i$  en la muestra

**FRECUENCIA RELATIVA ( $p_i$ ):** Cociente entre la frecuencia absoluta de un determinado valor  $X_i$  y el tamaño de la muestra ( $p_i = n_i/n$ ) **(PROPORCIÓN)**

**FRECUENCIA ABSOLUTA ACUMULADA ( $n_a$ ):** número de veces que se repite el valor  $X_i$  o cualquier otro valor inferior

**FRECUENCIA RELATIVA ACUMULADA ( $p_a$ ):** cociente entre la frecuencia absoluta acumulada de un determinado valor  $X_i$  y el tamaño de la muestra ( $p_a = n_a/n$ ) **(PROPORCIÓN ACUMULADA)**

## Tabla de distribución de frecuencias

Ej. Variable discreta: nº de hijos por familia

| nº hijos | frec.abs. | frec.rel.<br>o prop. | frec.abs.<br>acum. | frec.rel.<br>(prop.)acum. |
|----------|-----------|----------------------|--------------------|---------------------------|
| $X_i$    | $n_i$     | $p_i$                | $n_a$              | $p_a$                     |
| 4        | 1         | 0.05                 | 20                 | 1.00                      |
| 3        | 3         | 0.15                 | 19                 | 0.95                      |
| 2        | 7         | 0.35                 | 16                 | 0.80                      |
| 1        | 6         | 0.30                 | 9                  | 0.45                      |
| 0        | 3         | 0.15                 | 3                  | 0.15                      |
|          | 20        | 1.00                 |                    |                           |

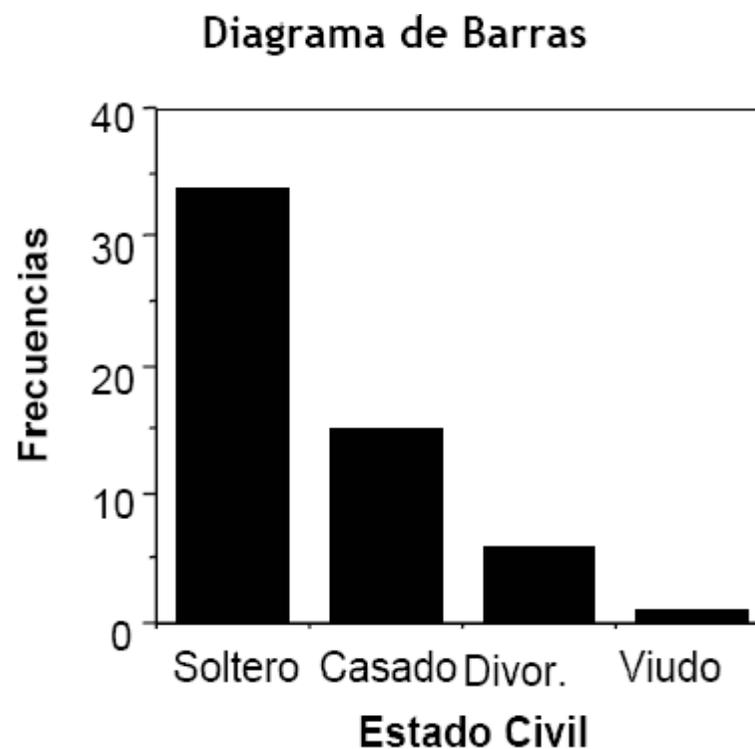
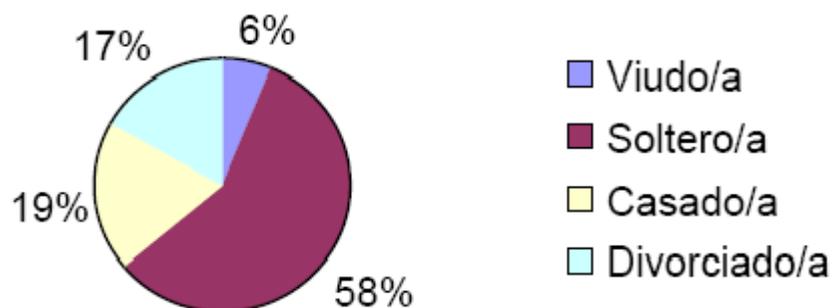
#### 4. Gráficas para variables cualitativas y cuantitativas.

| TIPOS DE VARIABLE  | GRAFICA                           | EJE ABSCISAS (x)   | EJE ORDENADAS (y)                     |
|--|-----------------------------------|--|---------------------------------------|
| Nominal<br>(sexo, estadocivil)   | Diagrama de barras                | Valores que puede adoptar la variable                    | Frecuencias                           |
|  | Pictograma o diagrama de sectores | --   | --                                    |
| Cuantitativa discreta (num. hijos) y ordinales (nivel educativo, de 1 a 4) | Diagrama de barras                | Valores que puede adoptar la variable                    | Frecuencias                           |
|  | Polígono de frecuencias           | Valores que puede adoptar la variable                    | Frecuencias                           |
| Cuantitativa continua (horas de estudio)                                   | Polígono de frecuencias           | Valores que puede adoptar la variable                    | Frecuencias                           |
|  | Histograma                        | Valores que puede adoptar la variable                    | Frecuencias                           |
| Informes psicopedagógicos o de rendimiento                                 | Perfil ortogonal                  | Diferentes momentos de la evaluación, diferentes pruebas | Valores que puede adoptar la variable |

## Variable categórica (NOMINAL)

1. Diagrama de barras  
(rectángulos)
2. Pictograma o diagrama de sectores

Pictograma o Diagrama de sectores



# Variables: ORDINALES CUANTITATIVAS DISCRETAS

1. Diagrama de barras
2. Diagrama de barras acumulado
3. Polígono de frecuencias

Polígono de freqüències

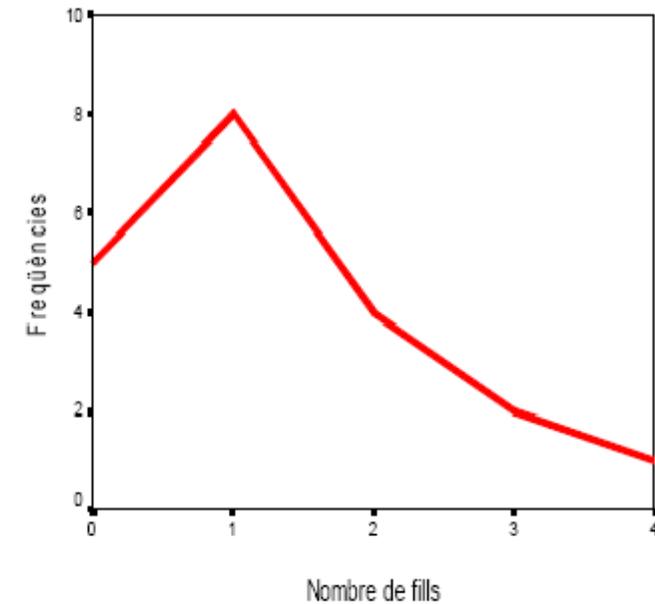


Diagrama de barres

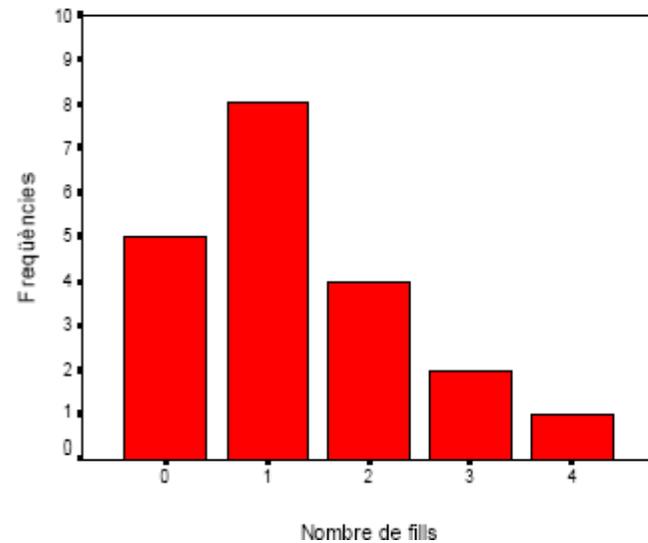
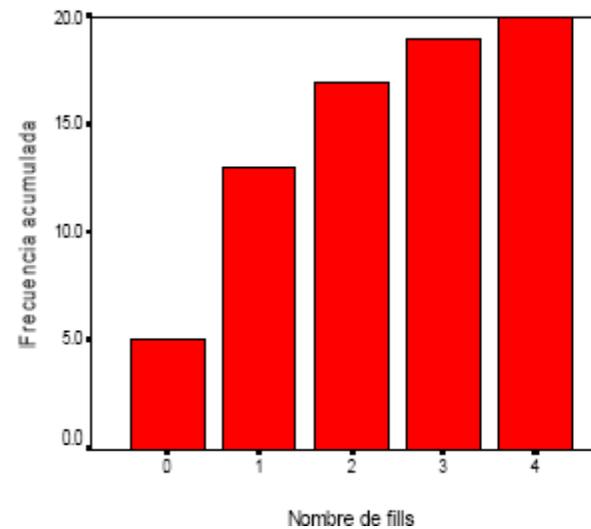


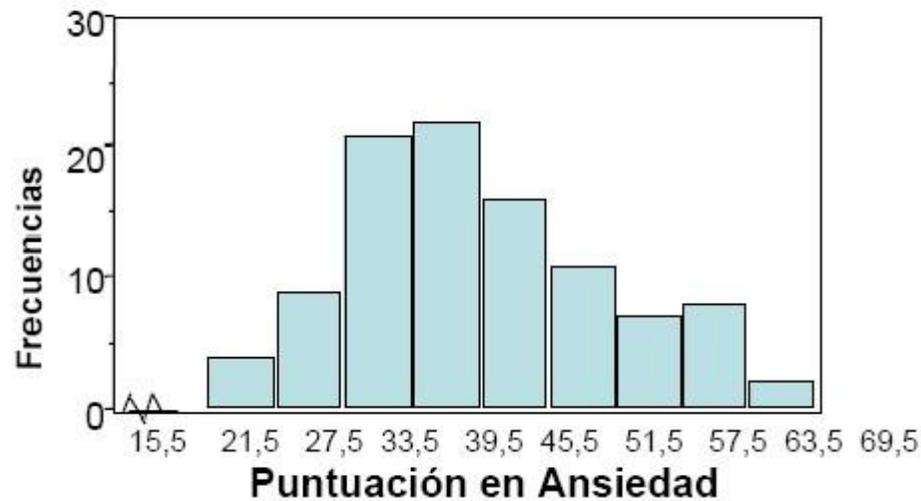
Diagrama de barres acumulatiu



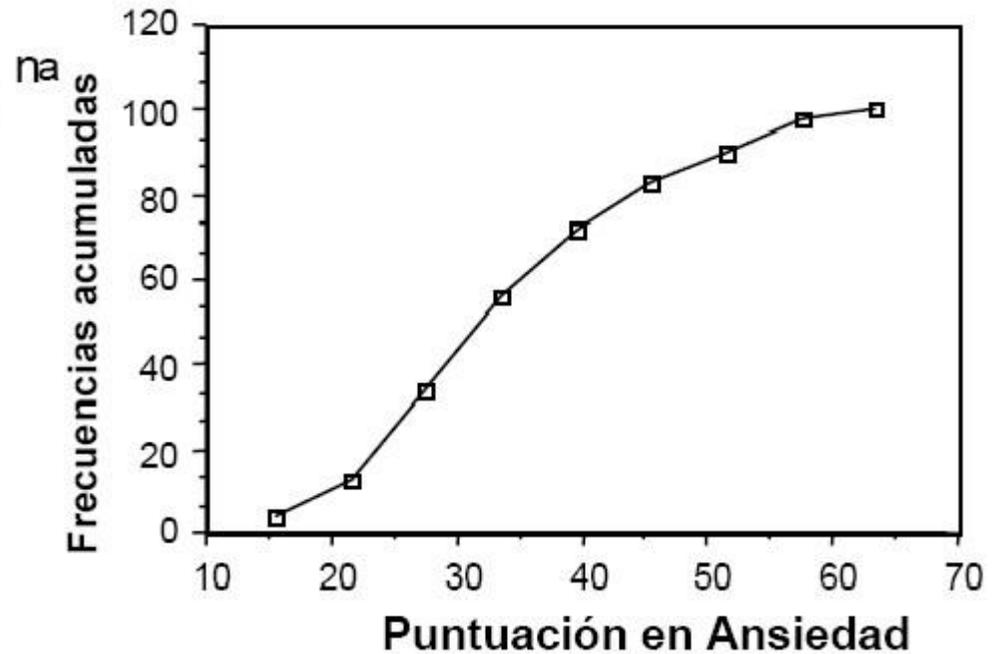
# Variable CUANTITATIVA CONTINUA (datos agrupados en intervalos)

1. Histograma
2. Polígono de frecuencias
3. Polígono de frecuencias acumuladas

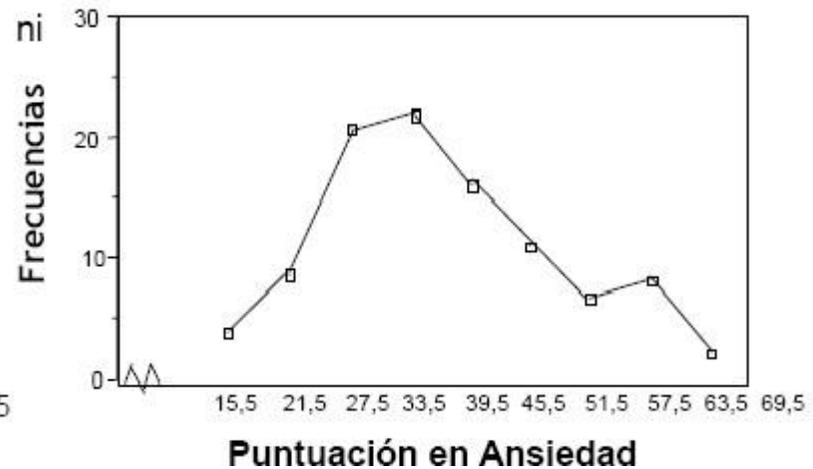
Histograma



Polígono de frecuencias acumuladas



Polígono de frecuencias



## Representaciones gráficas: Exploratorias: DIAGRAMA DE TALLO Y HOJAS

De cada puntuación se pueden diferenciar dos partes:

- 1) Tallo (primer o primeros dígitos) (común para diferentes datos)
- 2) Hoja (resto de los dígitos)

Ex. 10 12 21 21 22 25 31 32 34 37 38 40

1 | 02

2 | 1125

3 | 12478

4 | 0

1 | 02  
2 | 1125  
3 | 12478  
4 | 0

**Principal ventaja:** Nos permite identificar cada puntuación individual al tiempo que ofrece un dibujo de la distribución

# Representaciones gráficas: Exploratorias: DIAGRAMA DE CAJA Y BIGOTES

Se indica tanto la tendencia central (Mediana -la línea que corta la caja) como de variabilidad (las dos líneas que hacen los límites de la caja que son el percentil 75 y el percentil 25).

Además informan de las puntuaciones “atípicas”.

