Tema 1

Estadística descriptiva (unidimensional)

Población y muestra

Población

Conjunto de datos sobre el que queremos obtener conclusiones.

Lo que nos gustaría estudiar

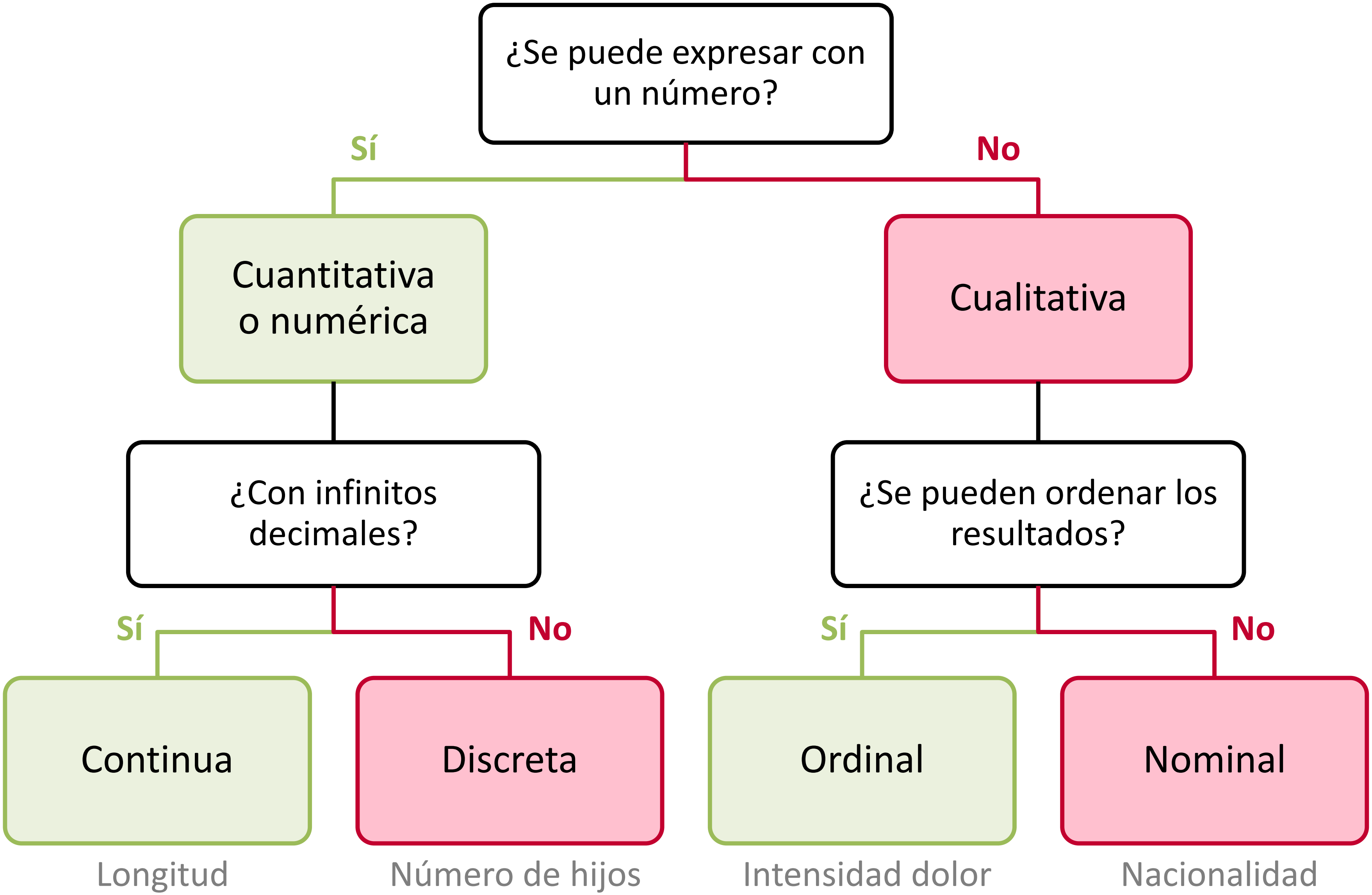
Muestra

Subconjunto de datos en el que hacemos las observaciones.

- Representativo
- Suficientemente grande

Encuesta electoral

Tipos de variables



Ejercicio

Clasifica las siguientes variables:

	¿Número?	¿Decimales? / ¿Orden?	Tipo de variable
Grupo sanguíneo			
Grado de satisfacción			
Número de averías			
Altura			

Ejercicio

Clasifica las siguientes variables:

	¿Número?	¿Decimales? / ¿Orden?	Tipo de variable
Grupo sanguíneo	No	Sin orden	Cualitativa nominal
Grado de satisfacción	No	Con orden	Cualitativa ordinal
Número de averías	Sí	Sin decimales	Cuantitativa discreta
Altura	Sí	Con decimales	Cuantitativa continua

Ejercicio

Clasifica las siguientes variables:

- Número de habitantes de un país.
- Sueldo en euros.

Variables que no son lo que parecen

Número de habitantes de un país

Es una variable **cuantitativa discreta** (sin decimales), pero debido a la gran cantidad de resultados posibles (es raro que dos países tengan el mismo número de habitantes) se suele tratar como si fuese **cuantitativa continua**.

Sueldo en euros

A pesar de tener decimales, sólo tiene dos (no un número infinito de ellos). A pesar de ser una variable **discreta**, por lo mismo que antes, se trata como si fuese **continua**.

Modalidades (posibles respuestas)

Agrupar en intervalos

Pasamos de una variable cuantitativa a una cualitativa.

Edad: menos de 20, de 20 a 50, más de 50.

Errores a evitar

Siempre tiene que haber una única respuesta válida:

Número de hijos: Ninguno, menos de 5, más de 2.

De los siguientes, qué le gusta: deporte, cine.

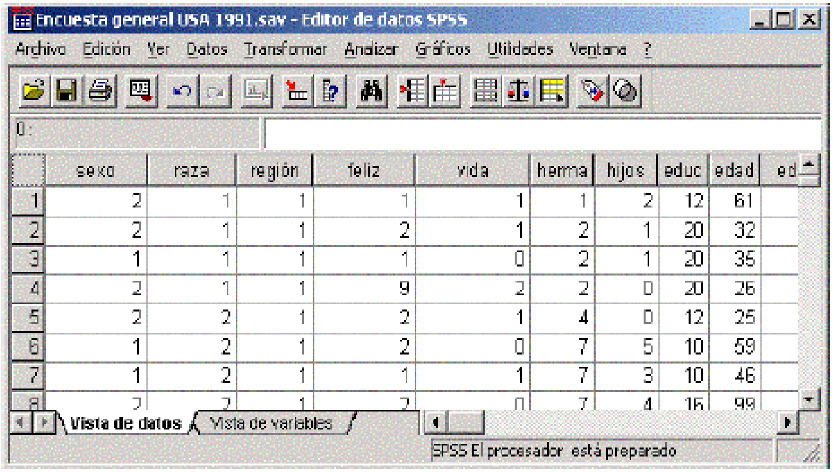
Crear dos variables

Codificar variables en SPSS

Asignar etiquetas a los posibles valores de las variables.

1: Hombre
2: Mujer

1: Muy feliz
2: Bastante feliz
3: No demasiado feliz



SPSS

Ver → Etiquetas valor

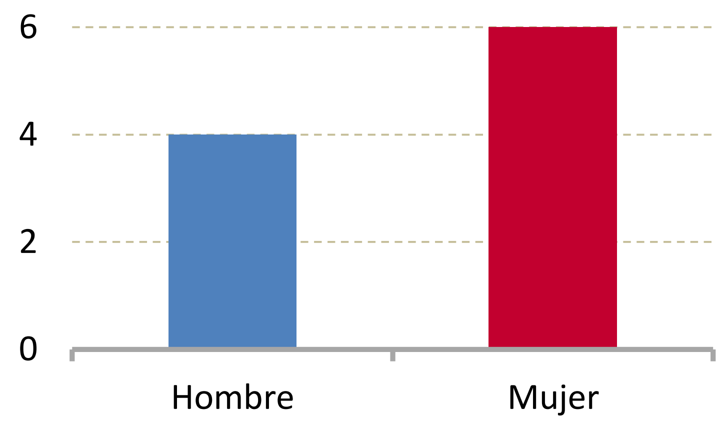
Tablas de frecuencias

Datos, frecuencias y diagramas

HHHHMMMMMM

Sexo	Frecuencia
Hombre	4
Mujer	6

HMMHMMMHHM



Frecuencias

Absoluta: número de individuos en cada modalidad (n_i)

Relativa: proporción en cada modalidad $\left(f_i = \frac{n_i}{N}\right)$

Acumulada: suma los casos anteriores $\left(F_i = \sum_{j=1}^{i} f_i\right)$

No tiene sentido con variables cualitativas nominales

Número de hijos		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	*Ninguno*	467	30,8	31,1	31,1
	Uno o dos	872	57,5	58,0	89,0
	Tres o más	165	10,9	11,0	100,0
	Total	1504	99,1	100,0	
Perdidos	No contesta	13	0,9		
Total		1517	100,0		

Frecuencias

Absoluta: número de individuos en cada modalidad (n_i)

Relativa: proporción en cada modalidad $\left(f_i = \frac{n_i}{N}\right)$

Acumulada: suma los casos anteriores $\left(F_i = \sum_{j=1}^{i} f_i\right)$

No tiene sentido con variables cualitativas nominales

SPSS

Analizar → Estadísticos descriptivos → Frecuencias

Ejercicio

¿Cuántas centrales eléctricas tienen menos de 2 averías?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Ejercicio

¿Cuántas centrales eléctricas tienen menos de 2 averías?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

674

Ejercicio

¿Qué porcentaje de las centrales tiene 5 o más averías?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Ejercicio

¿Qué porcentaje de las centrales tiene 5 o más averías?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

7.8

7.8

Ejercicio

¿Qué cantidad de averías es tal que al menos el 50% de la población tiene una cantidad inferior o igual?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Ejercicio

¿Qué cantidad de averías es tal que al menos el 50% de la población tiene una cantidad inferior o igual?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5 ≥ 50
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Agrupar en intervalos

$N =$ Número de datos

Número de intervalos

1. El número de intervalos esté entre 5 y 20.
2. Podemos seguir un indicador:
   - Fórmula de Sturgess: $1 + 3.3 \cdot \ln(N)$
   - Fórmula de la raíz: $\sqrt{N}$
3. Seleccionar las clases de forma que la amplitud de los intervalos sea similar.

Agrupar en intervalos

Marcas de clase

Valor intermedio del intervalo.

Lo utilizamos para hacer los cálculos necesarios.

Los cálculos serán siempre más precisos con los valores originales

SPSS

Transformar → Agrupación visual

Cambios de variables

$$Y = a \cdot X + b$$

Puede ser útil cuando queremos cambiar las unidades de una variable (temperatura, moneda…).

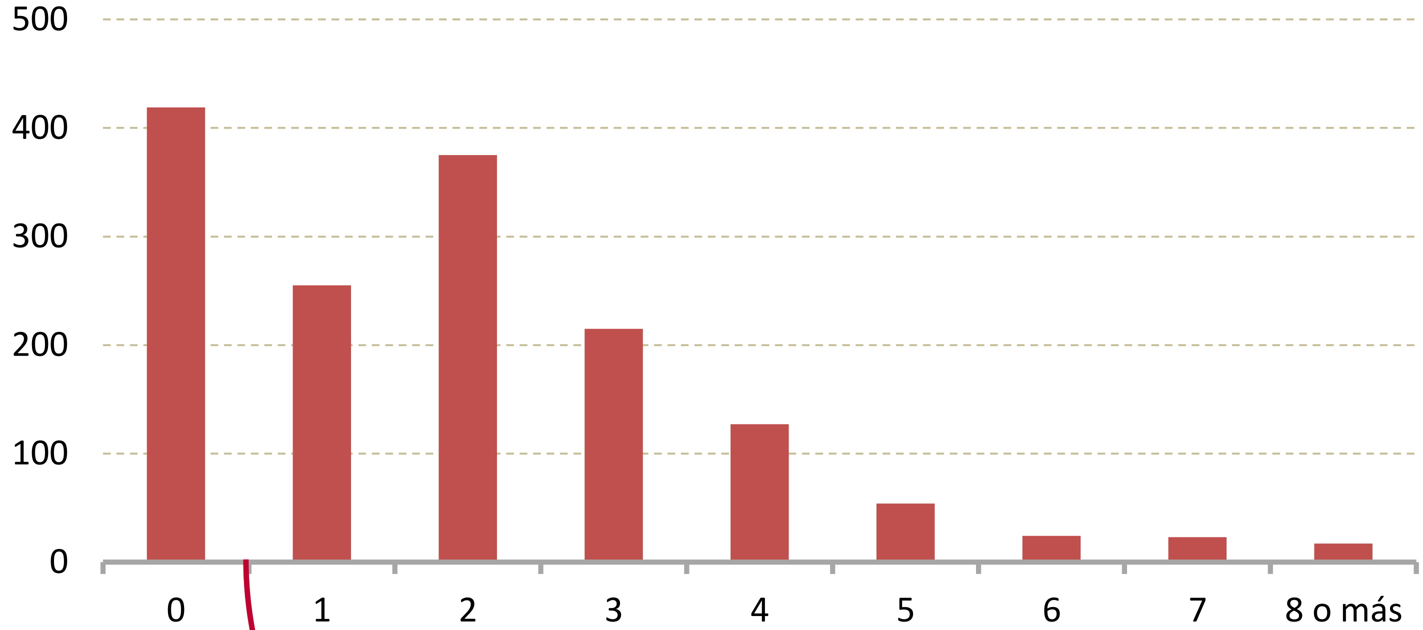
También nos sirve para crear ecuaciones teniendo en cuenta varias variables.

SPSS

Transformar → Calcular variable

Diagramas

Diagrama de barras

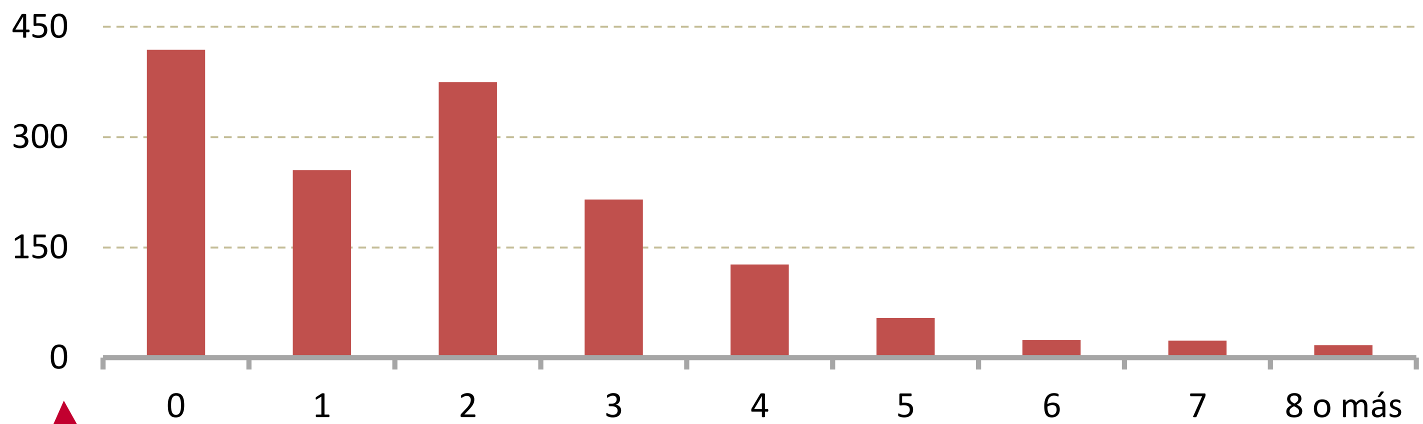


Espacio entre las barras, ya que no hay nada entre medias

SPSS

Analizar → Estadísticos descriptivos → Frecuencias

Diagrama de barras



Con frecuencia absoluta o relativa el gráfico es igual

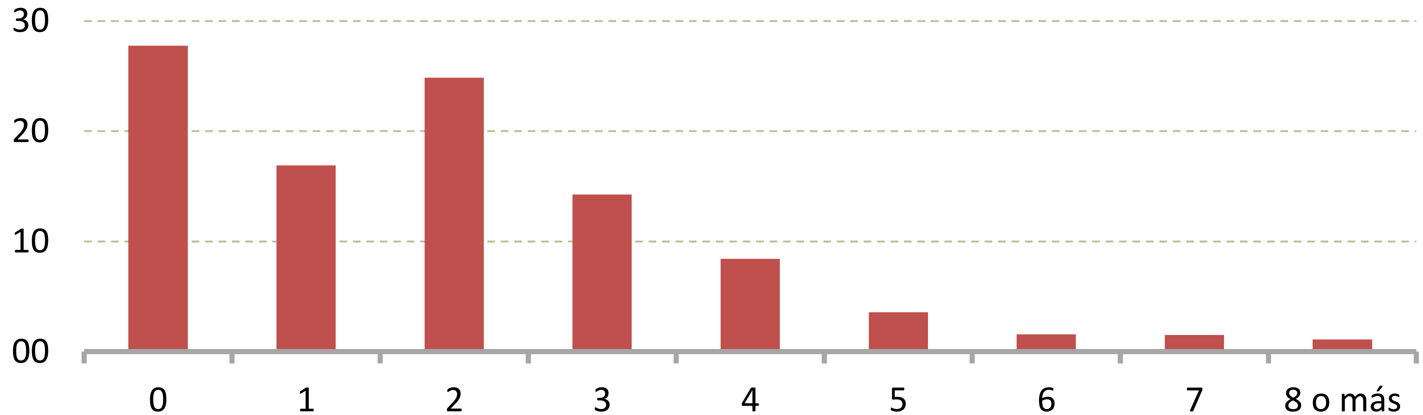
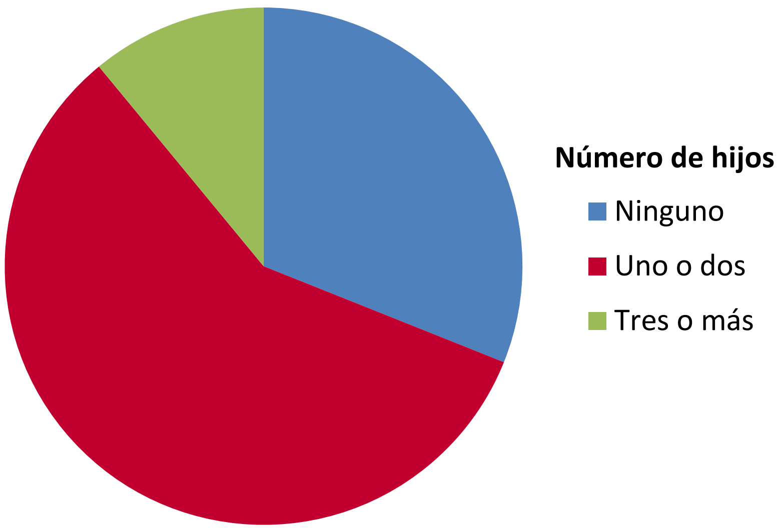
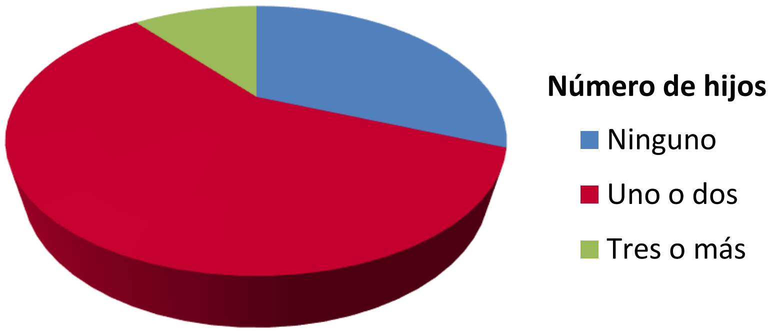
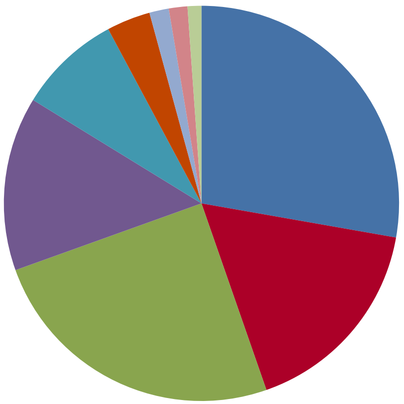


Diagrama de sectores o tarta





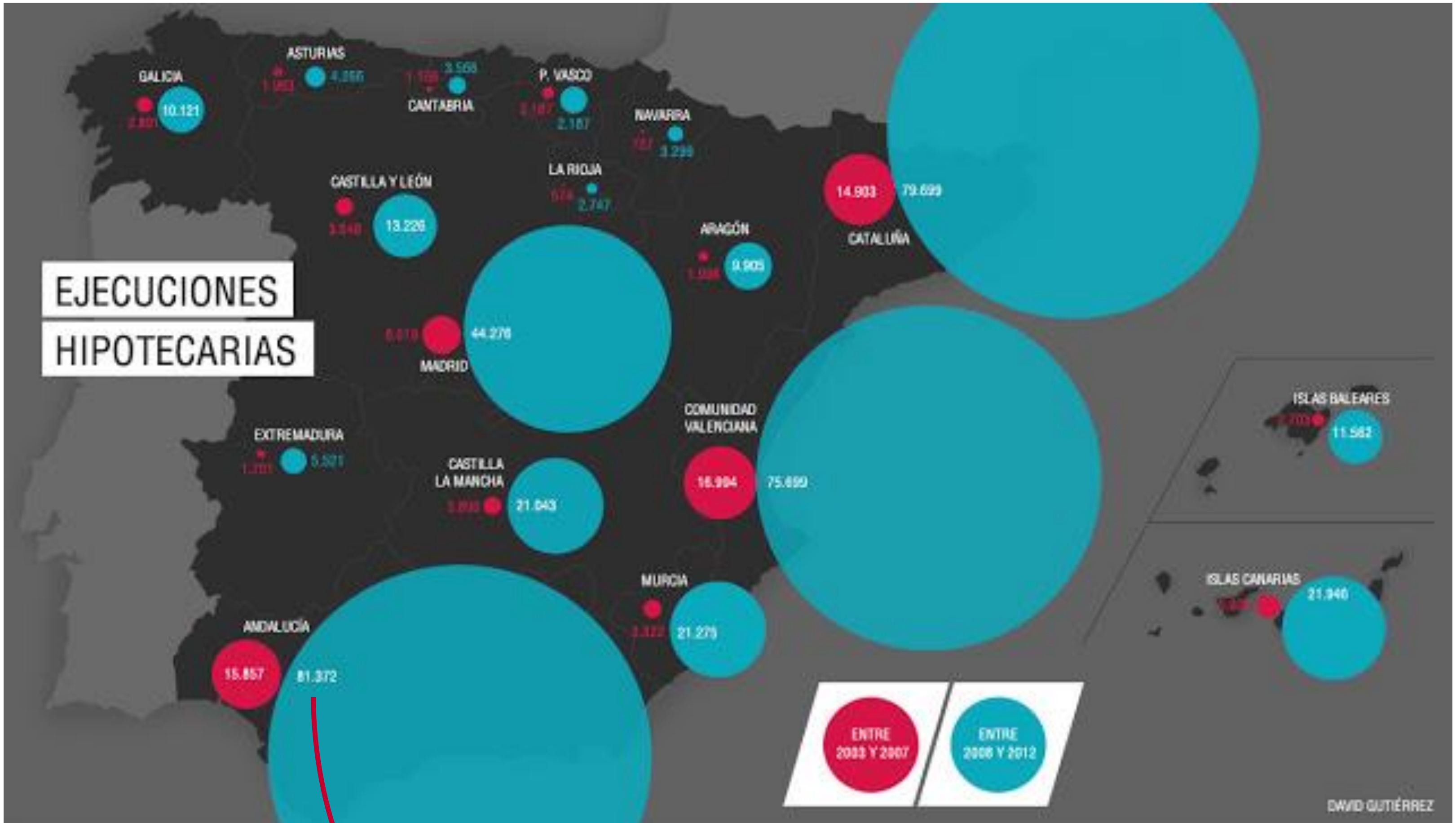
Evitar el 3D o usarlo con muchas clases



SPSS

Analizar → Estadísticos descriptivos → Frecuencias

Pictograma (áreas)



La cifra es 5,1 veces más grande.
El círculo es 26,3 veces más grande.

Diagrama de tallo y hojas

Tallo

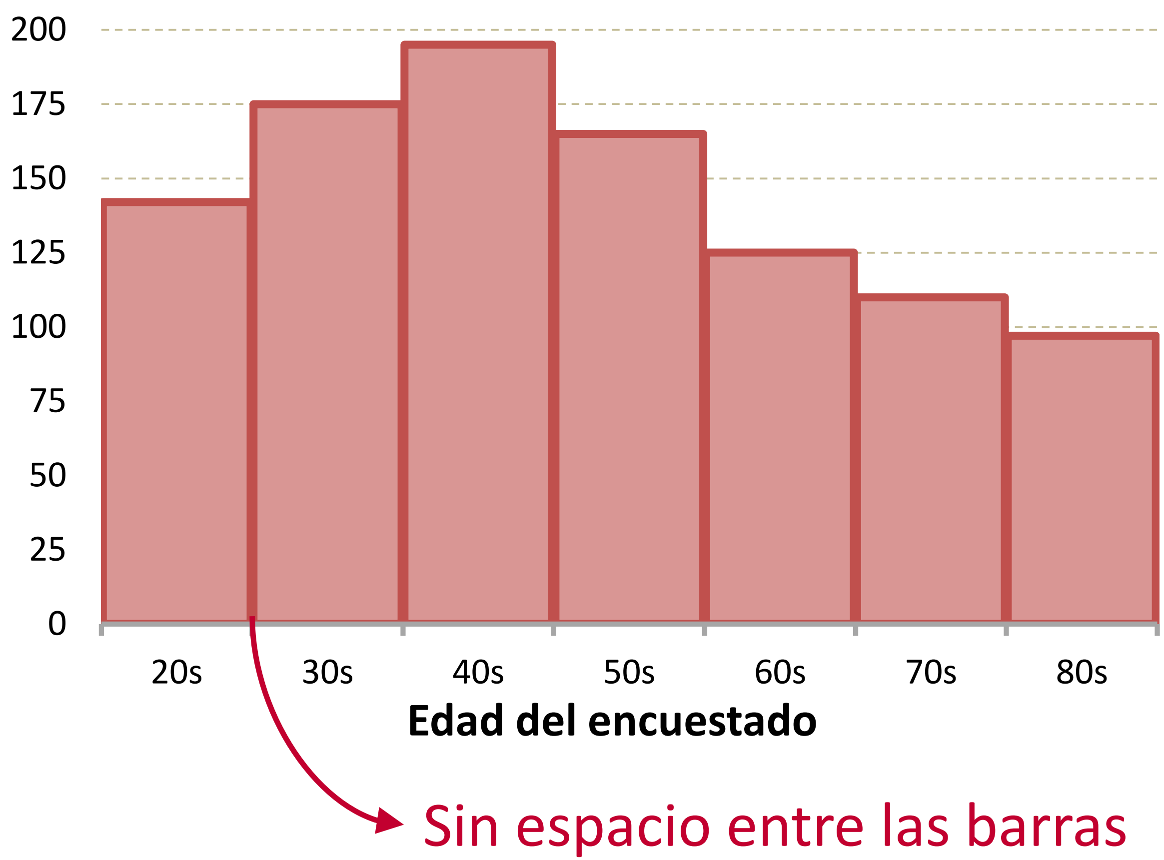
Lunes a viernes laborables (Vigente de 1 de septiembre a 15 de julio)

6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:
00	05^{p}	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
30	15^{p}	10^{p}	15	30	30	30	20	20	20	20	20	20	30	30	30	30	
55	30	20^{p}	30				40	40	40	40	40	40					
	45^{p}	30															
		45															

Hojas

Viendo el número de hojas de cada tallo podemos ver la densidad en cada hora

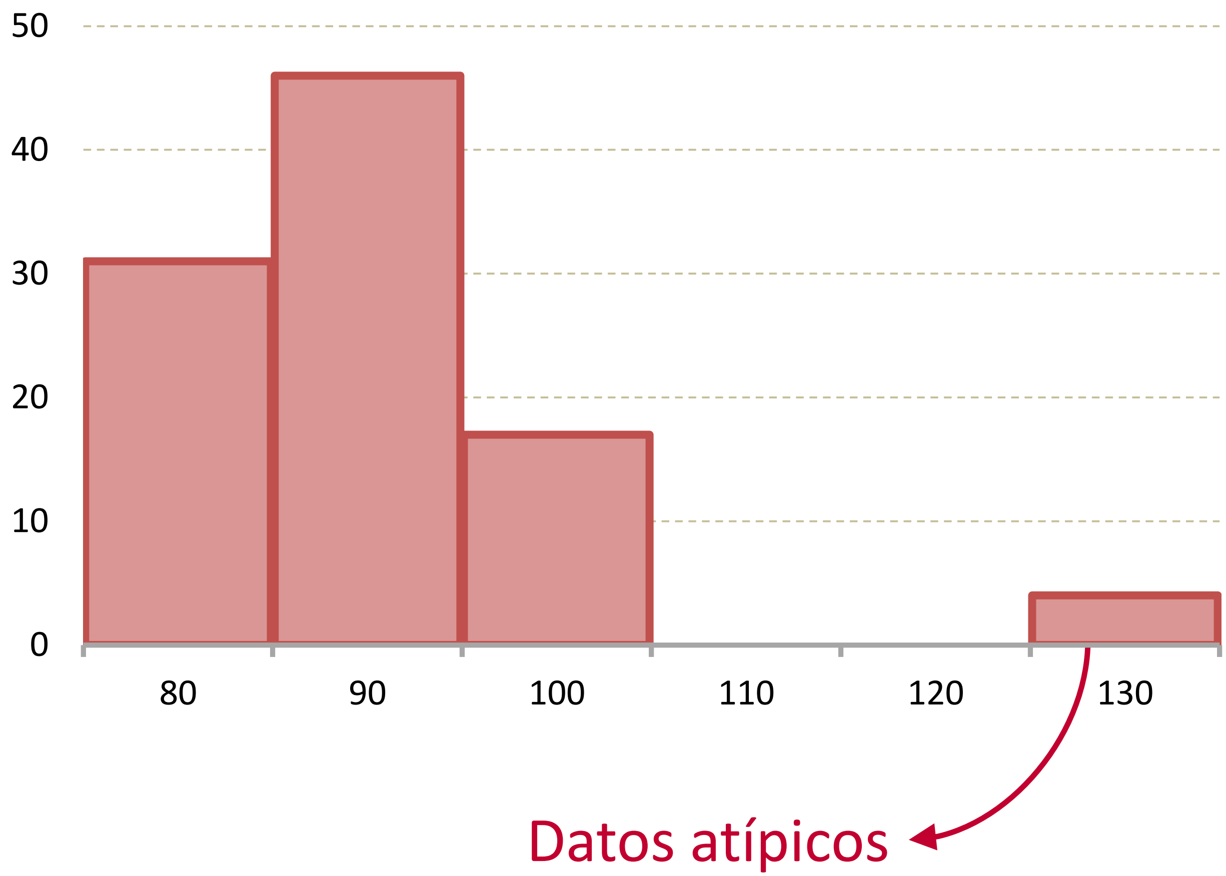
Histograma



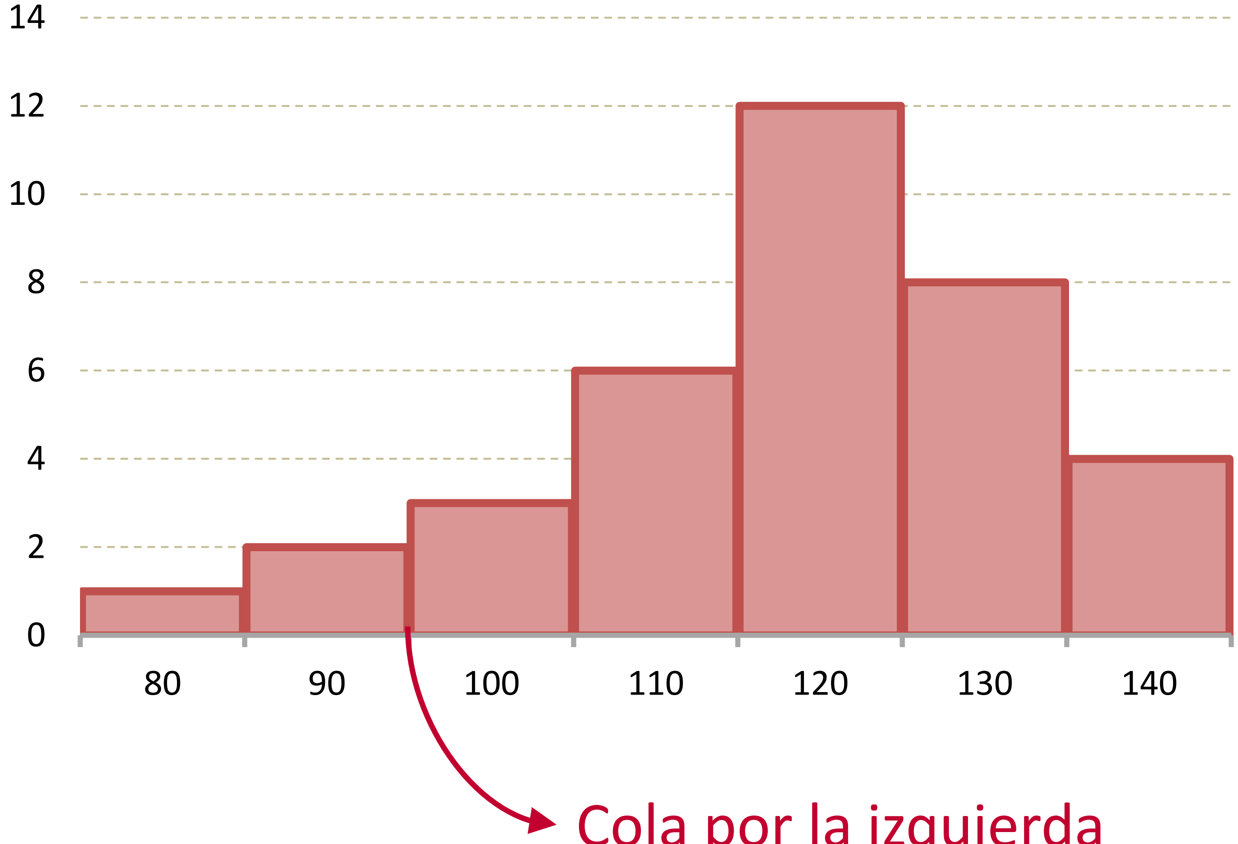
SPSS

Gráficos → Cuadros de diálogo antiguos → Histograma

Histograma. Datos atípicos



Histograma. Simetría o asimetría



Ejercicio

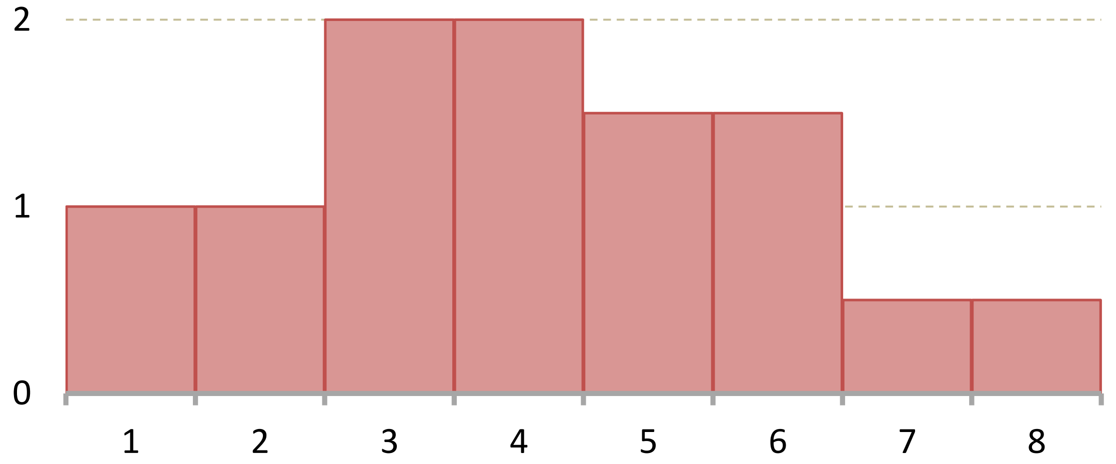
Dibuja un histograma para la siguiente distribución

	Frecuencia	Ancho del intervalo	Densidad
1-3	2		
3-5	4		
5-7	3		
7-9	1		
Total	10		

Ejercicio

Dibuja un histograma para la siguiente distribución

	Frecuencia	Ancho del intervalo	Densidad
1-3	2	2	1
3-5	4	2	2
5-7	3	2	1.5
7-9	1	2	0.5
Total	10		



Ejercicio

Dibuja un histograma para la siguiente distribución

	Frecuencia	Ancho del intervalo	Densidad
1-3	2		
3-6	3		
6-10	1		
10-12	4		
Total	10		

Ejercicio

Dibuja un histograma para la siguiente distribución

	Frecuencia	Ancho del intervalo	Densidad
1-3	2	2	1
3-6	3	3	1
6-10	1	4	0.25
10-12	4	2	2
Total	10		

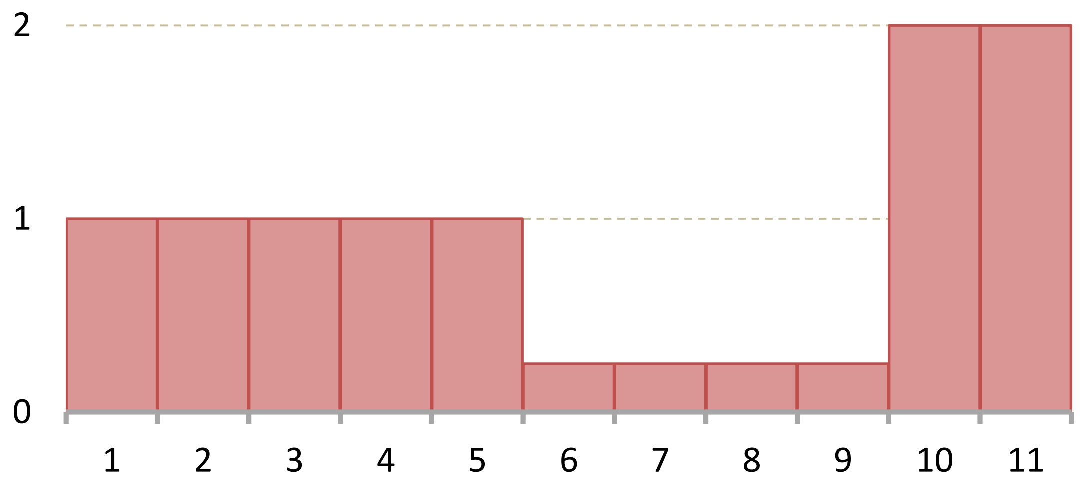


Diagrama de caja y bigotes

A pesar de su nombre, es bastante importante.

Veremos cómo representarlo más adelante.
Necesitamos ver varios conceptos antes.

Estadísticos

Parámetros y estadísticos

Parámetro

Cantidad numérica calculada sobre una **población**.

Tendríamos que observar todos sus elementos

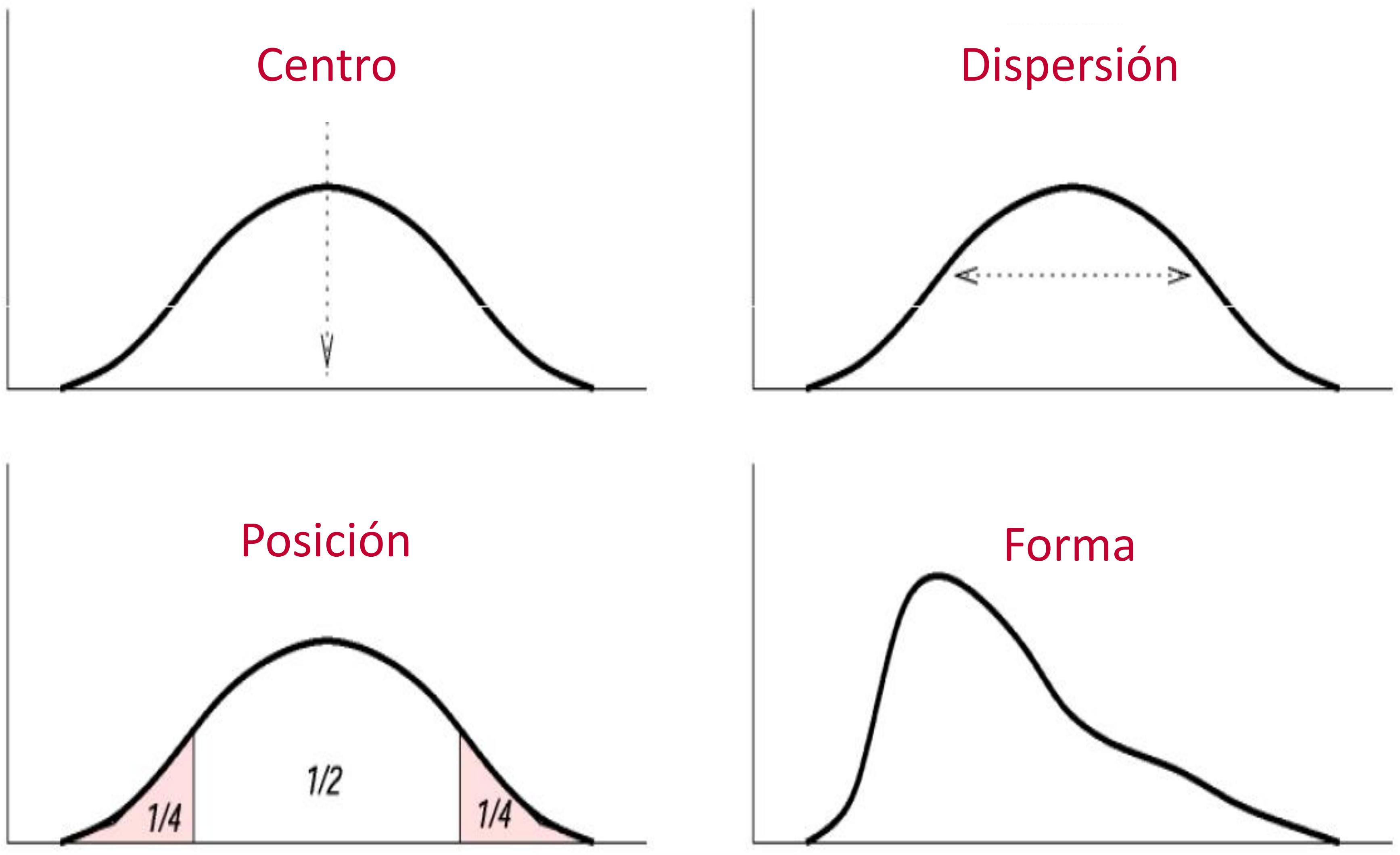
Estadístico

Cantidad numérica calculada sobre una **muestra**.

Es más común observar un conjunto de individuos

Inferencia: usarlo para estimar el parámetro

Estadísticos



Estadísticos de posición

Estadísticos de posición central

Centro

Expresan dónde está el 'centro' de la distribución.

Existen tres tipos de estadísticos de posición central:

- Media
- Mediana
- Moda

Cada estadístico tiene sus ventajas e inconvenientes

Media aritmética

$$\bar{x} = \frac{\sum x_i \cdot n_i}{N}$$

Intervalo	Valor medio x_i	Frecuencia n_i
1-3	2	2
3-5	4	4
5-7	6	3
7-9	8	1
		N=10

Media aritmética

Sumamos todos los valores.

El resultado lo dividimos entre el número de valores.

Media aritmética. Problemas

Todos los valores afectan al cálculo de la media.

Los valores extremos producen grandes cambios.

$$2,\ 3,\ 4,\ 5,\ 6,\ 6,\ 9 \quad \Longrightarrow \quad \bar{x} = 5$$

¿Son representativos de la distribución?

$$2,\ 3,\ 4,\ 5,\ 6,\ 6,\ 100 \quad \Longrightarrow \quad \bar{x} = 18$$

Ejercicio

Un artículo sube su precio un 20%.
El día siguiente su precio baja un 20%.

¿Cuál es su precio final?
¿Cuál es la media de los cambios?

Ejercicio

Un artículo valía 100 euros.
Un día su precio sube un 50%. El día siguiente, baja un 50%.

¿Cuál es su precio final?
¿Cuál es la media de los cambios?

Ejercicio

Un artículo valía 100 euros.
Un día su precio sube un 50%. El día siguiente, baja un 50%.

¿Cuál es su precio final?
¿Cuál es la media de los cambios?

$$100\text{ €} \xrightarrow{+50\%} 150\text{ €} \xrightarrow{-50\%} 75\text{ €}$$

$$100\text{ €} \cdot 1{,}5 \cdot 0{,}5 = 75\text{ €} \quad \rightarrow \quad x \cdot 1{,}5 \cdot 0{,}5 = 0.75 \cdot x$$

Variación media: $\sqrt[2]{1{,}5 \cdot 0{,}5} = \sqrt[2]{0{,}75} = 0.87$

Ha bajado un 13% de media en cada cambio

Media geométrica

$$G = \sqrt[N]{\prod x_i}$$

Media geométrica

Muy utilizada con proporciones.

Multiplicamos todos los valores.
Realizamos la raíz n-ésima del resultado.

Ejercicio

Queremos hacer un viaje de ida y vuelta con una velocidad media de 120 km/h.

En la ida hemos tenido mucho atasco, por lo que nuestra velocidad media ha sido de 60 km/h.

¿A qué velocidad debemos ir a la vuelta para que la velocidad media total sea de 120 km/h?

Ejercicio

Queremos hacer un viaje de ida y vuelta con una velocidad media de 120 km/h.

En la ida hemos tenido mucho atasco, por lo que nuestra velocidad media ha sido de 60 km/h.

¿A qué velocidad debemos ir a la vuelta para que la velocidad media total sea de 120 km/h?

Supongamos que cada viaje es de 120 km. Entre ida y vuelta, 240 km.

Para ir a 120 km/h de media, deberíamos tardar, entre ida y vuelta, 2 horas.

En la ida hemos ido a 60 km/h para recorrer 120 km, tardando 2 horas.

¡No nos queda tiempo para la vuelta!

Media armónica

$$H = \frac{N}{\sum \frac{1}{x_i}}$$

Media armónica

Muy utilizada con ratios.

Sumamos los inversos de todos los valores.
Realizamos la inversa del resultado.
Multiplicamos el resultado por el número de valores.

Mediana

Nos da un valor tal que la mitad de los datos están por encima y la otra mitad por debajo.

- Ordenamos todos los valores de menor a mayor.
- Tomamos el que esté en la posición central.
- Si no hay un valor central, realizamos la media entre los dos valores más cercanos al centro.

$$2, 3, 4, 5, 8, 8, 9 \quad \Longrightarrow \quad Me = 5$$

$$2, 3, 4, 5, 6, 8, 8, 9 \quad \Longrightarrow \quad Me = 5{,}5$$

Ejercicio

Calcula la mediana de esta distribución:

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Ejercicio

Calcula la mediana de esta distribución:

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5 ≥ 50
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Mediana. Problemas

Sólo depende del valor que esté en la posición central,
por lo que los valores extremos no le afectan.

2, 3, 4, 5, 6, 8, 8, 9 $\Longrightarrow$ $Me = 5{,}5$

¿Son representativos de la distribución?

4, 4, 5, 5, 12, 62, 123 $\Longrightarrow$ $Me = 5$

Moda

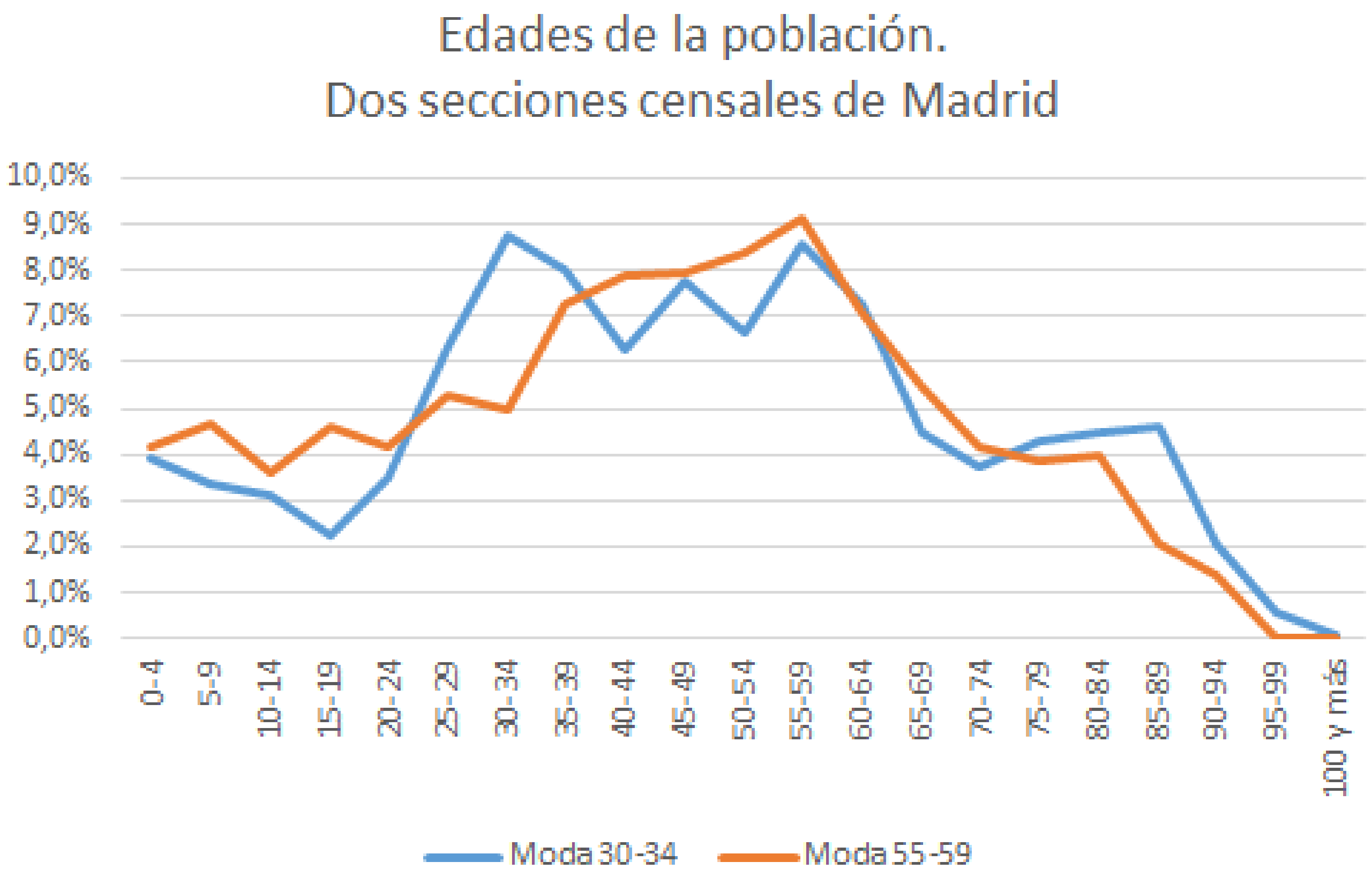
Es el valor que más se repite.

En caso de utilizar intervalos, es el intervalo que contiene el mayor número de valores.

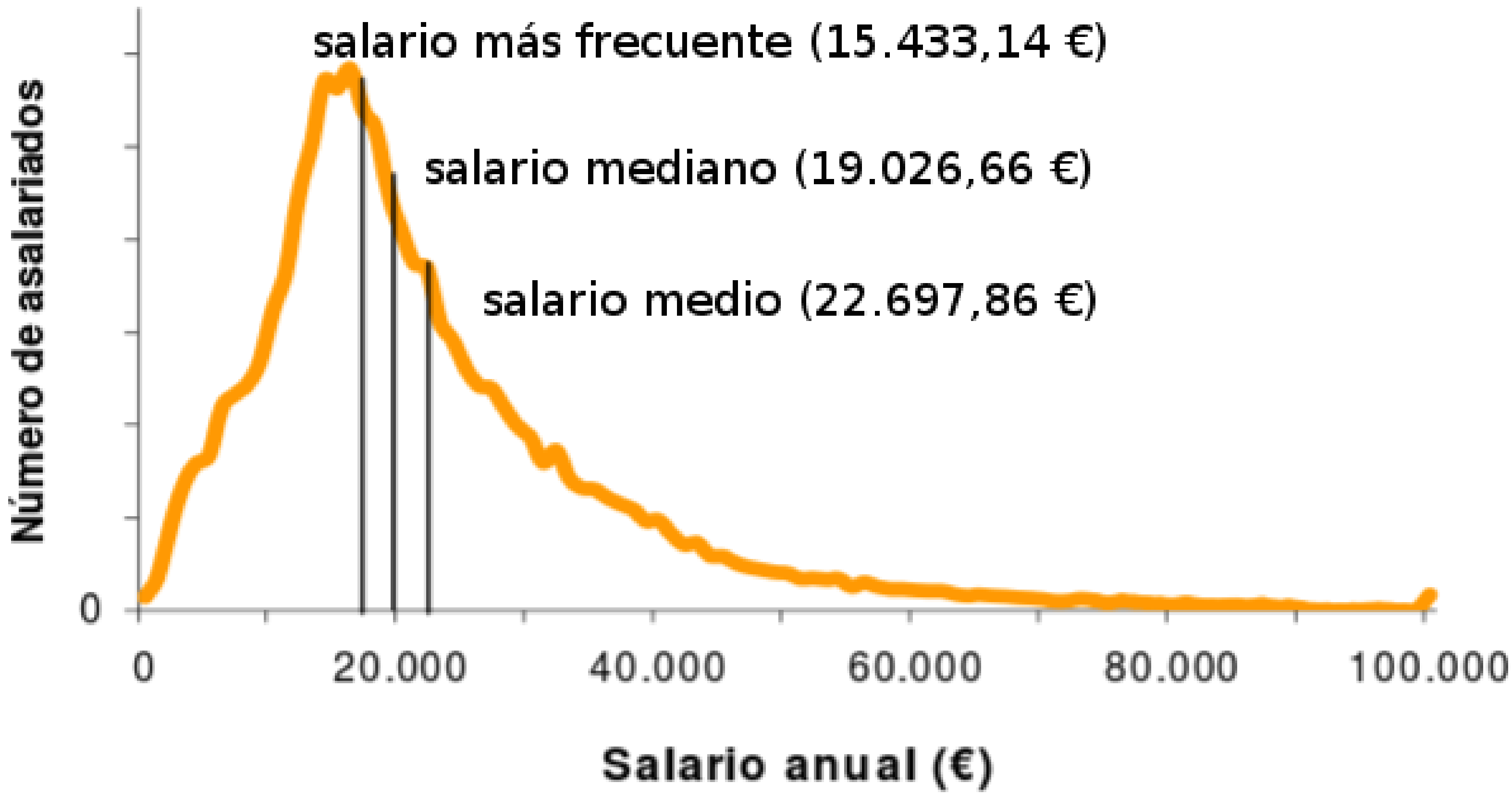
Puede variar según cómo hagamos los intervalos

Intervalo	**Frecuencia** n_i
1-3	2
3-5	4
5-7	3
7-9	1
	N=10

Moda. Problemas



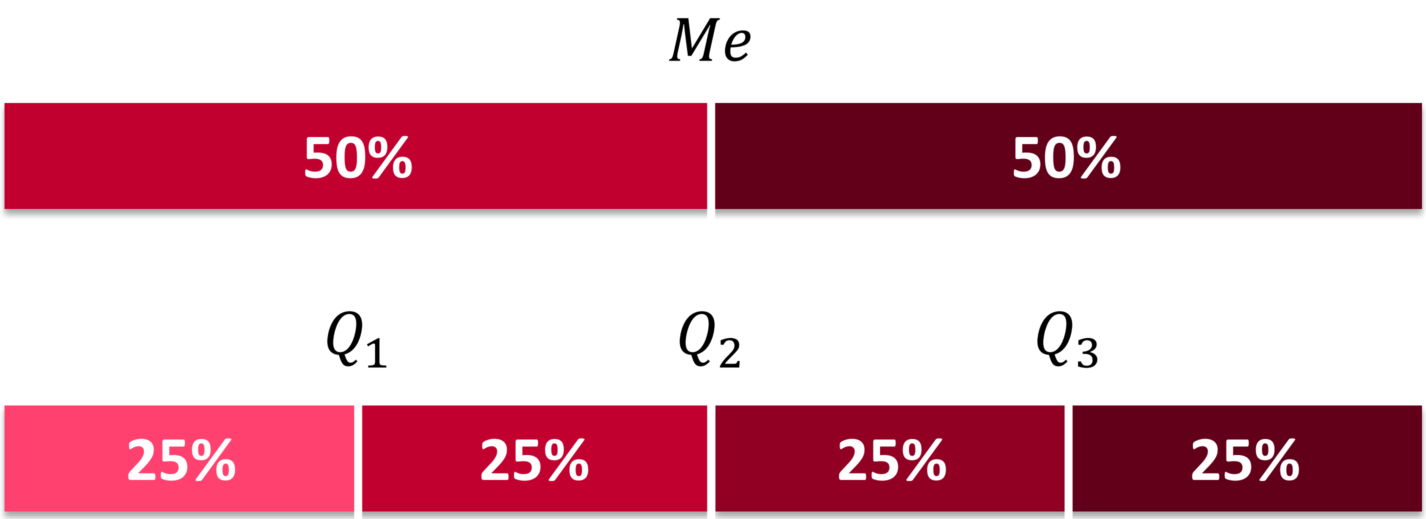
Salario anual en España



Otros estadísticos de posición

Cuartiles

- Ordenamos los datos de menor a mayor.
- Elegimos la mediana (dividimos los datos en dos).
- Cada mitad la dividimos en dos y tenemos los cuartiles.



Ejercicio

Calcula los cuartiles de esta distribución:

3, 4, 4, 5, 5, 5, 7, 8 , 12, 13, 15

Ejercicio

Calcula los cuartiles de esta distribución:

3, 4, 4, 5, 5, 5, 7, 8 , 12, 13, 15

3, 4, 4, 5, 5, 5, 7, 8 , 12, 13, 15

$Me = Q_2$

3, 4, 4, 5, 5, 5, 7, 8 , 12, 13, 15

Q_1 $Me = Q_2$ Q_3

Ejercicio

Calcula los cuartiles de esta distribución:

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Ejercicio

Calcula los cuartiles de esta distribución:

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
$Q_1 =$ *0*	419	27.8	27.8 ≥ 25
1	255	16.9	44.7
$Q_2 =$ *2*	375	24.9	69.5 ≥ 50
$Q_3 =$ *3*	215	14.2	83.8 ≥ 75
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Otros estadísticos de posición

Percentiles

La idea es similar a la mediana y a los cuartiles

- Ordenamos los datos de menor a mayor.
- En vez de dividir en dos partes (mediana) o en cuatro (cuartiles), dividimos en 100 partes.

Percentil 30: valor que es superior al 30% de los valores.

Percentil 30

30% **70%**

Ejercicio

¿Qué valor es superado por el 10% de los datos?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Ejercicio

¿Qué valor es superado por el 10% de los datos?

Número de averías	Frecuencia	Porcentaje	Porcentaje acumulado
0	419	27.8	27.8
1	255	16.9	44.7
2	375	24.9	69.5
3	215	14.2	83.8
4	127	8.4	92.2 ≥ 90
5	54	3.6	95.8
6	24	1.6	97.3
7	23	1.5	98.9
8 o más	17	1.1	100.0
Total	1509	100.0	

Estadísticos de dispersión

Rango

Es la diferencia que hay entre el valor mínimo y el máximo.

Sólo depende de esos dos valores, y no de si el resto de datos están agrupados o no.

Se ve muy afectado por casos extremos.

2, 3, 4, 5, 6, 6, 7, 8, 8, 9, 47 $\Longrightarrow$ $Rango = 45$

2, 6, 10, 13, 17, 21, 24, 27, 30, 32, 37 $\Longrightarrow$ $Rango = 35$

Rango intercuartil

$$RI = Q_3 - Q_1$$

Es la diferencia que hay entre primer y tercer cuartil.

Sólo depende de esos dos valores e ignora lo que pasa en la mitad de los datos (el 25% menor y el 25% mayor).

No se ve afectado por casos extremos.

2, 3, 4, 5, 6, 6, 7, 8, 8, 9, 47 $\implies$ $RI = 4$

2, 6, 10, 13, 17, 21, 24, 27, 30, 32, 37 $\implies$ $RI = 20$

Varianza y desviación típica

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}$$

Varianza (s^2)

Nos da una magnitud de cómo de lejos están los datos respecto a la media.

Tiene en cuenta todos los datos.

Desviación típica (s)

Es la raíz cuadrada de la varianza.

Coeficiente de variación

$$CV = \frac{s}{|\bar{x}|}$$

Nos da una magnitud de cómo de grande es la desviación típica respecto a la media.

Nos sirve para comparar la dispersión de dos distribuciones distintas.

Es una cantidad adimensional.

Ejercicio

En una clase la altura media es 170 cm,
con una desviación típica de 17 cm.

El peso medio es de 70 kg,
con una desviación típica de 14 kg.

¿Cuál de las dos variables es más dispersa?

Ejercicio

En una clase la altura media es 170 cm,
con una desviación típica de 17 cm.

El peso medio es de 70 kg,
con una desviación típica de 14 kg.

¿Cuál de las dos variables es más dispersa?

No tiene sentido comparar las desviaciones típicas directamente.
Si la altura estuviese en metros, la desviación típica sería de 0,17 m.

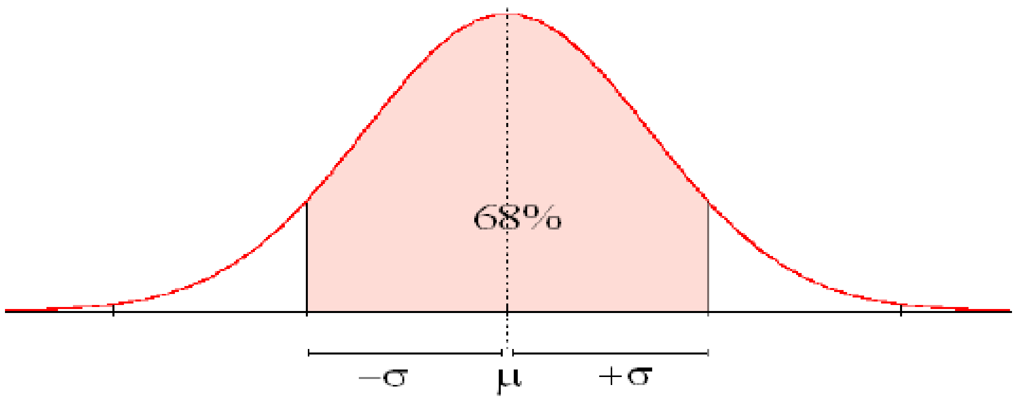
$$CV_{altura} = \frac{17\ cm}{170\ cm} = 0{,}1$$

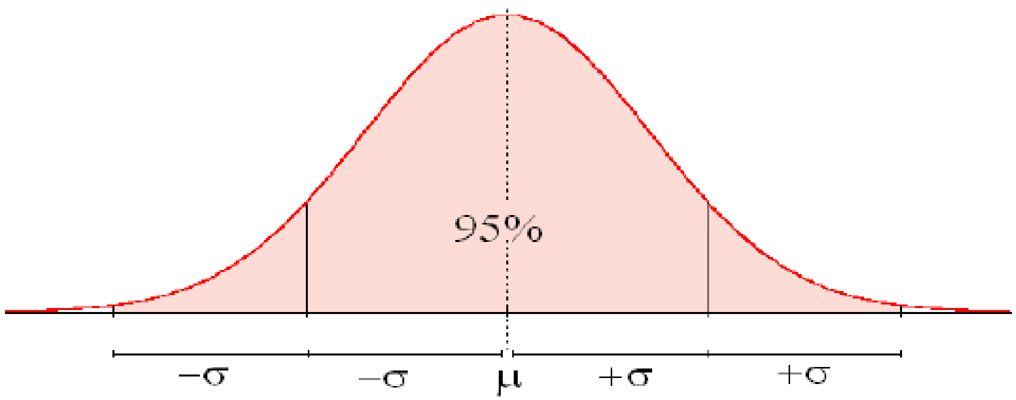
$$CV_{peso} = \frac{14\ kg}{70\ kg} = 0{,}2$$

Más dispersa

Interpretación de la dispersión

La desviación típica nos da una idea de cuán alejados o dispersos están los datos respecto a la media.





Interpretación de la dispersión

Desigualdad de Chebyshev

El intervalo:

$$[\bar{x} - k \cdot s, \quad \bar{x} + k \cdot s,]$$

Incluye al menos al

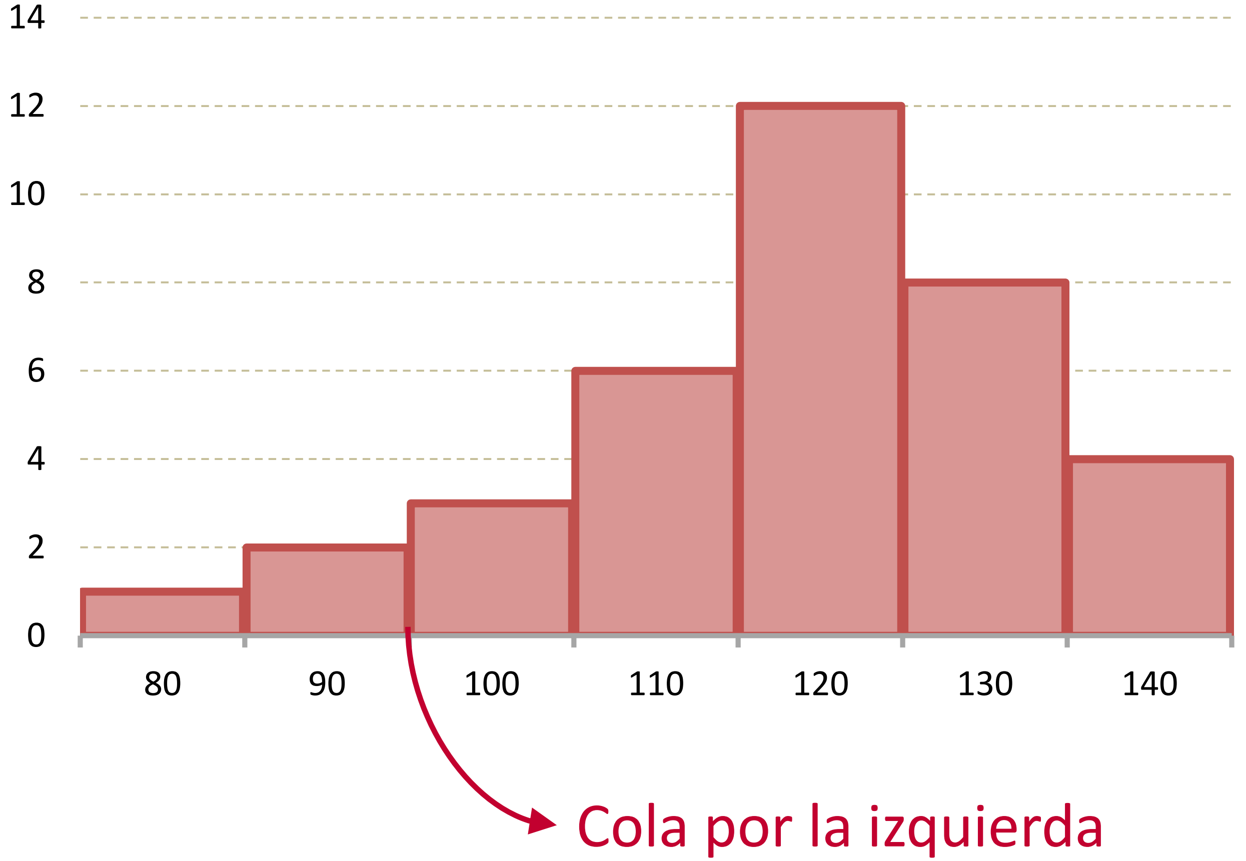
$$\left(1 - \frac{1}{k^2}\right)\%$$

de los datos.

$$\text{Si k=2} \quad \Longrightarrow \quad \textit{Incluye al menos el } 75\% \textit{ de los datos}$$

Estadísticos de forma

Simetría o asimetría

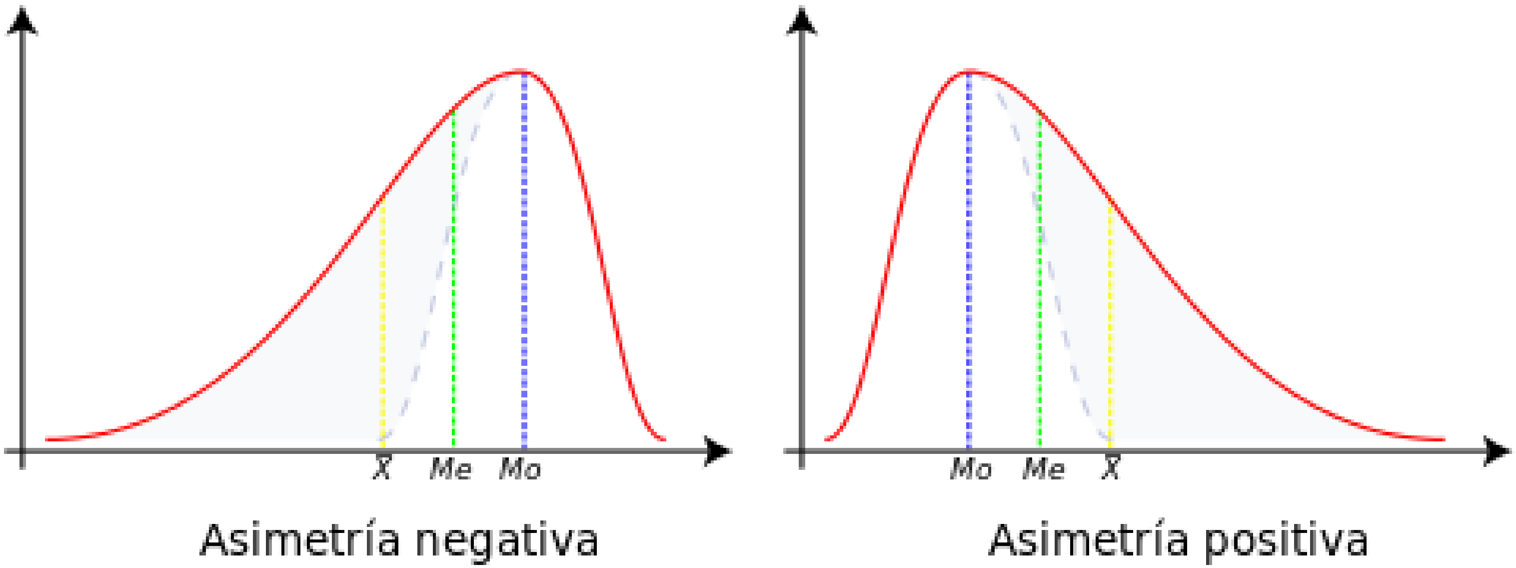


Simetría o asimetría

$$v_1 = \frac{\bar{x} - Mo}{s} \qquad v_2 = \frac{3 \cdot (\bar{x} - Me)}{s}$$

Comparar media con moda y mediana

$$g_1 = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{N}}{s^3}$$

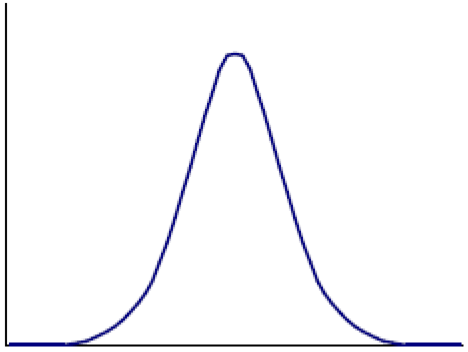


$g_1 < 0$ $\qquad$ $g_1 > 0$

Curtosis

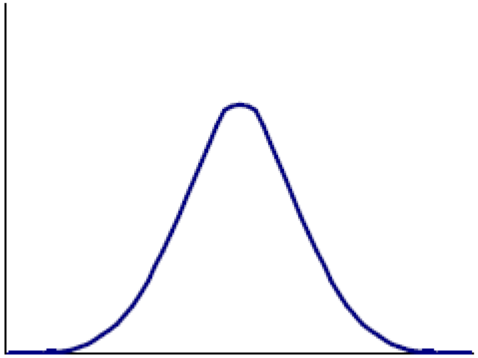
$$g_2 = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{N}}{s^4} - 3$$

Nos da una magnitud de cuán picuda es la distribución.



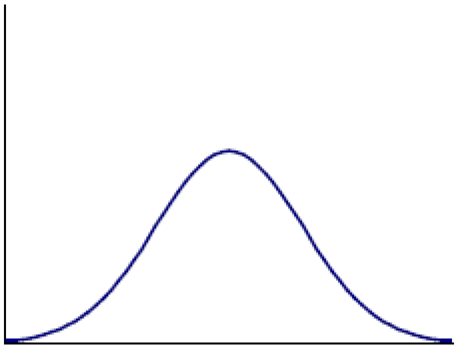
Leptocúrtica

$g_2 > 0$



Mesocúrtica

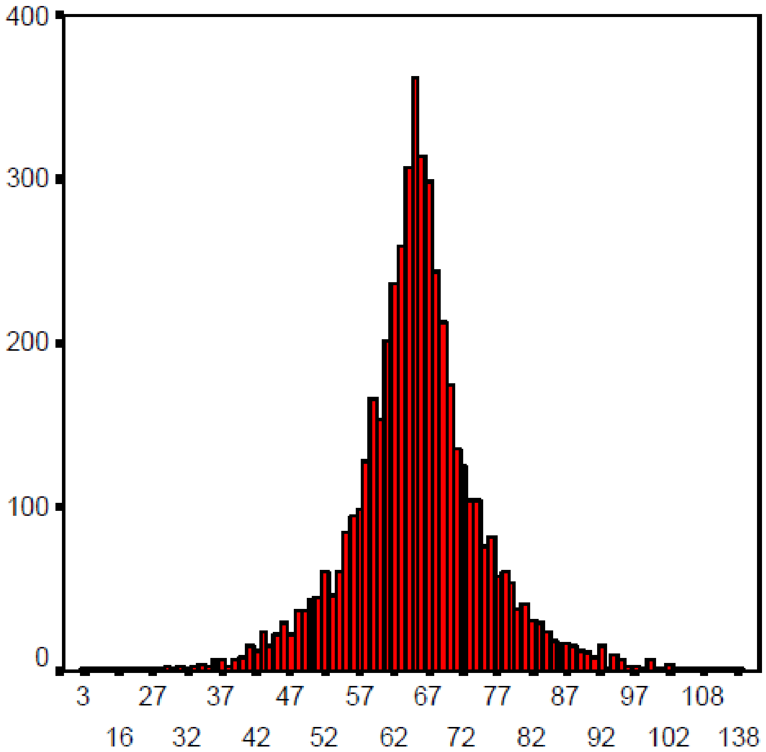
$g_2 = 0$



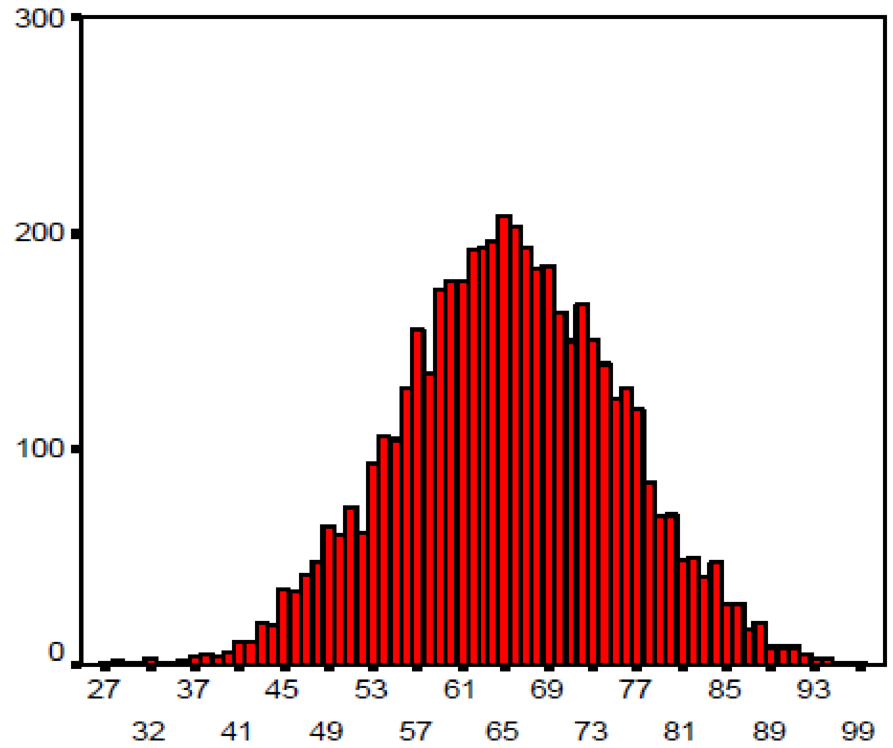
Platicúrtica

$g_2 < 0$

Curtosis



Leptocúrtica



Mesocúrtica

Ambas tienen la misma media y desviación típica.

Tablas de estadísticos

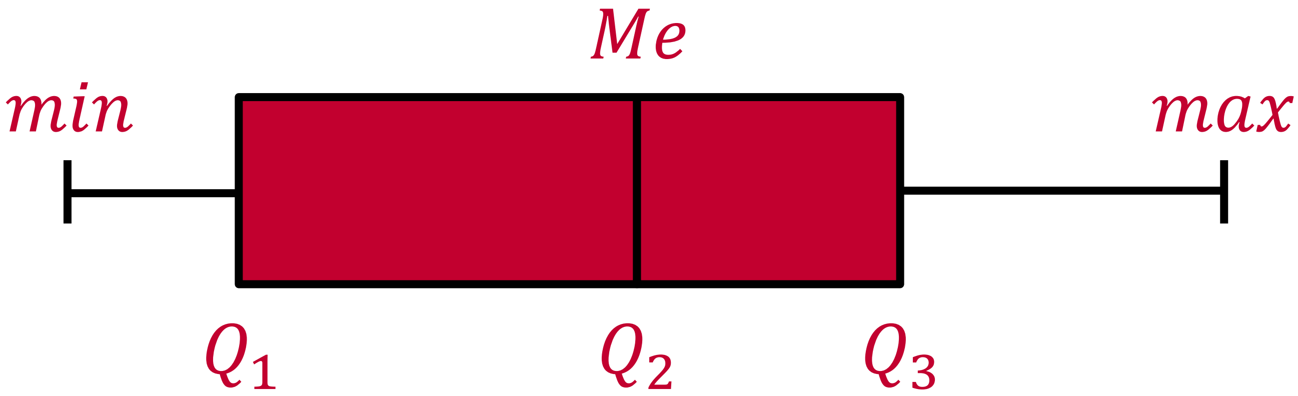
			Estadístico	Error típ.
	Media		1,90	,045
	Intervalo de confianza para la media al 95%	Límite inferior	1,81	
		Límite superior	1,99	
	Media recortada al 5%		1,75	
	Mediana		2,00	
	Varianza		3,114	
	Desv. típ.		1,765	
	Mínimo		0	
	Máximo		8	
	Rango		8	
	Amplitud intercuartil		3,00	
	Asimetría		1,034	,063
	Curtosis		1,060	,126

SPSS

Analizar → Estadísticos descriptivos → Explorar

Diagrama de caja y bigotes

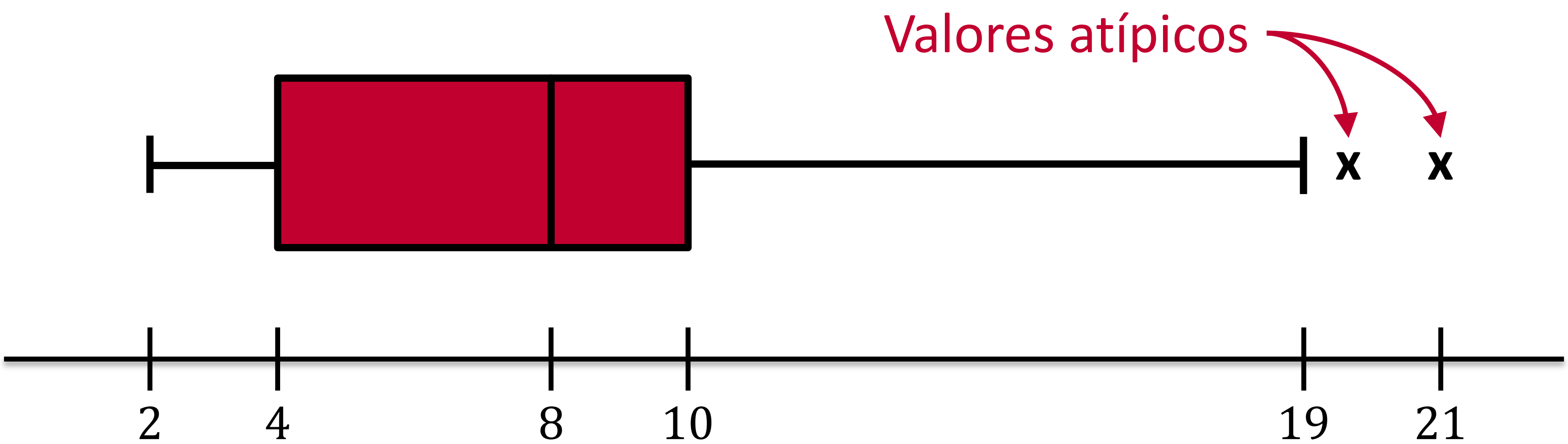
Diagrama de caja y bigotes



SPSS

Analizar → Estadísticos descriptivos → Explorar

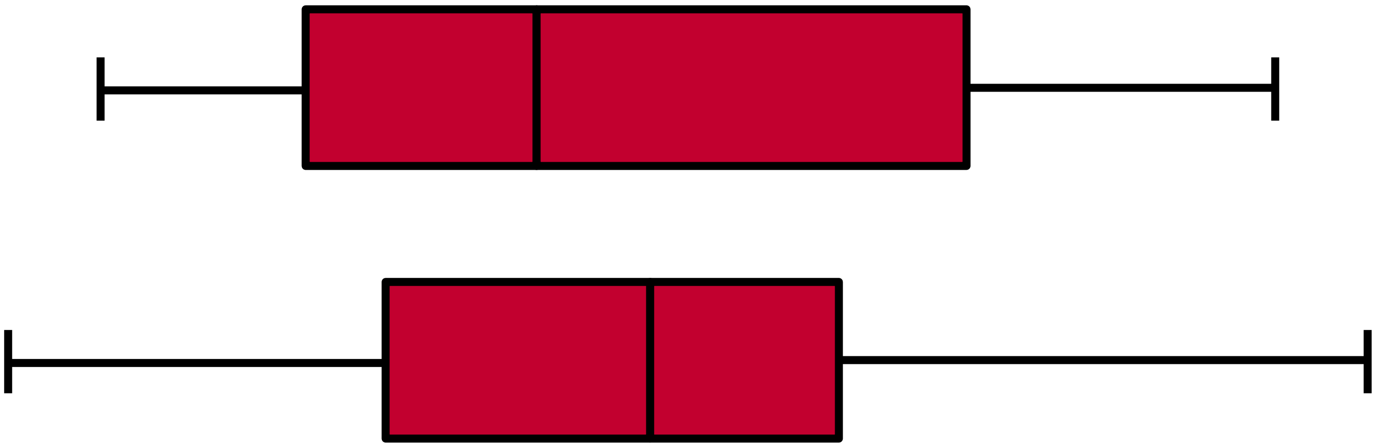
Datos atípicos



El ancho de cada bigote no puede ser mayor que 1,5 veces el ancho de la caja (rango intercuartil).

Los valores que el bigote no abarca son datos atípicos.

Dispersión

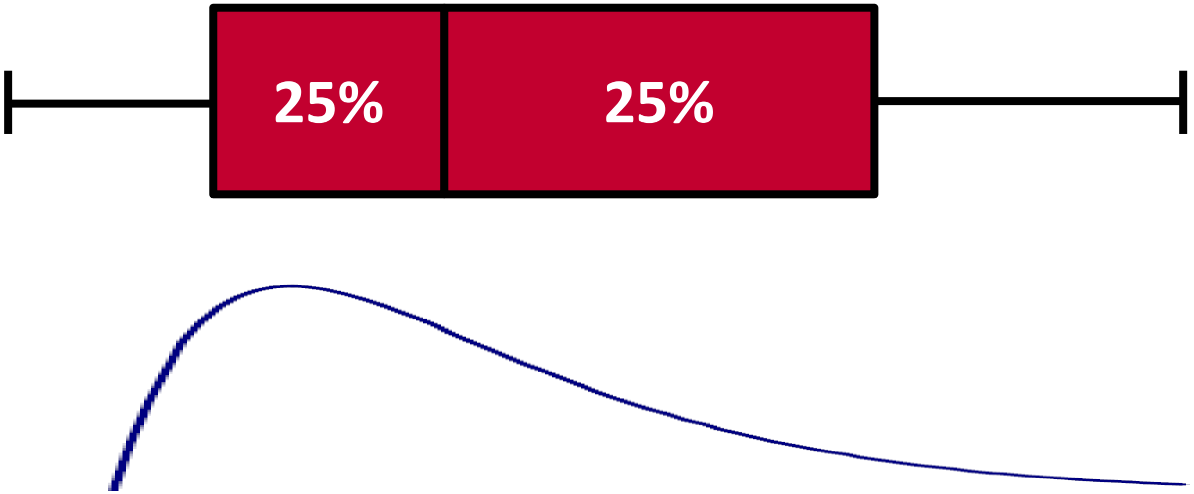


Rango intercuartílico

Podemos comparar el rango intercuartílico
de dos distribuciones comparando el ancho de la caja.

No nos fijamos en los bigotes,
ya que se ven afectados por los datos extremos.

Simetría o asimetría



Para estudiar la simetría nos fijamos en la caja.
No en los bigotes ya que dependen de los valores extremos.

En cada parte de la caja hay la misma cantidad de datos.
Si una parte es más pequeña es porque hay más densidad.

Cómo manipular las estadísticas para que digan lo que tú quieres

(sin llegar a mentir)

Comparar términos absolutos

Madrid, Andalucía y Cataluña concentran la mitad de la recaudación en Sucesiones

Madrid, Cataluña, Andalucía y Comunidad Valenciana concentran el 63% del ahorro en planes de pensiones

Lifestyle

Cataluña, Madrid y Andalucía, las comunidades que más viajan el fin de semana

Madrid, Cataluña y Andalucía, las comunidades más activas en comercio electrónico

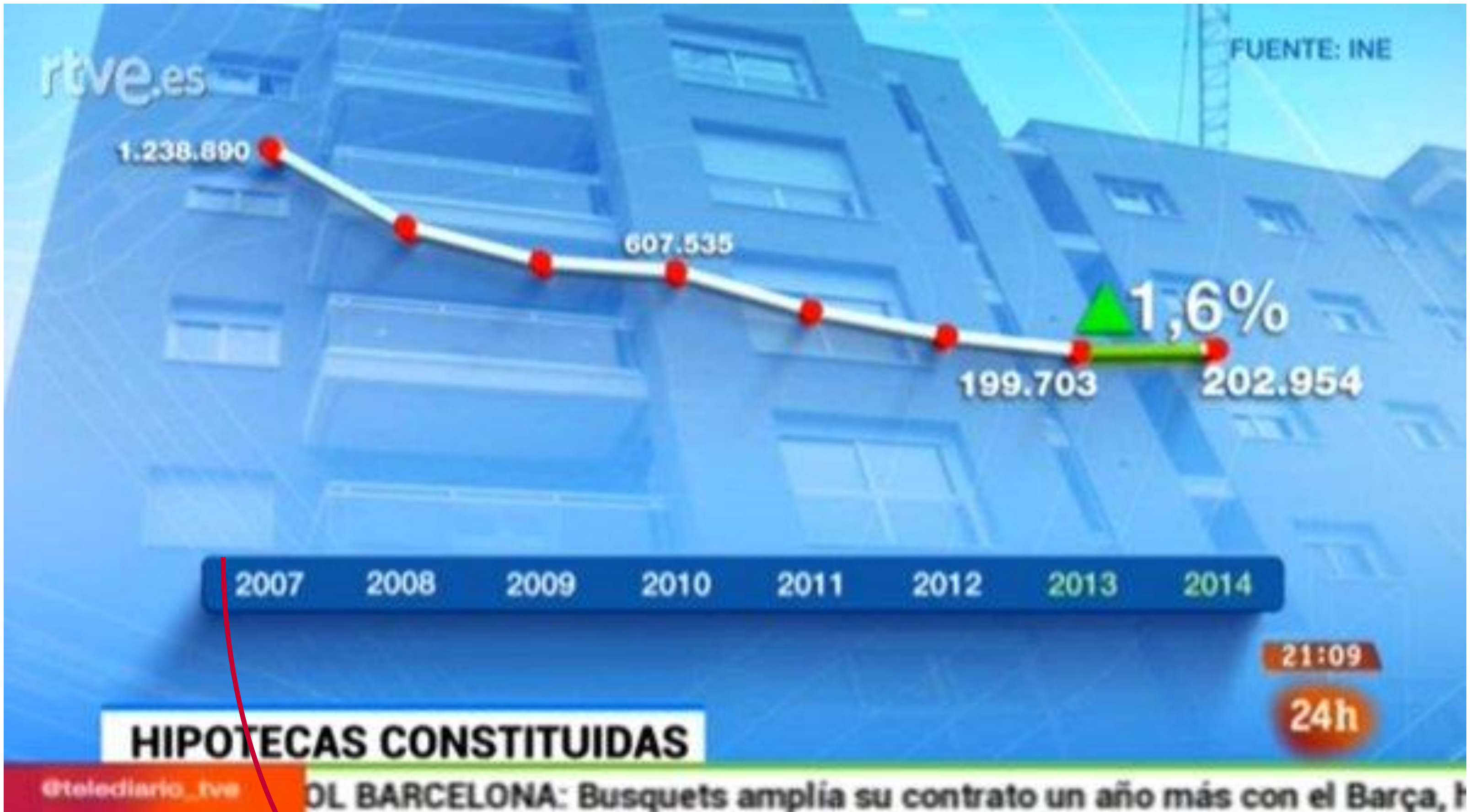
Cataluña, Madrid y Andalucía lideran el gasto absoluto en sanidad

Madrid, Cataluña y Andalucía, donde más dinero devuelve Hacienda

Eje al antojo y dos únicos valores

Con el eje puesto en un valor arbitrario

Poner el eje al antojo



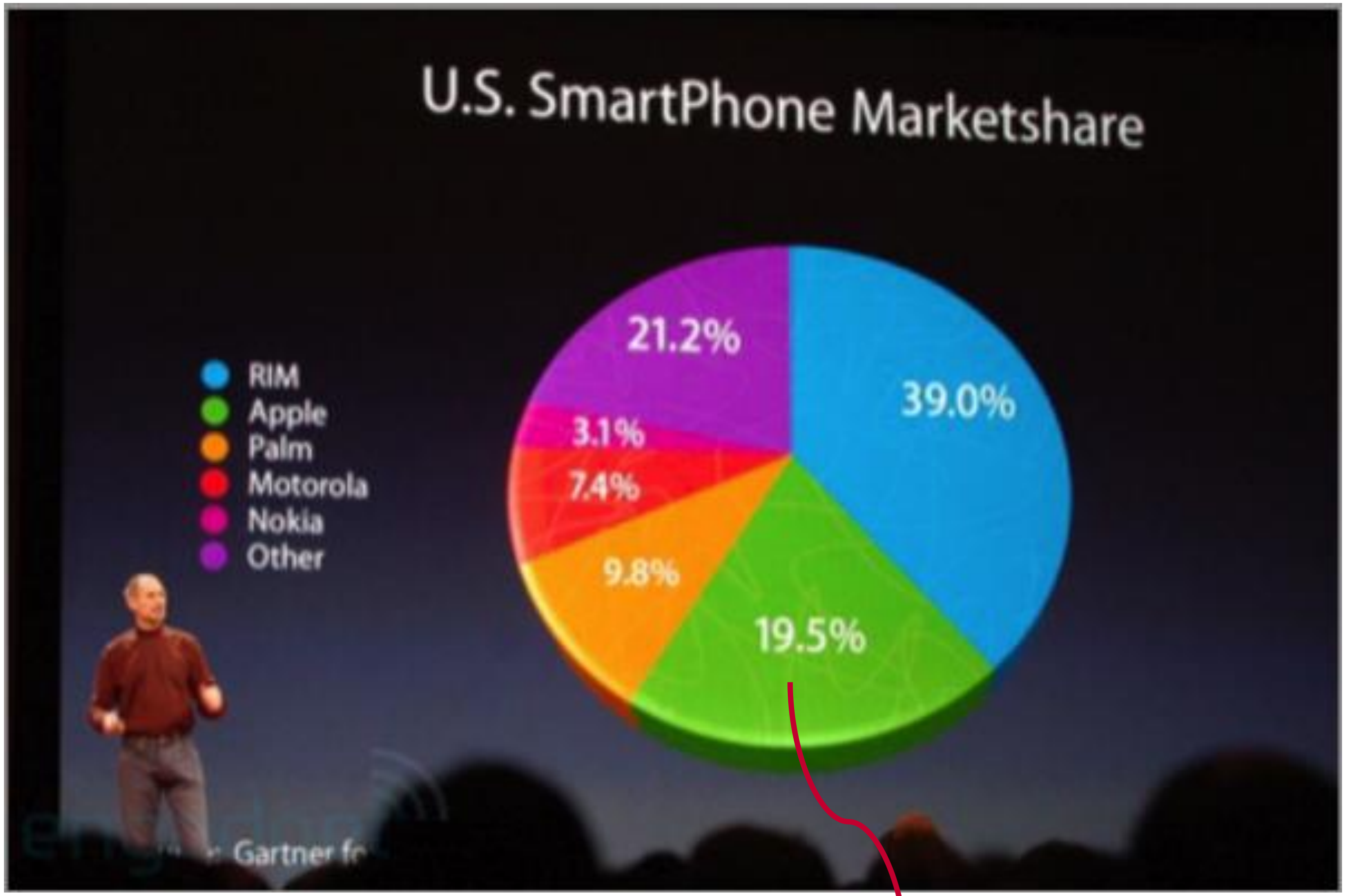
Con el eje en valores negativos

Eje al antojo y dos únicos valores

Eje al antojo y dos únicos valores

Con el eje en el 49%

Usar el 3D para distorsionar



El verde es más pequeño que el morado

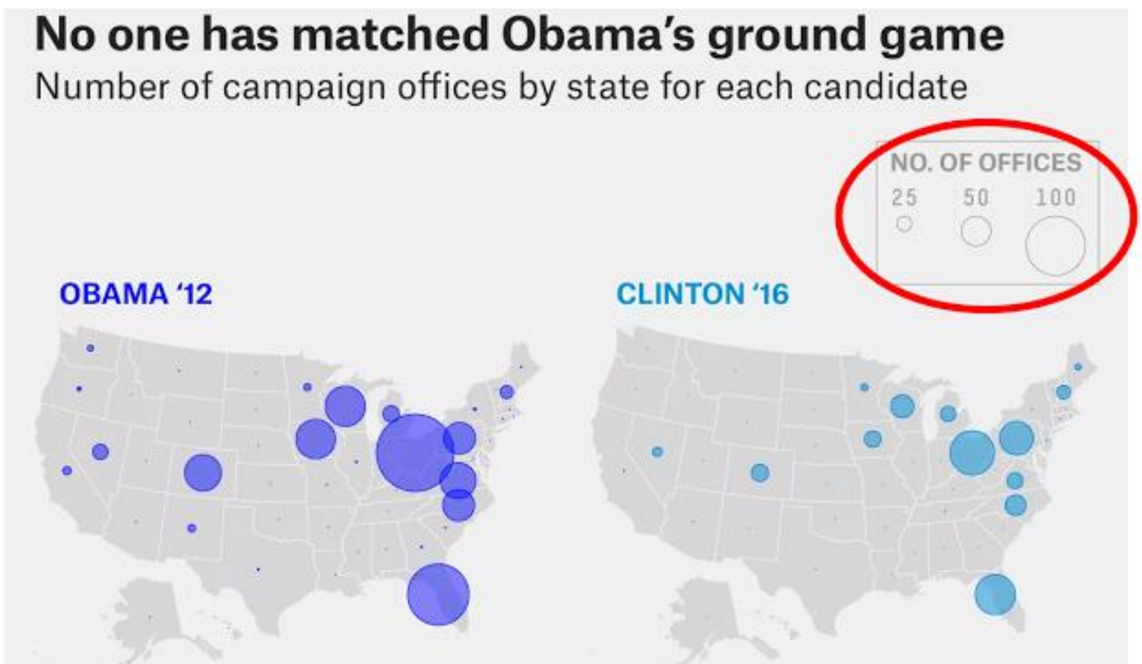
Cómo manipular las estadísticas para que digan lo que tú quieres

(con ‘errores’)

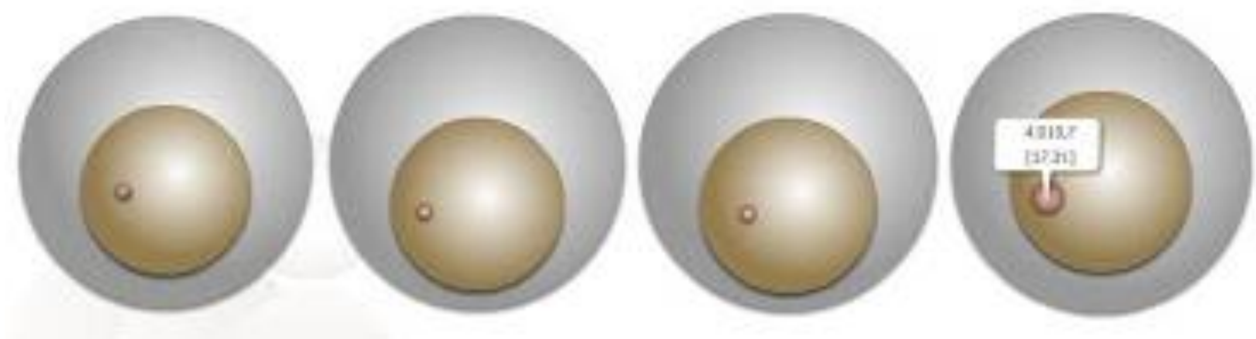
Los datos no se corresponden

Los datos no se corresponden

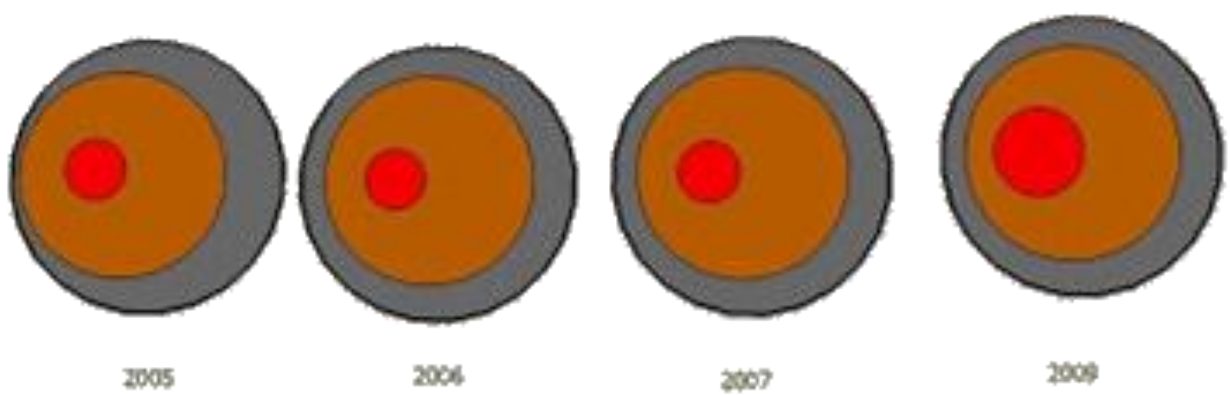
Comparar radios y no áreas



Comparar radios y no áreas

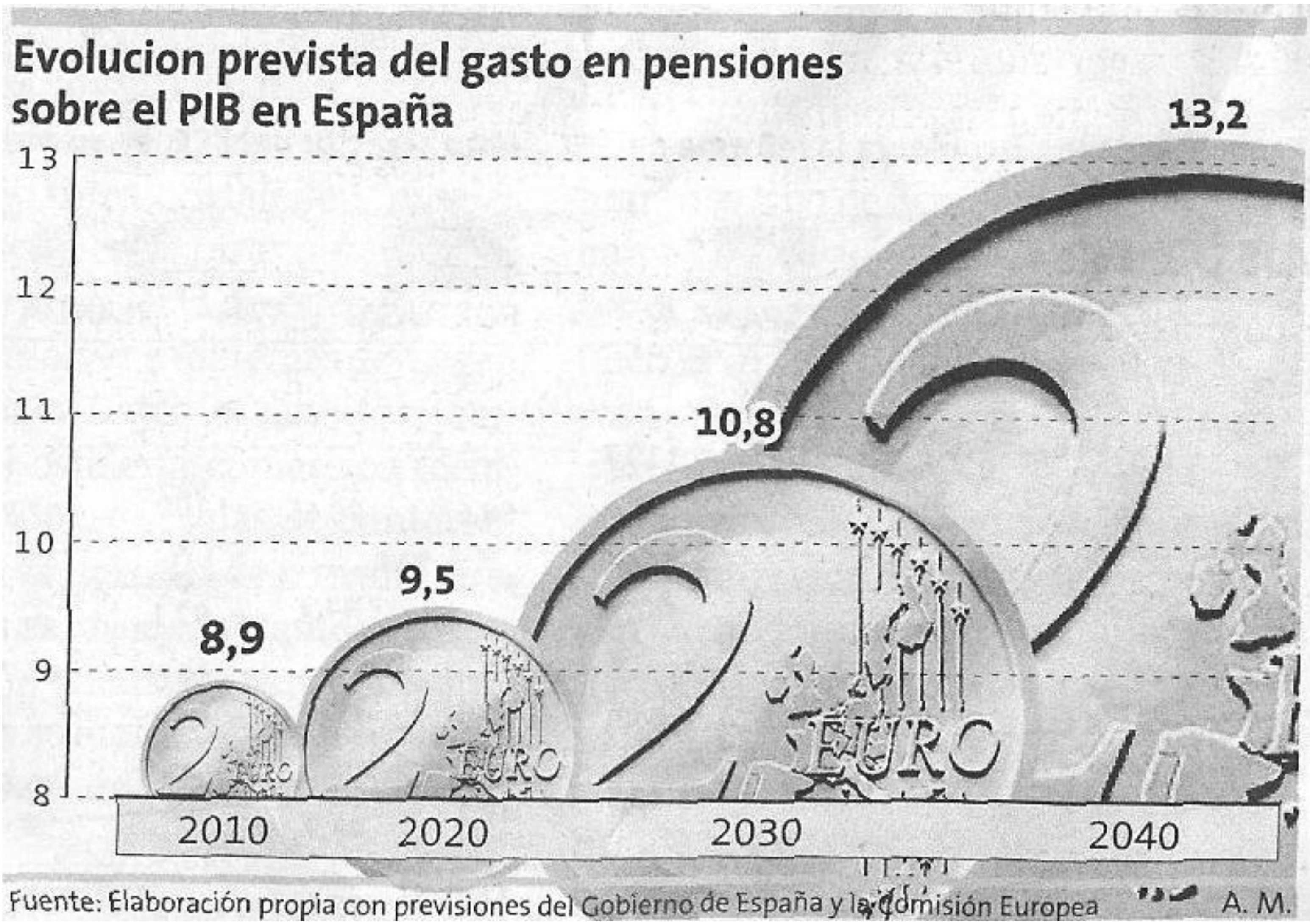


Bolas: población, población activa y parados

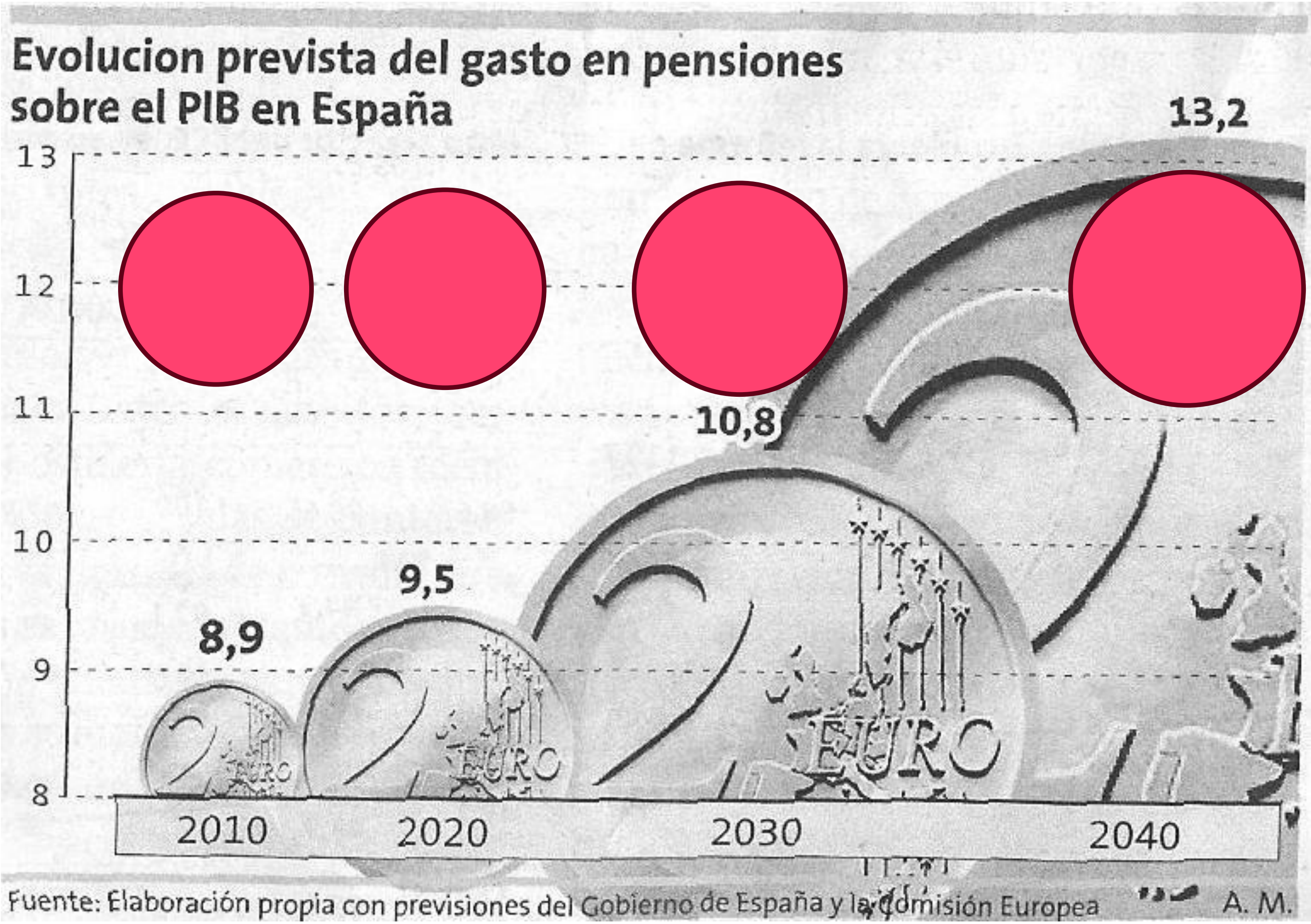


Hacer gráficos con el Paint

Hacer gráficos con el Paint



Hacer gráficos con el Paint



Hacerlo todo mal

1. Los años están descolocados (el último es el central).
2. La botella central está desalineada con las otras dos.
3. Realmente la botella central es más grande que las otras dos.
4. Las botellas se llenan en torno al 20%.
5. El titular dice lo contrario que los datos: en el último año en ningún país ha subido el consumo.