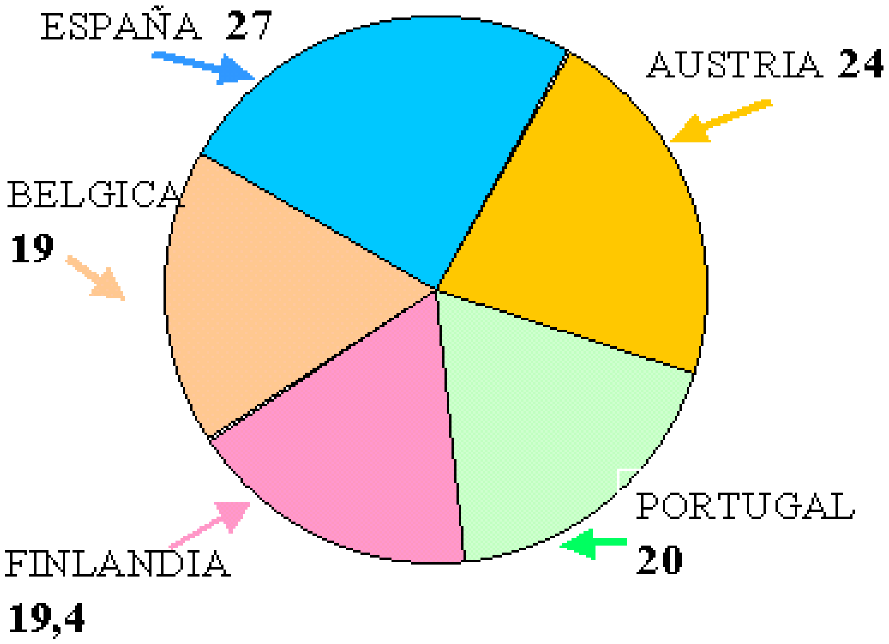
Tablas y gráficas estadísticas (I)

España es el país de Europa con un mayor número de donantes de órganos.

En la siguiente **tabla estadística** se muestran los cinco países con mayor número de donantes:

País	Nº de donantes por millón de hab.
España	27
Austria	24
Portugal	20
Finlandia	19,4
Bélgica	19

La misma información se resume en el correspondiente **gráfico de sectores.**



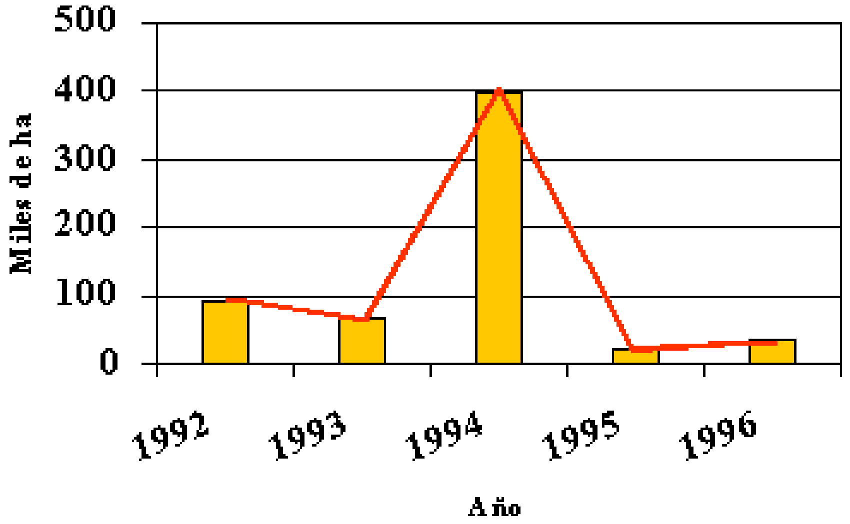
Tablas y gráficas estadísticas (II)

Los incendios forestales en España durante el año 1996 descendieron espectacularmente.

La tabla siguiente muestra la superficie total quemada, en hectáreas, en el quinquenio 1992–1996.

Año	Superficie quemada (en ha)
1992	93 243
1993	67 499
1994	399 626
1995	23 093
1996	35 100

La misma información se muestra en el **diagrama de barras** correspondiente.



Uniendo los extremos de las barras obtenemos el **polígono de frecuencias.**

Agrupación de datos en intervalos

160, 167, 163, 148, 151, 158, 166, 166, 157, 153, 151, 150, 155, 164, 162, 166, 171, 167, 165, 152, 150, 147, 152, 162, 155, 158, 158, 164, 157, 155, 160, 154, 153, 156, 160, 159, 159, 158, 163, 161

Los 40 datos anteriores pueden agruparse formando intervalos o clases.

Por ejemplo, en el intervalo [150–155) hay 9 datos: se incluye 150, se excluye 155.

A continuación se forma la tabla:

Datos	Frecuencia absoluta
[145–150)	2
[150–155)	9
[155–160)	12
[160–165)	10
[165–170)	6
[170–175)	1

Si el número de datos es grande conviene agruparlos en **intervalos o clases**.
Todas las clases deben tener la misma **amplitud**.
Los puntos medios de cada clase se llaman **marcas de clase**.

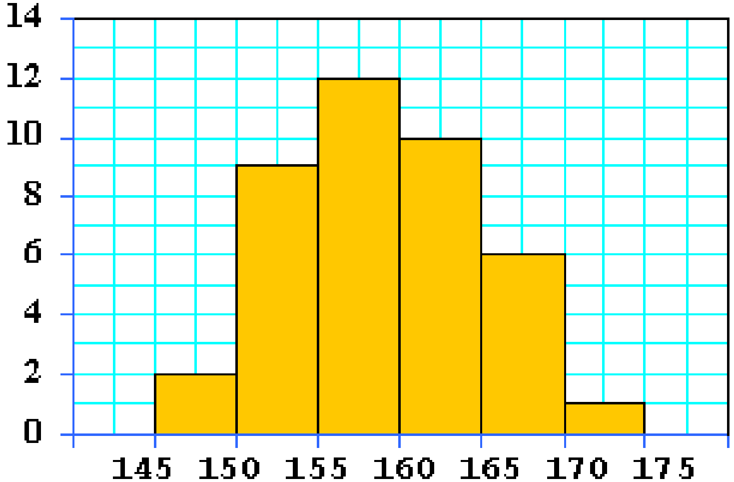
La marca de clase del intervalo [150–155) es

$$\frac{150+155}{2} = 152{,}5$$

Histograma y polígono de frecuencias

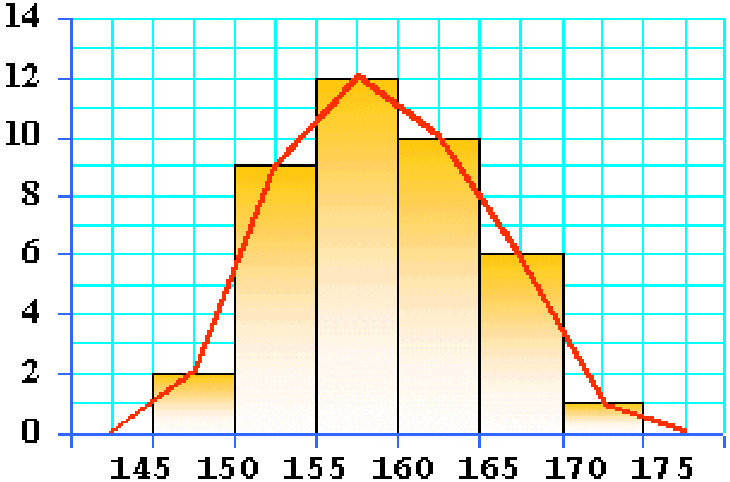
Para representar gráficamente la situación dada por la tabla estadística anterior, construimos el gráfico de la figura 1.

Figura 1



Este gráfico se llama **histograma** de frecuencias absolutas.

Figura 2



Uniendo los puntos medios de los lados superiores se obtiene el **polígono de frecuencias absolutas**.

Media aritmética. Datos no agrupados

Ejemplo. Los goles marcados por un equipo de fútbol en 20 partidos vienen dados en las dos primeras filas de la siguiente tabla:

Nº de goles	Nº de partidos	Datos x Frecuencia
0	6	0
1	4	4
2	4	8
3	3	9
4	2	8
5	1	5
Total	20	34

Para calcular la media aritmética:

1.º Se multiplican los datos por sus frecuencias absolutas respectivas, y se suman.
2.º El resultado se divide por el total de datos.

Datos × frecuencias

$$\text{Media} = \frac{34}{20} = 1{,}7$$

Total de datos

Media aritmética. Datos agrupados en intervalos

Los pesos de 35 niños al nacer vienen dados en las filas primera y tercera de la siguiente tabla:

Peso (en kg)	Marca de clase (Datos)	Nº de niños (Frecuencia absoluta)	Datos x Frecuencia absoluta
[1,5–2,0)	1,75	1	1,75
[2,0–2,5)	2,25	3	6,75
[2,5–3,0)	2,75	9	24,75
[3,0–3,5)	3,25	12	39
[3,5–4,0)	3,75	6	22,5
[4,0–4,5)	4,25	4	17
TOTAL		35	111,75

Para calcular la media aritmética:

1.º Se multiplican los datos (marcas de clase) por sus frecuencias absolutas respectivas, y se suman.
2.º El resultado se divide por el total de datos.

El peso medio es:

$$\frac{111{,}75}{35} = 3{,}19 \text{ kg}$$

La moda

La **moda** de un conjunto de datos es el dato que tiene mayor frecuencia.

Ejemplo 1. Las ventas de calzado de unos grandes almacenes son:

Nº de calzado	38	39	40	41	42	43	44	45
Nº de pares vendidos	16	21	30	35	29	18	10	7

El número de calzado más frecuente es el 41.

Lo compran 35 personas

La moda es 41.

Si los datos se encuentran agrupados en intervalos tomaremos como moda la marca de clase que tiene mayor frecuencia (**clase modal**).
La moda puede no existir (si todos los datos tiene igual frecuencia), puede ser única (unimodal), tener dos modas (bimodal), etc.

La mediana

La **mediana** de un conjunto de datos es un valor tal que el número de datos menores que él es igual al número de datos mayores que él.

Ejemplo 1: Las tallas, en centímetros, de 11 patinadores son:

169, 150, 162, 155, 157, 153, 164, 153, 170, 167, 172

1º. Ordenamos los datos:

150, 153, 153, 155, 157, **162**, 164, 167, 169, 170, 172

2º. El dato que queda en el centro es 162.

La mediana es 162 cm.

Ejemplo 2: **Si el número de datos fuese par**, como sucedería si se incorpora otra patinadora que mide 171 cm, **la mediana es la media aritmética de los dos valores centrales.**

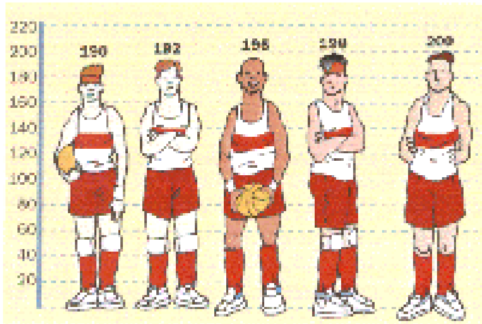
150, 153, 153, 155, 157, **162**, **164**, 167, 169, 170, 171, 172

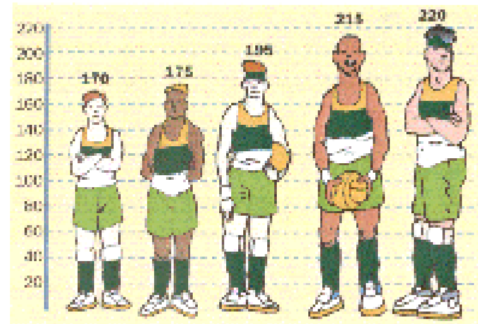
La mediana es: $\dfrac{162+164}{2}=163\ \text{cm}$

Rango o recorrido

Dos equipos de baloncesto tienen las siguientes tallas en cm:

Equipo A:
190
192
195
198
200





Equipo B:
170
175
195
215
220

Ambos equipos tienen la misma media:

$$\frac{190+192+195+198+200}{5}=195 \text{ cm} \qquad \frac{170+175+195+215+220}{5}=195 \text{ cm}$$

En el equipo A las tallas están muy próximas a la media: es **poco disperso**.

En el equipo B las tallas están muy separadas de la media: es **muy disperso**.

La medida de dispersión más sencilla es el **rango o recorrido**, que es la diferencia entre el mayor y el menor valor de los datos.

Rango de A = 200 – 190 = 10 cm.

Rango de B = 220 – 170 = 50 cm.

Desviación respecto a la media

Para los dos equipos de baloncesto se tiene:

EQUIPO A	
Datos	Datos – Media
190	190 – 195 = –5
192	192 – 195 = –3
195	195 – 195 = 0
198	198 – 195 = 3
200	200 – 195 = 5

Media = 195

EQUIPO B	
Equipo B	Datos – Media
170	170 – 195 = –25
175	175 – 195 = –20
195	195 – 195 = 0
215	215 – 195 = 20
220	220 – 195 = 25

Media = 195

¿En qué equipo se obtiene diferencias mayores, en A o en B?

En el B.

Las diferencias entre cada dato y la media aritmética del grupo se llaman **desviaciones respecto a la media**.

Estas diferencias pueden ser positivas, negativas o nulas.

Desviación media

Completamos la tabla anterior con una columna que exprese el valor absoluto de las desviaciones.

EQUIPO A		
Datos	Datos – Media	|Datos – Media|
190	190 – 195 = –5	5
192	192 – 195 = –3	3
195	195 – 195 = 0	0
198	198 – 195 = 3	3
200	200 – 195 = 5	5
Sumas:	0	16

EQUIPO B		
Equipo B	Datos – Media	|Datos – Media|
170	170 – 195 = –25	25
175	175 – 195 = –20	20
195	195 – 195 = 0	0
215	215 – 195 = 20	20
220	220 – 195 = 25	25
Sumas:	0	90

La media de los valores absolutos de las desviaciones es:

$$\frac{16}{5} = 3{,}2 \qquad \frac{90}{5} = 18$$

Las tallas del equipo B están más dispersas o menos concentradas que las del equipo A. (La desviación media de B es mayor que la de A)

Desviación media de un conjunto de datos es la media aritmética de los valores absolutos de las desviaciones respecto a la media.

Resolución de problemas

PROBLEMA

La edad media de las siete primeras personas que acudieron a la fiesta de cumpleaños del abuelo de Olga es de 21 años. Después llegaron Luis y Ana, y la edad media creció a 23 años. Y al llegar el abuelo de Olga, la edad media fue de 29 años. ¿Qué edad tiene el abuelo de Olga?

Leer atentamente el enunciado

Con 7 personas la edad media es 21 años.
Con 9 personas la edad media es 23 años.
Con 9 personas y el abuelo la edad media es 29 años.

Hay que hallar la edad de abuelo.

¿Qué se puede deducir de los datos?

Si la edad media de 7 personas es 21, suman en total: $7 \cdot 21 = 147$ años.
Si la edad media de 9 personas es 23, suman en total: $9 \cdot 23 = 207$ años.

Representamos por x la edad del abuelo

Realizar los cálculos

$$\left.\begin{array}{l}\text{Nueve personas suman: 207 años.}\\ \text{El abuelo tiene x años.}\end{array}\right\}\ \text{Entre las 10 personas, suman: } 207 + x$$

Su edad media es 29 años: $\longrightarrow \dfrac{207+x}{10} = 29$

$\Rightarrow$ $207 + x = 290$ $\Rightarrow$ $x = 83$ años.