

12 Estadística

ANALIZA Y SACA CONCLUSIONES

¿Cómo interpretas la definición de estadística que da el autor en la primera estrofa? ¿Cómo definirías tú lo que es la estadística?

Como una herramienta que permite hacer una estimación de las características de una población.

La estadística es una rama de las matemáticas que se encarga de estudiar determinadas características de una población y que nos permite extraer conclusiones sobre la misma.

¿Qué es el promedio? ¿Hay situaciones en las que no tenga sentido calcular el promedio? Pon algún ejemplo.

El promedio es el cociente entre la suma de diferentes valores y la cantidad de valores que se suman.

Carece de sentido considerar el promedio cuando estudiamos valores no numéricos. Por ejemplo, si preguntamos por el color del pelo de unos alumnos de 2.º de ESO, las posibles respuestas son *moreno*, *castaño*, *pelirrojo* o *rubio*. En este caso no tiene sentido calcular el promedio.

REFLEXIONA Y PON EN COMÚN

En los estudios y datos que proporciona la estadística, ¿juega un papel importante la interpretación que se haga de ellos? ¿Muestran la realidad fielmente?

La interpretación que se hace de los datos juega un papel relevante. Dependiendo del ámbito en que se realice el estudio, una visión sesgada de los resultados del mismo puede hacer que se obtengan interpretaciones casi opuestas.

Por ejemplo, las estadísticas afirman que el paro se ha reducido a 4 000 000 de desempleados. Para el Gobierno es un gran dato, ya que gracias a sus reformas están consiguiendo reducir el paro. Sin embargo, para la oposición, lo que muestra esta estadística es que tenemos un paro enorme y que las medidas del Ejecutivo no están teniendo el resultado prometido.

Y TÚ, ¿QUÉ OPINAS?

Para “hacer la cuenta general”, la estadística precisa recoger gran cantidad de datos. En la actualidad, el uso de internet y las redes sociales ha favorecido que esta recogida de datos sea masiva. ¿Crees que es realista utilizar las redes sociales para recoger datos? ¿Son fiables los datos recogidos? Debate con tus compañeros.

En mi opinión, los datos recogidos de las redes sociales no son fiables. Los perfiles falsos, la creación de múltiples perfiles por parte de un único usuario o el anonimato son factores que hacen que los datos que se recogen en las redes sociales no resulten muy realistas y carezcan de la seguridad suficiente para obtener conclusiones fiables.

Actividades propuestas

1. **Antes de sacar al mercado un nuevo zumo, se selecciona a 500 personas para estudiar su aceptación. ¿Cuáles son la población, la muestra y los individuos del estudio?**

La población es el conjunto de las personas entre las que se seleccionan aquellas sobre las que se va a estudiar la aceptación del zumo. La muestra son las 500 personas seleccionadas para el estudio, y los individuos son cada una de las personas que realizan el estudio.

2. **Indica si la muestra que se ha seleccionado en cada caso es representativa y de qué tipo es la variable estudiada:**

a) **Para conocer el número de SMS que se envían al día en un barrio se ha preguntado a 100 personas menores de 25 años y a 20 mayores de 25.**

b) **Para conocer los gustos musicales de los alumnos de un instituto se ha preguntado a 10 alumnos de cada curso.**

a) No se trata de una muestra representativa puesto que se seleccionan para el estudio cinco veces más de personas de la población menores de 25 años que mayores de 25. Es una variable estadística cuantitativa discreta.

b) Es una muestra representativa porque se ha seleccionado el mismo número de alumnos en todos los cursos. Se trata de una variable estadística cualitativa.

3. Una asociación ecologista ha realizado una encuesta sobre el medio de transporte que utiliza la población habitualmente para ir a trabajar. Las preguntas son:
1. Sexo
 2. Edad
 3. Medio de transporte habitual (coche, metro, bus, bici, etc.)
 4. Distancia de casa al lugar de trabajo
 5. Tiempo de viaje en horas

Indica el tipo de cada una de las variables estudiadas.

Cualitativa: 1, 3

Cuantitativa discreta: 2, 5

Cuantitativa continua: 4

4. Se ha preguntado a 25 personas cuántas veces han ido al cine en el último mes. Las respuestas han sido las siguientes:

1, 0, 2, 1, 0, 1, 1, 2, 3, 0, 2, 1, 1, 2, 0, 4, 2, 3, 4, 1, 3, 2, 3, 0, 1

- a) ¿De qué tipo es la variable?
 - b) Elabora una tabla de frecuencias.
 - c) ¿Cuál es el valor de la variable más frecuente?
- a) Variable cuantitativa discreta.
- b)

x_i	f_i	h_i	F_i	H_i
0	5	0,20	5	0,20
1	8	0,32	5 + 8 = 13	0,52
2	6	0,24	13 + 6 = 19	0,76
3	4	0,16	19 + 4 = 23	0,92
4	2	0,08	23 + 2 = 25	1
$N = 25$		1		

- c) El valor de la variable más frecuente es 1.

5. El número de portátiles Normal, Premium y Supra reparados en los últimos siete días por un servicio técnico son:

*N, P, P, S, P, S, P, N, S, S, S, N, N, P,
P, N, S, P, P, S, N, P, N, P, P, N, P, N, N, S*

- a) ¿De qué tipo es la variable estudiada?
 - b) Presenta los datos en una tabla de frecuencias.
 - c) ¿Qué ordenadores parecen ser los más defectuosos?
- a) Variable cualitativa.
- b)

Tipo de portátil	f_i	h_i	F_i	H_i
Normal	10	0,33	10	0,33
Premium	12	0,40	10 + 12 = 22	0,73
Supra	8	0,27	22 + 8 = 30	1
$N = 30$		1		

- c) Los ordenadores más defectuosos son los de la clase Premium.

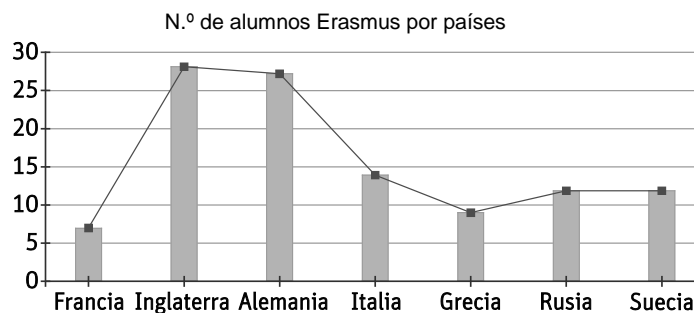
6. En la siguiente tabla se recoge el país de procedencia de los alumnos Erasmus que han llegado a una universidad de la ciudad:

País	f_i
Francia	7
Inglaterra	28
Alemania	27
Italia	14
Grecia	9
Rusia	12
Suecia	12

- a) Construye la tabla de frecuencias.
 b) Representa estos datos mediante un diagrama de barras y un polígono de frecuencias.
 a)

País	f_i	h_i	F_i	H_i
Francia	7	0,064	7	0,064
Inglaterra	28	0,257	7 + 28 = 35	0,321
Alemania	27	0,248	35 + 27 = 62	0,569
Italia	14	0,128	62 + 14 = 76	0,697
Grecia	9	0,083	76 + 9 = 85	0,780
Rusia	12	0,110	85 + 12 = 97	0,890
Suecia	12	0,110	97 + 12 = 109	1
$N = 109$		1		

b)



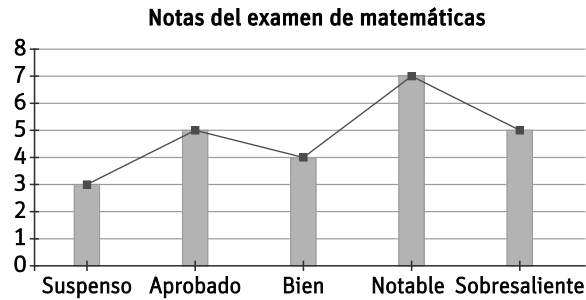
7. Las 24 notas del último examen de matemáticas realizado por los alumnos de 2.º de ESO han sido las siguientes:

Suspense: 3 Aprobado: 5 Bien: 4
 Notable: 7 Sobresaliente: 5

- a) Construye la tabla de frecuencias asociada.
 b) Representa la distribución en un diagrama de barras y en un polígono de frecuencias.
 c) Representa los datos en un diagrama de sectores.
 a)

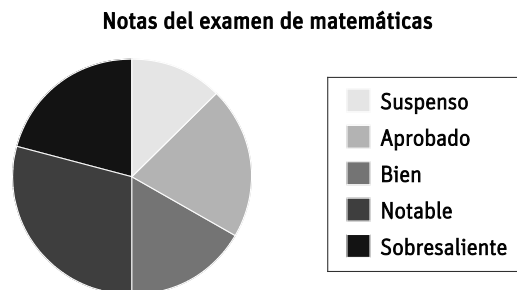
Nota	f_i	h_i	F_i	H_i
Suspense	3	0,125	3	0,125
Aprobado	5	0,208	3 + 5 = 8	0,333
Bien	4	0,167	8 + 4 = 12	0,500
Notable	7	0,292	12 + 7 = 19	0,792
Sobresaliente	5	0,208	19 + 5 = 24	1
$N = 24$		1		

b)



c)

Nota	f _i	h _i	Amplitud del sector
Suspenso	3	0,125	$0,125 \cdot 360^\circ = 45^\circ$
Aprobado	5	0,208	$0,208 \cdot 360^\circ = 74,88^\circ$
Bien	4	0,167	$0,167 \cdot 360^\circ = 60,12^\circ$
Notable	7	0,292	$0,292 \cdot 360^\circ = 105,12^\circ$
Sobresaliente	5	0,208	$0,208 \cdot 360^\circ = 74,88^\circ$
	N = 24	1	360°



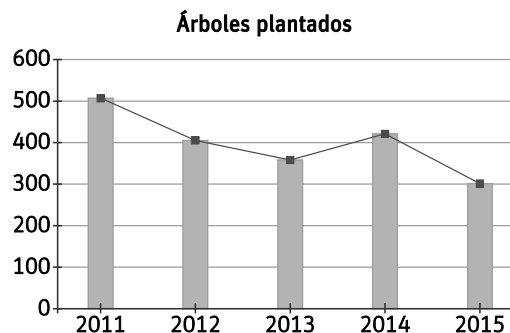
8. Con el fin de preservar la riqueza de nuestros bosques, se ha llevado a cabo una política de plantación de árboles en los últimos años:

Año	2011	2012	2013	2014	2015
N.º de árboles	510	405	356	421	298

a) Representa estos datos mediante un diagrama de barras y un polígono de frecuencias.

b) Construye el diagrama de sectores de los datos asociados.

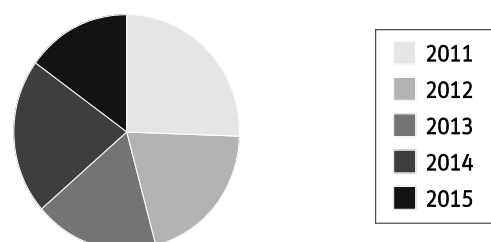
a)



b)

Árboles	f _i	h _i	Amplitud del sector
2011	510	0,256	$0,256 \cdot 360^\circ = 92,16^\circ$
2012	405	0,204	$0,204 \cdot 360^\circ = 73,44^\circ$
2013	356	0,179	$0,179 \cdot 360^\circ = 64,44^\circ$
2014	421	0,212	$0,212 \cdot 360^\circ = 76,32^\circ$
2015	298	0,150	$0,150 \cdot 360^\circ = 54^\circ$
	1990	1	

Árboles plantados



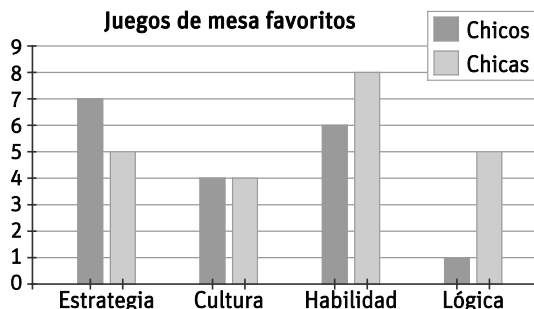
9. La siguiente tabla muestra los datos recogidos al preguntar a 40 alumnos sobre su tipo de juego de mesa favorito:

Juego de mesa	Chicas	Chicos
Estrategia	5	7
Cultura	4	4
Habilidad	8	6
Lógica	5	1

a) Haz un diagrama de barras, con diferentes barras para chicos y chicas.

b) Representa los resultados totales en un diagrama de sectores.

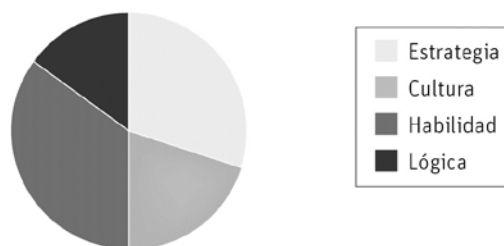
a)



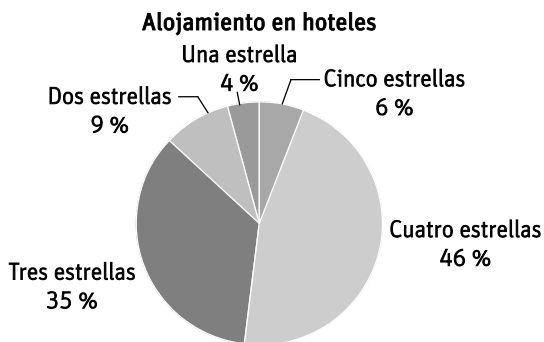
b) Realizamos una nueva tabla con los datos totales:

Juego de mesa	Total	h_i	Amplitud del sector
Estrategia	12	0,30	$0,3 \cdot 360^\circ = 108^\circ$
Cultura	8	0,20	$0,2 \cdot 360^\circ = 72^\circ$
Habilidad	14	0,35	$0,35 \cdot 360^\circ = 126^\circ$
Lógica	6	0,15	$0,15 \cdot 360^\circ = 54^\circ$

Juegos de mesa favoritos



10. Los 92 000 turistas que el año pasado visitaron nuestras playas se alojaron en hoteles según este gráfico.



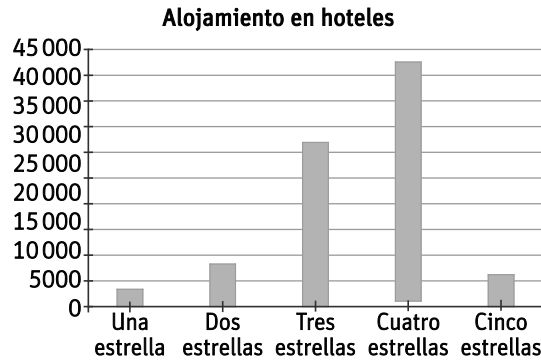
a) Construye la tabla de frecuencias.

b) Representa los datos en un diagrama de barras.

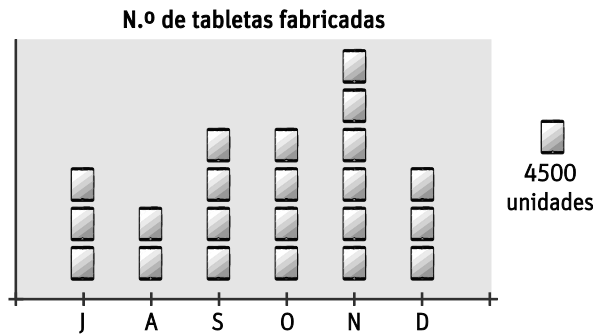
a)

Hotel	f_i	h_i	F_i	H_i
Una estrella	4 % de 92 000 = 3680	0,04 (4 %)	3680	0,04
Dos estrellas	9 % de 92 000 = 8280	0,09 (9 %)	11 960	0,13
Tres estrellas	35 % de 92 000 = 32 200	0,35 (35 %)	44 160	0,48
Cuatro estrellas	46 % de 92 000 = 42 320	0,46 (46 %)	86 480	0,94
Cinco estrellas	6 % de 92 000 = 5520	0,06 (6 %)	92 000	1
	$N = 92 000$	1		

b)



11. El número de tabletas que se han producido en una fábrica a lo largo de los últimos seis meses del año viene representado a través del siguiente pictograma:

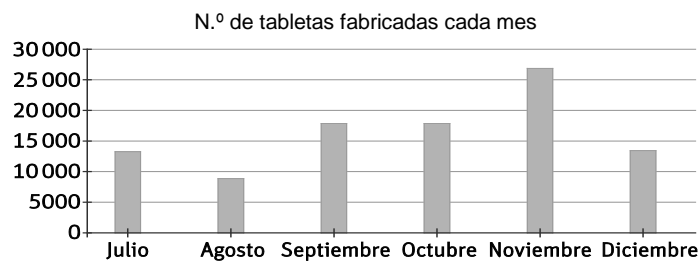


- a) Construye la tabla de frecuencias de los datos representados.
- b) Representa los datos en un diagrama de barras.
- c) Representa los datos en un diagrama de sectores.

a)

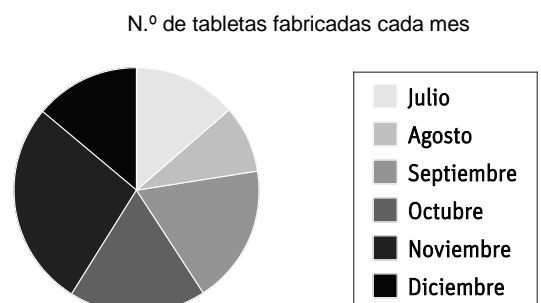
Meses	f_i	h_i	F_i	H_i
Julio	$4500 \cdot 3 = 13\ 500$	0,136	13 500	0,136
Agosto	$4500 \cdot 2 = 9000$	0,091	22 500	0,227
Septiembre	$4500 \cdot 4 = 18\ 000$	0,182	40 500	0,409
Octubre	$4500 \cdot 4 = 18\ 000$	0,182	58 500	0,591
Noviembre	$4500 \cdot 6 = 27\ 000$	0,273	85 500	0,864
Diciembre	$4500 \cdot 3 = 13\ 500$	0,136	99 000	1
	$N = 99\ 000$	1		

b)



c)

Meses	f_i	h_i	Amplitud del sector
Julio	13 500	0,136	$0,136 \cdot 360^\circ = 48,96^\circ$
Agosto	9000	0,091	$0,091 \cdot 360^\circ = 32,76^\circ$
Septiembre	18 000	0,182	$0,182 \cdot 360^\circ = 65,52^\circ$
Octubre	18 000	0,182	$0,182 \cdot 360^\circ = 65,52^\circ$
Noviembre	27 000	0,273	$0,273 \cdot 360^\circ = 98,28^\circ$
Diciembre	13 500	0,136	$0,136 \cdot 360^\circ = 48,96^\circ$
	$N = 99\ 000$	1	



12. Los días al mes que salen a correr varias personas son:

Edad	f_i
[0, 5)	8
[5, 10)	11
[10, 15)	15
[15, 20)	9
[20, 25)	5

- a) Completa la tabla de frecuencias.
 b) ¿Cuántos corredores corren entre 5 y 20 días?
 a)

Edad	Marca de clase, x_i	f_i	h_i	F_i	H_i
[0, 5)	2,5	8	0,167	8	0,167
[5, 10)	7,5	11	0,229	19	0,396
[10, 15)	12,5	15	0,313	34	0,709
[15, 20)	17,5	9	0,188	43	0,897
[20, 25)	22,5	5	0,104	48	1
		$N = 48$	1		

- b) Para determinar cuántas personas corren entre 5 y 20 días debemos sumar las frecuencias absolutas de los intervalos [5, 10), [10, 15) y [15, 20). Por tanto, $11 + 15 + 9 = 35$ personas corren entre 5 y 20 días.

13. Los kilogramos de masa al nacer de 24 niños son:

1,9 3,5 2,8 2,1 1,8 3,6 4,1 4,2
 2,7 3 2,7 2 3,2 3,3 4,2 2,2
 3,6 2,1 2,8 1,8 3,5 3,7 3,1 3,4

- a) Agrupa los datos de 0,5 en 0,5 y elabora la tabla.
 b) ¿Cuántos bebés han pesado más de 3 kg?
 a) Como el dato más pequeño es 1,9 kg y el más grande es 4,2 kg, vamos a considerar intervalos de amplitud 0,5 comenzando por 1,5 y finalizando en 4,5. Así, debemos determinar la cantidad de bebés cuyo peso al nacer se encuentra en cada uno de los intervalos.

Peso	Marca de clase, x_i	f_i	h_i	F_i	H_i
[1,5, 2)	1,75	3	0,125	3	0,125
[2, 2,5)	2,25	4	0,167	7	0,292
[2,5, 3)	2,75	4	0,167	11	0,459
[3, 3,5)	3,25	5	0,208	16	0,667
[3,5, 4)	3,75	5	0,208	21	0,875
[4, 4,5)	4,25	3	0,125	24	1
		$N = 24$	1		

- b) Debemos contar los bebés cuyo peso se encuentra en los intervalos [3, 3,5), [3,5, 4) y [4, 4,5). Así, $5 + 5 + 3 = 13$ bebés han pesado más de 3 kg al nacer.

14. El número de fotografías que ha hecho Gorka con su teléfono móvil durante los 20 días que ha estado de intercambio en París han sido:

31 60 57 49 29 14 63 38 50 19
48 23 28 19 21 22 36 18 46 27

- a) Agrupa los datos en intervalos de amplitud 10 y haz la tabla de frecuencias.
b) ¿Cuál es la marca de clase más frecuente?
c) ¿Cuántos días ha hecho menos de 40 fotos?
- a) Puesto que el menor número de fotos en un día es 14, y el mayor, 63, vamos a agrupar los datos en intervalos de longitud 10 comenzando en 10 y finalizando en 70. Por tanto, tenemos que determinar el número de días cuya cantidad de fotos se encuentra en cada intervalo.

N.º de fotos	Marca de clase, x_i	f_i	h_i	F_i	H_i
[10, 20)	15	4	0,20	4	0,20
[20, 30)	25	6	0,30	10	0,50
[30, 40)	35	3	0,15	13	0,65
[40, 50)	45	3	0,15	16	0,80
[50, 60)	55	2	0,10	18	0,90
[60, 70)	65	2	0,10	20	1
		$N = 20$	1		

- b) La marca de clase más frecuente es aquella que tiene el término f_i (frecuencia absoluta) más grande. Por tanto, la marca de clase con mayor frecuencia es 25.
c) El número de días que ha hecho menos de 40 fotos se obtiene sumando las frecuencias absolutas correspondientes a los intervalos [10, 20), [20, 30) y [30, 40). Por tanto, $4 + 6 + 3 = 13$ días hizo menos de 40 fotos.

15. Al realizar el inventario de un almacén se han pesado todos los paquetes que se encuentran en existencias:

Masa (kg)	f_i
[0, 10)	9
[10, 20)	12
[20, 30)	14
[30, 40)	3
[40, 50)	1

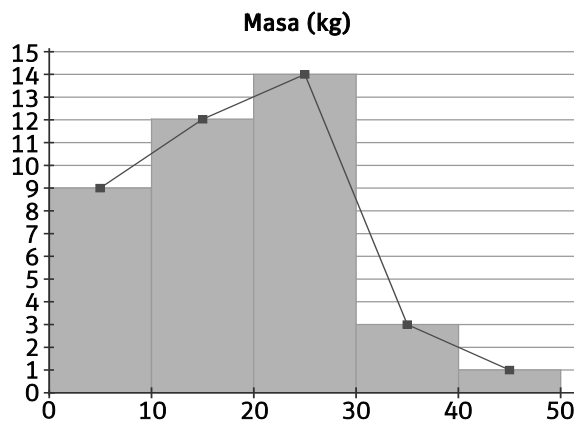
- a) Completa la tabla de frecuencias e indica las marcas de clase.
b) ¿Cuántos paquetes pesan menos de 30 kg?
c) Representa los datos en un histograma y construye su polígono de frecuencias.

a)

Masa (kg)	Marca de clase, x_i	f_i	h_i	F_i	H_i
[0, 10)	5	9	0,231	9	0,231
[10, 20)	15	12	0,308	21	0,539
[20, 30)	25	14	0,359	35	0,898
[30, 40)	35	3	0,077	38	0,975
[40, 50)	45	1	0,026	39	1
		$N = 39$	1		

- b) Tenemos que sumar la frecuencia absoluta asociada a los intervalos [0, 10), [10, 20) y [20, 30). Por tanto, $9 + 12 + 14 = 35$ paquetes pesan menos de 30 kg.

c)



16. Las estaturas en centímetros de los 30 alumnos de una clase son:

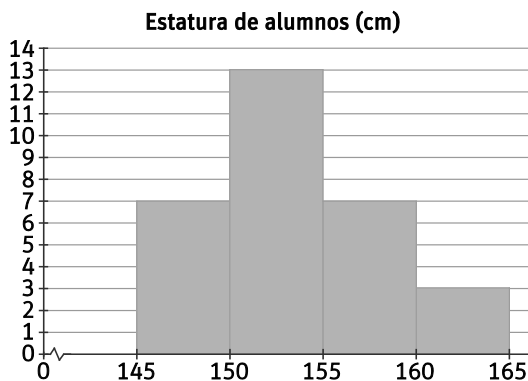
148 149 150 150 151 152
 151 153 154 148 155 156
 146 151 160 158 149 155
 153 161 148 152 158 156
 150 153 157 148 161 150

- a) Agrupa los datos en cuatro intervalos de amplitud 5.
 - b) Elabora la tabla de frecuencias.
 - c) Representa el histograma.
- a) Puesto que el menor de los datos es 146 cm, y el mayor, 161 cm, para agruparlos en cuatro intervalos de amplitud 5 vamos a considerar los intervalos [145, 150), [150, 155), [155, 160) y [160, 165).

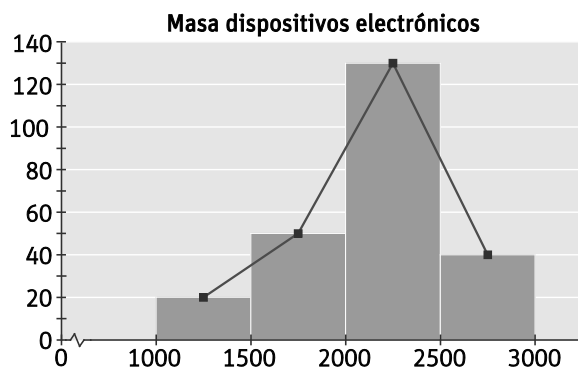
b)

Estatura (cm)	Marca de clase, x_i	f_i	h_i	F_i	H_i
[145, 150)	147,5	7	0,23	7	0,23
[150, 155)	152,5	13	0,43	20	0,66
[155, 160)	157,5	7	0,23	27	0,89
[160, 165)	162,5	3	0,1	30	1
		$N = 30$	1		

c)



17. En el histograma se representa la masa de diferentes marcas y modelos de dispositivos electrónicos:



- a) Elabora la tabla de frecuencias absolutas y relativas.
- b) ¿Cuántos dispositivos tienen una masa menor de 1 kg?

a)

Masa (gramos)	Marca de clase, x_i	f_i	h_i
[1000, 1500)	1250	20	0,083
[1500, 2000)	1750	50	0,208
[2000, 2500)	2250	130	0,542
[2500, 3000)	2750	40	0,167
		$N = 240$	1

- b) Como la masa de todos los dispositivos electrónicos representados en este histograma es superior a 1000 gramos, ninguno de ellos tiene una masa menor de 1 kg.

18. En un reconocimiento médico se ha determinado el nivel de calcio en sangre de varios pacientes:

9,7 9,6 8,2 7,1 7,8 9,8 8,9 10,1
 10 8,5 8,9 9,1 9,5 9,6 10 7,6
 8,3 8,3 8,4 10,1 10,2 9,3 9,4 9,4
 9,5 7,7 8,1 8,6 8,1 8 9,2 9
 9,9 8,8 8,8 7,9 7,6 8,4 8,5 9,5
 9,1 9,4 9,9 10,2 8,6 9,6 9,5 10

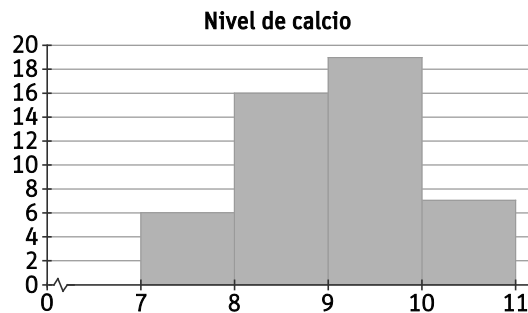
a) Construye la tabla de frecuencias de datos agrupados.

b) Representa los datos en un histograma.

a) El menor nivel de calcio en sangre es 7,1, y el mayor, 10,2. Por tanto, y puesto que no nos dicen nada al respecto, vamos a agrupar los datos en intervalos de amplitud 1 comenzando en 7 y finalizando en 11.

Calcio en sangre	Marca de clase, x_i	f_i	h_i	F_i	H_i
[7, 8)	8,5	6	0,125	6	0,125
[8, 9)	9,5	16	0,333	22	0,458
[9, 10)	10,5	19	0,396	41	0,854
[10, 11)	11,5	7	0,146	48	1
		$N = 48$	1		

b)



19. Las alturas de los miembros de la familia Vera son:

	Madre	Padre	Hugo	Eva	Fran	Ada
Altura	175	180	165	150	150	156

Calcula la moda, la media y la mediana.

La moda se corresponde con la estatura que tiene mayor frecuencia. En este caso, $M_0 = 150$, que coincide con las alturas de Eva y Fran.

Para calcular la media de las alturas, dividiremos la suma de todas las alturas entre el número de miembros de la familia.

$$\bar{x} = \frac{175 + 180 + 165 + 150 + 150 + 156}{6} = \frac{976}{6} = 162,66\dots$$

Para determinar la mediana, ordenaremos los datos de menor a mayor. Posteriormente, puesto que tenemos un número par de datos, la mediana es la media aritmética de los dos datos centrales.

150, 150, 156, 165, 175, 180

$$M = \frac{156 + 165}{2} = 160,5$$

20. Actividad resuelta.

21. En un campeonato de baile, la nota de cada ejercicio se calcula eliminando la puntuación más baja y la más alta de las dadas por los jueces y haciendo la media de las restantes. Las puntuaciones de los jueces a un bailarín han sido:

8,9 9,2 9,6 8,8 9,4 9,8 9,5

- a) ¿Qué nota sacará el bailarín? ¿Sería favorable si se hiciese la media con todas las calificaciones?
 b) ¿Y si la nota más baja y la más alta valieran el doble?

- a) Para calcular la nota del bailarín, eliminamos la puntuación más baja y la más alta. En este caso eliminamos las puntuaciones 8,8 y 9,8. Por tanto, la nota obtenida por el bailarín es:

$$\frac{8,9+9,2+9,6+9,4+9,5}{5} = \frac{46,6}{5} = 9,32$$

Calculemos ahora la nota considerando todas las notas obtenidas:

$$\frac{8,9+9,2+9,6+8,8+9,4+9,8+9,5}{7} = \frac{65,2}{7} = 9,314$$

En este caso no resulta favorable para el bailarín considerar todas las notas.

- b) Supongamos ahora que la nota más baja vale $8,8 \cdot 2 = 17,6$, y la más alta, $9,8 \cdot 2 = 19,8$. Entonces, la nota obtenida por el bailarín será:

$$\frac{8,9+9,2+9,6+17,6+9,4+19,8+9,5}{7} = \frac{83,8}{7} = 11,971$$

En el caso de que las notas más baja y más alta valieran el doble, la nota del bailarín aumentaría considerablemente.

22. En un concurso de debates se valoran cuatro aspectos con distintos pesos: con 4 el respeto, con 3 la comunicación, con 2 la claridad de las ideas y con 1 el lenguaje gestual. El equipo A obtuvo 8, 7, 8 y 6 puntos, respectivamente, en cada categoría, y el equipo B, 7, 10, 8 y 5 puntos. ¿Quién ganó el debate?

En este caso, como cada uno de los aspectos que se valoran en el concurso de debates tiene un peso asignado, debemos calcular la puntuación media ponderada de cada uno de los equipos.

$$\text{Media A: } \frac{4 \cdot 8 + 3 \cdot 7 + 2 \cdot 8 + 1 \cdot 6}{4 + 3 + 2 + 1} = \frac{75}{10} = 7,5$$

$$\text{Media B: } \frac{4 \cdot 7 + 3 \cdot 10 + 2 \cdot 8 + 1 \cdot 5}{4 + 3 + 2 + 1} = \frac{79}{10} = 7,9$$

Por tanto, el equipo B ganó el debate.

23. Actividad resuelta.

24. Las notas, sobre 100 puntos, obtenidas por los aspirantes presentados a una oposición han sido:

Nota	[0, 20)	[20, 40)	[40, 60)	[60, 80)	[80, 100]
f_i	5	28	52	65	30

- a) Halla el intervalo modal, la media y el intervalo mediano.
 b) ¿Cuántos aspirantes han obtenido una nota por encima de la media? ¿Y por encima del intervalo mediano?

Nota	Marca de clase, x_i	f_i	h_i	F_i	H_i
[0, 20)	10	5	0,028	5	0,028
[20, 40)	30	28	0,156	33	0,184
[40, 60)	50	52	0,289	85	0,473
[60, 80)	70	65	0,361	150	0,834
[80, 100]	90	30	0,166	180	1
		$N = 180$	1		

- a) El intervalo modal es el intervalo con mayor frecuencia. En este caso, el intervalo [60, 80).

La media se calcula a partir de las marcas de clase de cada uno de los intervalos:

$$\frac{10 \cdot 5 + 30 \cdot 28 + 50 \cdot 52 + 70 \cdot 65 + 90 \cdot 30}{180} = \frac{10740}{180} = 59,666\dots$$

Por último, puesto que tenemos 180 datos, el intervalo mediano será aquel en el que se encuentran los datos que ocupan las posiciones centrales, es decir, los datos que ocupan los lugares 90 y 91. En el caso que nos ocupa, ambos datos se encuentran en el intervalo [60, 80), por lo que se trata del intervalo mediano.

- b) Puesto que la media es 59,666..., podemos considerar que nos están preguntando por los aspirantes cuyas notas se encuentran en los intervalos [60, 80) y [80, 100]. Por tanto, la cantidad de aspirantes que han obtenido una nota superior a la media es 65 + 30 = 95.

Por otro lado, puesto que el intervalo mediano es [60, 80), los aspirantes que han sacado una nota por encima del intervalo mediano son los que se encuentran en el intervalo [80, 100], es decir, 30 aspirantes.

25. Actividad interactiva.

26. En la siguiente tabla se recogen las estaturas, en centímetros, de los jugadores de dos equipos de baloncesto de la ciudad:

Equipo	J1	J2	J3	J4	J5
A	190	192	195	198	200
B	180	185	195	210	200

- a) Calcula la media y la mediana.
 b) Calcula el recorrido y la desviación media.
 c) Calcula la varianza y la desviación típica.
 d) ¿Qué conclusiones puedes sacar en vista de los resultados obtenidos para cada equipo?

a) Media A: $\frac{190 + 192 + 195 + 198 + 200}{5} = \frac{975}{5} = 195$

Media B: $\frac{180 + 185 + 195 + 210 + 200}{5} = \frac{970}{5} = 194$

Mediana A: 195

Mediana B: 195

- b)

Equipo A:

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
190	-5	5	25
192	-3	3	9
195	0	0	0
198	3	3	9
200	5	5	25
		16	68

Recorrido A: 200 - 190 = 10

Desviación media A: $\frac{16}{5} = 3,2$

c) Varianza A: $s^2 = \frac{68}{5} = 13,6$

Desviación típica A: $s = \sqrt{13,6} = 3,69$

Equipo B:

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
180	-14	14	196
185	-9	9	81
195	1	1	1
210	16	16	256
200	6	6	36
		46	570

Recorrido B: 210 - 180 = 30

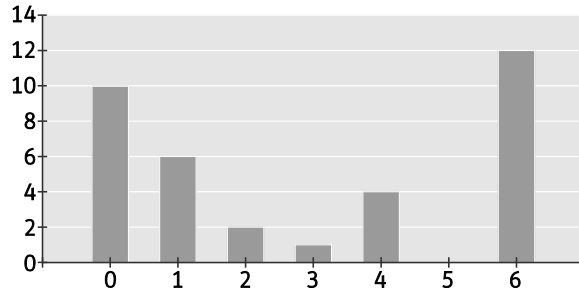
Desviación media B: $\frac{46}{5} = 9,2$

Varianza B: $s^2 = \frac{570}{5} = 114$

Desviación típica B: $s = \sqrt{114} = 10,68$

- d) La media y la mediana del equipo A coinciden, mientras que las del equipo B son distintas. Además, el recorrido, la desviación media, la varianza y la desviación típica son más grandes para el equipo B que para el equipo A. Por tanto, podemos concluir que las estaturas de los jugadores del equipo A se encuentran más próximas a la media, mientras que las del equipo B están más dispersas.

27. Observa los datos representados:



- a) Construye la tabla de frecuencias y calcula la media.
 b) Calcula los parámetros de dispersión.

a)

x_i	f_i	h_i	F_i	H_i
0	10	0,286	10	0,286
1	6	0,171	16	0,457
2	2	0,057	18	0,514
3	1	0,029	19	0,543
4	4	0,114	23	0,657
5	0	0	23	0,657
6	12	0,343	35	1
$N = 35$		1		

Media:

$$\frac{0 \cdot 10 + 1 \cdot 6 + 2 \cdot 2 + 3 \cdot 1 + 4 \cdot 4 + 5 \cdot 0 + 6 \cdot 12}{35} = \frac{101}{35} = 2,886$$

b)

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
0	-2,886	2,886	8,329
1	-1,886	1,886	3,557
2	-0,886	0,886	0,785
3	0,114	0,114	0,013
4	1,114	1,114	1,241
5	2,114	2,114	4,469
6	3,114	3,114	9,697
		12,114	28,091

Desviación media: $\frac{12,114}{7} = 1,731$

Varianza: $s^2 = \frac{28,091}{7} = 4,013$

Desviación típica: $s = \sqrt{4,013} = 2,003$

28. Para hacer un estudio sobre la calidad de la prensa escrita de una ciudad de 1 000 000 de habitantes se han seleccionado 2000 personas, de las cuales 150 tienen estudios superiores, y el resto, estudios medios. Además, 995 son mayores de 45 años, y el resto, menores.

- a) Indica cuáles son la población y la muestra del estudio.
 b) ¿La muestra seleccionada con respecto al nivel de estudios es representativa? ¿Y con respecto a la edad?
- a) La población está formada por el millón de habitantes de la ciudad, mientras que la muestra son las 2000 personas que han sido seleccionadas para realizar el estudio.
 b) La muestra no es representativa respecto al nivel de estudios, puesto que se seleccionan muchas más personas con estudios medios que con estudios superiores.

Sin embargo, respecto a la edad, se trata de una muestra representativa, dado que prácticamente la mitad son mayores de 45 años, y la otra mitad, menores.

29. En un formulario para matricularse en un curso de informática se pide completar varios campos donde figuran las siguientes variables.

- a) Edad.
- b) País de nacimiento.
- c) Programas informáticos que maneja.
- d) Número de dispositivos informáticos que utiliza.
- e) Tipo de dispositivos informáticos que utiliza.

¿De qué tipo es cada una de las variables?

- a) Variable cuantitativa discreta
- b) Variable cualitativa
- c) Variable cualitativa
- d) Variable cuantitativa discreta
- e) Variable cualitativa

30. Indica si cada una de las variables que aparecen a continuación son cualitativas o cuantitativas, y en ese caso, si son continuas o discretas.

- a) Número de asistentes a un festival de música cada año.
- b) Precio de la entrada al festival cada año.
- c) Tipo de música de los grupos que tocan en el festival.
- d) Número de grupos que actúan en el festival cada año.

- a) Variable cuantitativa discreta
- b) Variable cuantitativa continua
- c) Variable cualitativa
- d) Variable cuantitativa discreta

31. El número de personas que habitan en cada una de las viviendas de un edificio es:

4	4	2	1	2	2	4	4	2	3
4	3	2	5	2	4	5	3	6	4
2	1	3	2	2	1	4	4	5	2
1	1	4	3	3	2	2	1	2	1

- a) ¿De qué tipo es la variable estadística estudiada?
- b) Elabora una tabla de frecuencias a partir de los datos recogidos.
- c) ¿Cuál es el valor de la variable más frecuente?

- a) Variable cuantitativa discreta
- b)

Personas por vivienda	f_i	h_i	F_i	H_i
1	7	0,175	7	0,175
2	13	0,325	20	0,500
3	6	0,150	26	0,650
4	10	0,250	36	0,900
5	3	0,075	39	0,975
6	1	0,025	40	1
	$N = 40$	1		

- c) El valor de la variable más frecuente es el 2.

32. Para analizar los resultados de la evaluación de un curso de 2.º de ESO, el tutor ha realizado un recuento de las asignaturas suspendidas de cada alumno. Los resultados son:

4 4 2 1 2 2 4 4 2 3
 4 3 2 5 2 4 5 3 6 4
 2 1 3 2 2 1 4 4 5 2

- a) Construye la tabla de frecuencias.
 b) El alumno puede pasar de curso si tiene menos de tres asignaturas suspendidas. ¿Qué porcentaje de alumnos estarían en esa situación?

a)

Asignaturas suspendidas	f_i	h_i	F_i	H_i
1	3	0,100	3	0,100
2	10	0,333	13	0,433
3	4	0,133	17	0,566
4	9	0,300	26	0,866
5	3	0,100	29	0,966
6	1	0,034	30	1
	$N = 30$	1		

- b) Los alumnos que tienen menos de tres asignaturas suspendidas son aquellos que han suspendido una o dos asignaturas. Por tanto, $3 + 10 = 13$ alumnos pueden pasar de curso (este valor puede obtenerse observando el segundo dato de la columna de frecuencias absolutas acumuladas, F_i).

El porcentaje de alumnos en esta situación es del $\frac{13}{30} \cdot 100 = 43,3\%$ (este valor puede obtenerse observando el segundo dato de la columna de las frecuencias relativas acumuladas, H_i).

33. Se ha medido el ritmo cardiaco de un grupo de personas en estado de reposo obteniéndose los siguientes resultados:

Pulsaciones	f_i
[50, 60)	5
[60, 70)	12
[70, 80)	14
[80, 90)	8
[90, 100)	9
[100, 110)	2

- a) Completa la tabla de frecuencias e indica las marcas de clase.
 b) Si un ritmo cardiaco normal se encuentra entre las 60 y 100 pulsaciones por minuto, ¿cuántas personas se encuentran fuera de la normalidad?

a)

Pulsaciones	Marcas de clase, x_i	f_i	h_i	F_i	H_i
[50, 60)	55	5	0,10	5	0,10
[60, 70)	65	12	0,24	17	0,34
[70, 80)	75	14	0,28	31	0,62
[80, 90)	85	8	0,16	39	0,78
[90, 100)	95	9	0,18	48	0,96
[100, 110)	105	2	0,04	50	1
		$N = 50$			

- b) Las personas que se encuentran fuera de la normalidad son aquellas cuyas pulsaciones se encuentran en los intervalos [50, 60) y [100, 110). Por tanto, $5 + 2 = 7$ personas tienen un ritmo cardiaco fuera de lo normal.

34. Un estudio estadístico sobre la duración, en horas, de la batería de varios móviles ha arrojado los siguientes datos:

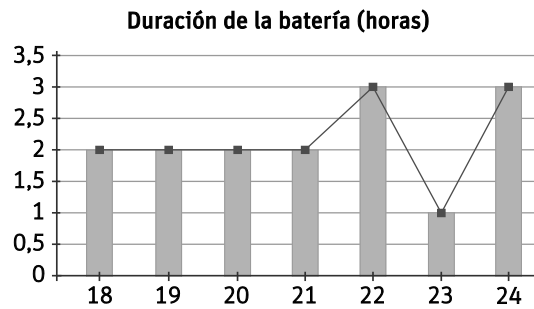
18 20 22 19 21 21 20 24
24 18 23 22 19 24 22

- Construye una tabla de frecuencias.
- Representa los datos en un diagrama de barras y un polígono de frecuencias.
- Representa los datos en un diagrama de sectores.

a)

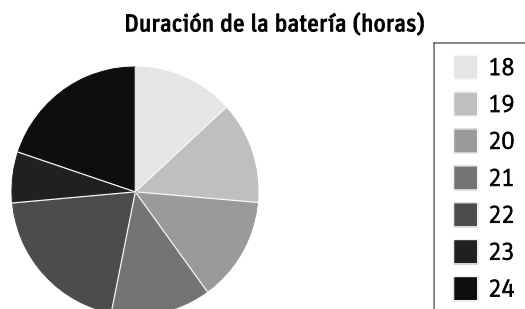
Duración (horas)	f_i	h_i	F_i	H_i
18	2	0,133	2	0,133
19	2	0,133	4	0,266
20	2	0,133	6	0,399
21	2	0,133	8	0,532
22	3	0,200	11	0,732
23	1	0,067	12	0,8
24	3	0,200	15	1
$N = 15$		1		

b)



c)

Duración (horas)	f_i	h_i	Amplitud del sector
18	2	0,133	48°
19	2	0,133	48°
20	2	0,133	48°
21	2	0,133	48°
22	3	0,200	72°
23	1	0,068	24°
24	3	0,200	72°
$N = 15$		1	



35. El número de ausencias laborales en una empresa de 24 empleados en un año ha sido:

2 1 3 4 2 3 1 3 1 4 0 0
3 3 4 3 0 3 2 1 0 0 4 2

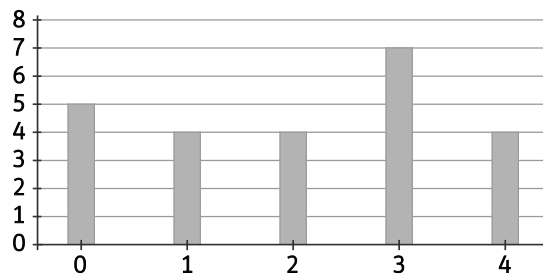
a) Representa los datos en un diagrama de barras y en un diagrama de sectores.

b) ¿Qué porcentaje de empleados se ha ausentado menos de dos días?

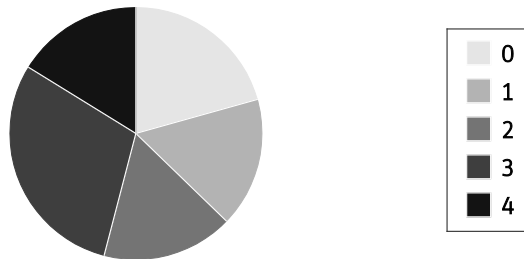
a)

Ausencias laborales	f_i	h_i	F_i	H_i	Amplitud del sector
0	5	0,208	5	0,208	75°
1	4	0,167	9	0,375	60°
2	4	0,167	13	0,542	60°
3	7	0,291	20	0,833	105°
4	4	0,167	24	1	60°
$N = 24$		1			

Ausencias laborales



Ausencias laborales



b) Los empleados que se han ausentado menos de dos días son aquellos que lo han hecho 0 o 1 días. Así, para determinar el porcentaje de empleados que se ausentaron menos de dos días debemos observar el segundo dato de la columna de las frecuencias relativas acumuladas, H_i , es decir, un 37,5 %.

36. El tiempo, en minutos, que han tardado ciertos atletas en completar una carrera ha sido:

98	100	105	102	95	89	110
99	95	94	94	89	86	85
81	121	80	90	100	103	103
106	108	113	112	111	94	92
87	86	86	114	108	108	101

a) Agrupa los datos en intervalos de amplitud 5.

b) Elabora la tabla de frecuencias.

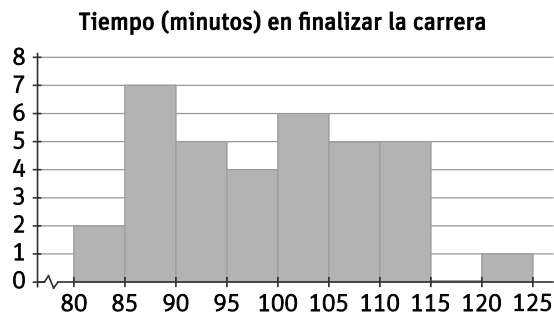
c) Representa el histograma.

a) Puesto que el menor de los datos es 80 minutos, y el mayor, 121 minutos, vamos a agrupar los datos en intervalos de amplitud 5 que comiencen por 80 y finalicen en 125.

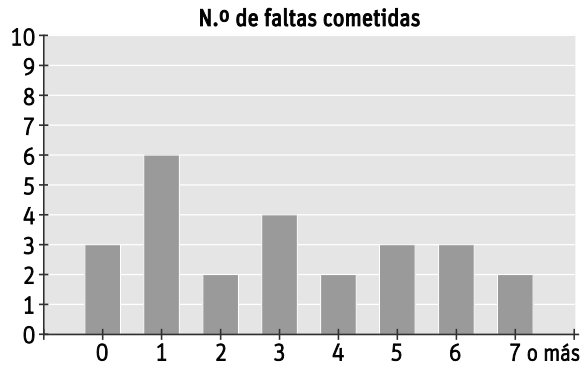
b)

Tiempo (minutos)	Marca de clase, x_i	f_i	h_i	F_i	H_i
[80, 85)	82,5	2	0,057	2	0,057
[85, 90)	87,5	7	0,200	9	0,257
[90, 95)	92,5	5	0,143	14	0,400
[95, 100)	97,5	4	0,114	18	0,514
[100, 105)	102,5	6	0,171	24	0,685
[105, 110)	107,5	5	0,143	29	0,828
[110, 115)	112,5	5	0,143	34	0,971
[115, 120)	117,5	0	0	34	0,971
[120, 125)	122,5	1	0,029	35	1
		$N = 35$	1		

c)



37. Entre un grupo de estudiantes se ha hecho una prueba de ortografía consistente en escribir un texto dictado. El número de faltas cometidas por cada uno se recoge en el siguiente gráfico.



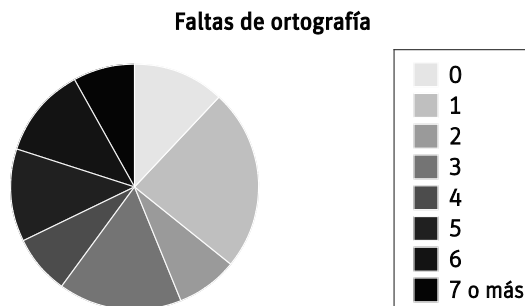
- Construye una tabla de frecuencias.
- Si la prueba se aprueba con menos de tres faltas de ortografía, ¿cuántos alumnos la han superado?
- Construye un diagrama de sectores a partir de los datos.

a)

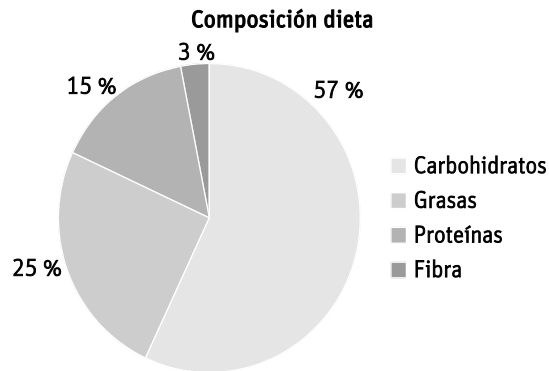
Faltas de ortografía	f_i	h_i	F_i	H_i	Amplitud del sector
0	3	0,12	3	0,12	43,2°
1	6	0,24	9	0,36	86,4°
2	2	0,08	11	0,44	28,8°
3	4	0,16	15	0,60	57,6°
4	2	0,08	17	0,68	28,8°
5	3	0,12	20	0,80	43,2°
6	3	0,12	23	0,92	43,2°
7 o más	2	0,08	25	1	28,8°
	$N = 25$	1			

b) Si la prueba se aprueba con menos de tres faltas de ortografía, debemos contar el número de estudiantes que tuvieron 0, 1 o 2 faltas de ortografía. Así, $3 + 6 + 2 = 11$ alumnos superaron la prueba (también se puede mirar el tercer dato de la columna de las frecuencias absolutas acumuladas, F_i).

c)



38. Según las recomendaciones sobre aportaciones a la dieta (*RDA, Recommended Dietary Allowances*) publicadas por la Academia Nacional de Ciencias de EE. UU., una dieta equilibrada debería tener la siguiente composición:



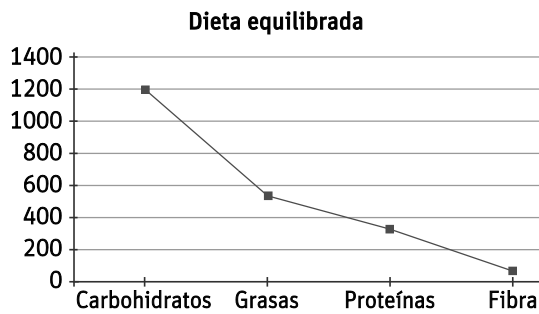
Si el total de calorías diarias debe ser de 2100:

- a) Construye una tabla de frecuencias asociada al gráfico.
- b) Representa los datos en un polígono de frecuencias.

a)

Dieta	f_i	h_i	F_i	H_i
Carbohidratos	57 % de 2100 = 1197	0,57 (57 %)	1197	0,57
Grasas	25 % de 2100 = 525	0,25 (25 %)	1722	0,82
Proteínas	15 % de 2100 = 315	0,15 (15 %)	2037	0,97
Fibra	3 % de 2100 = 63	0,03 (3 %)	2100	1
	$N = 2100$	1		

b)



39. En una encuesta sobre hábitos alimenticios se han recogido los siguientes datos sobre el número de piezas de fruta que se consumen diariamente.

Piezas	0	1	2	3	4	5
f_i	14	22	18	10	6	2

- a) Halla la moda, la media y la mediana.
 b) Representa los datos en un diagrama de sectores.

a)

Piezas de fruta	f_i	h_i	F_i	H_i	Amplitud del sector
0	14	0,194	14	0,194	70°
1	22	0,306	36	0,500	110°
2	18	0,250	54	0,750	90°
3	10	0,139	64	0,889	50°
4	6	0,083	70	0,972	30°
5	2	0,028	72	1	10°
	$N = 72$	1			

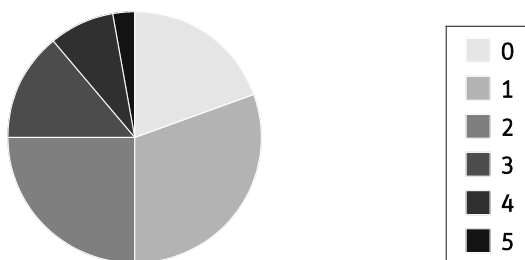
La moda es el dato de mayor frecuencia absoluta. Por tanto, $M_0 = 1$.

$$\text{Media: } \frac{0 \cdot 14 + 1 \cdot 22 + 2 \cdot 18 + 3 \cdot 10 + 4 \cdot 6 + 5 \cdot 2}{72} = \frac{122}{72} = 1,69 \text{ piezas de fruta diarias}$$

Puesto que tenemos una cantidad par de datos, 72, la mediana se obtiene como la media de los datos centrales, es decir, de los datos que ocupan las posiciones 36 y 37. Observando la columna de las frecuencias absolutas acumuladas, tenemos que el dato 36 vale 1, mientras que el dato 37 vale 2, por lo que la mediana será la media entre 1 y 2, es decir, $M = 1,5$.

b)

Piezas de fruta diarias



40. El número de viajes que ha realizado un coche grúa de una empresa pública a lo largo de la semana ha sido:

L	M	X	J	V	S	D
6	7	6	4	3	10	12

Y los realizados por una empresa privada:

L	M	X	J	V	S	D
3	0	5	6	7	9	11

Para cada empresa:

- a) Halla la desviación media y el rango.
 b) Calcula la varianza y la desviación típica.
 c) ¿Qué empresa tiene valores menos dispersos?

a) Media (empresa pública): $\frac{6+7+6+4+3+10+12}{7} = \frac{48}{7} = 6,86$ viajes diarios

Media (empresa privada): $\frac{3+0+5+6+7+9+11}{7} = \frac{41}{7} = 5,86$ viajes diarios

Empresa pública:

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
6	-0,86	0,86	0,7396
7	0,14	0,14	0,0196
6	-0,86	0,86	0,7396
4	-2,86	2,86	8,1796
3	-3,86	3,86	14,8996
10	3,14	3,14	9,8596
12	5,14	5,14	26,4196
	16,86		60,8572

Desviación media: $\frac{16,86}{7} = 2,41$

Rango: $12 - 3 = 9$

b) Empresa pública:

Varianza: $s^2 = \frac{60,8572}{7} = 8,69$

Desviación típica: $s = \sqrt{8,69} = 2,95$

Empresa privada:

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
3	-2,86	2,86	8,1796
0	-5,86	5,86	34,3396
5	-0,86	0,86	0,7396
6	0,14	0,14	0,0196
7	1,14	1,14	1,2996
9	3,14	3,14	9,8596
11	5,14	5,14	26,4196
	19,14		80,8572

Desviación media: $\frac{19,14}{7} = 2,73$

Rango: $11 - 0 = 11$

Empresa privada:

Varianza: $s^2 = \frac{80,8572}{7} = 11,55$

Desviación típica: $s = \sqrt{11,55} = 3,40$

- c) La empresa pública tiene los valores menos dispersos, puesto que todos los parámetros de dispersión que hemos calculado son más pequeños (desviación media, rango, varianza y desviación típica).

41. Un pediatra hace un estudio sobre la edad, en meses, a la que los bebés comienzan a caminar y obtiene los siguientes datos de 12 bebés: 12, 13, 12, 11, 15, 12, 13, 10, 14, 11, 12 y 13 meses.

a) Halla la moda y la media de los datos.

b) Calcula la mediana.

c) Halla la desviación media.

d) Calcula la varianza y la desviación típica.

a) Moda: $M_0 = 12$ meses

Media: $\bar{x} = \frac{1 \cdot 10 + 2 \cdot 11 + 4 \cdot 12 + 3 \cdot 13 + 1 \cdot 14 + 1 \cdot 15}{12} = \frac{148}{12} = 12,33$ meses

x_i	f_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
10	1	-2,33	2,33	5,4289
11	2	-1,33	1,33	1,7689
12	4	-0,33	0,33	0,1089
13	3	0,67	0,67	0,4489
14	1	1,67	1,67	2,7889
15	1	2,67	2,67	7,1289
			9	17,6734

- b) La mediana es la media de los datos que ocupan las posiciones centrales, es decir, la media de los datos que ocupan los puestos 6 y 7. Puesto que ambos datos valen 12, $M = 12$.

c) Desviación media: $\frac{9}{12} = 0,75$

d) Varianza: $s^2 = \frac{17,6734}{12} = 1,47$

Desviación típica: $s = \sqrt{1,47} = 1,21$

42. En el colegio de Lola, su profesora de matemáticas puntúa los exámenes de 0 a 100. Lola tiene una media de 60 puntos de sus primeros cuatro exámenes y en el quinto examen sacó 80 puntos.

a) ¿Cuál es la media de las notas de Lola después de los cinco exámenes?

b) Si el último examen es global y vale un 60 %, ¿qué nota final ha obtenido?

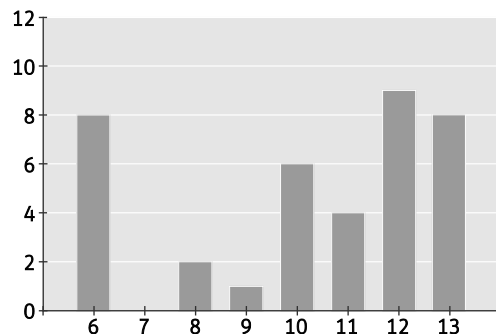
a) Después de los cinco exámenes, la nota media de los cuatro primeros cuenta 4/5 de la nota, mientras que la nota del último examen cuenta 1/5. Por tanto, debemos calcular una media ponderada de las notas.

$$\text{Media: } \bar{x} = \frac{0,8 \cdot 60 + 0,2 \cdot 80}{0,8 + 0,2} = 64$$

b) La nota final también se calcula haciendo una media ponderada.

$$\text{Nota final: } \frac{0,4 \cdot 60 + 0,6 \cdot 80}{0,4 + 0,6} = 72$$

43. Determina la media, la mediana y la moda de los datos representados en el siguiente diagrama de barras.



a) Construye la tabla de frecuencias asociada.

b) Halla el recorrido y la desviación media.

c) Calcula la varianza y la desviación típica.

d) Construye el diagrama de sectores asociado a los datos.

a)

x _i	f _i	h _i	F _i	H _i	Amplitud del sector
6	8	0,21	8	0,21	75,6°
7	0	0	8	0,21	0°
8	2	0,0526	10	0,2626	18,9°
9	1	0,0263	11	0,2889	9,4°
10	6	0,1578	17	0,4467	56,8°
11	4	0,105	21	0,5517	37,9°
12	9	0,23684	30	0,79	85,2°
13	8	0,21	38	1	75,6°
N = 38		1			

$$\text{Media: } \bar{x} = \frac{6 \cdot 8 + 7 \cdot 0 + 8 \cdot 2 + 9 \cdot 1 + 10 \cdot 6 + 11 \cdot 4 + 12 \cdot 9 + 13 \cdot 8}{38} = \frac{389}{38} = 10,24$$

Mediana: M = 11

Moda: M₀ = 12

b)

x _i	x _i - \bar{x}	x _i - \bar{x}	(x _i - \bar{x}) ²
6	-4,24	4,24	17,9776
7	-3,24	3,24	10,4976
8	-2,24	2,24	5,0176
9	-1,24	1,24	1,5376
10	-0,24	0,24	0,0576
11	0,76	0,76	0,5776
12	1,76	1,76	3,0976
13	2,76	2,76	7,6176
		16,48	46,3808

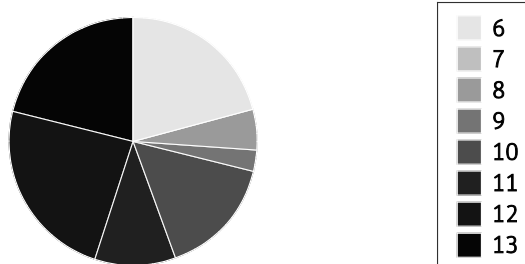
Recorrido: 13 - 6 = 7

$$\text{Desviación media: } \frac{16,48}{38} = 0,43$$

c) Varianza: $s^2 = \frac{46,3808}{38} = 1,22$

Desviación típica: $s = \sqrt{1,22} = 1,1$

d)



44. El ahorro a lo largo de un año conseguido por varias familias de una urbanización viene expresado en la siguiente tabla.

Ahorro (€)	N.º de familias
[0, 600)	11
[600, 1200)	15
[1200, 1800)	25
[1800, 2400)	39
[2400, 3000)	10

- a) Halla el intervalo modal y la media.
 b) Representa el histograma y el polígono de frecuencias.

a)

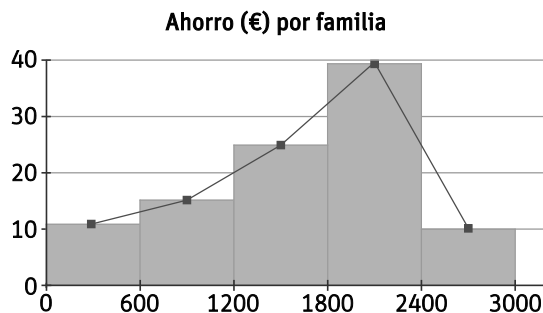
Ahorro (€)	Marca de clase, x_i	N.º de familias
[0, 600)	300	11
[600, 1200)	900	15
[1200, 1800)	1500	25
[1800, 2400)	2100	39
[2400, 3000)	2700	10
		$N = 100$

El intervalo modal es el intervalo de mayor frecuencia, es decir, [1800, 2400).

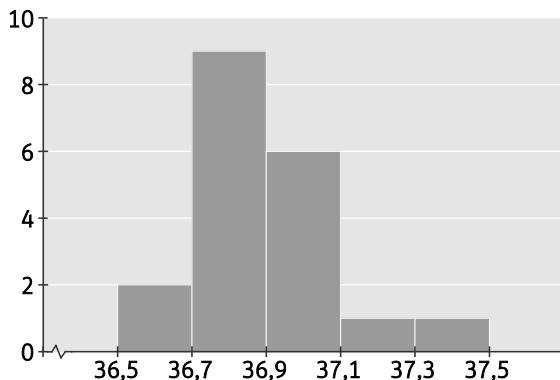
Para calcular la media, vamos a tener en cuenta la marca de clase de cada intervalo.

$$\text{Media: } \bar{x} = \frac{300 \cdot 11 + 900 \cdot 15 + 1500 \cdot 25 + 2100 \cdot 39 + 2700 \cdot 10}{100} = \frac{163200}{100} = 1632 \text{ €}$$

b)



45. En el siguiente histograma se representa la temperatura corporal habitual de un grupo de personas.



- a) Construye la tabla de frecuencias asociada.
- b) Calcula el intervalo modal.
- c) Halla la media y el intervalo mediano.
- d) ¿Cuál es el recorrido o rango de los datos?

a)

Temperatura corporal	Marca de clase, x_i	f_i	h_i	F_i	H_i
[36,5, 36,7)	36,6	2	0,11	2	0,11
[36,7, 36,9)	36,8	9	0,47	11	0,58
[36,9, 37,1)	37	6	0,32	17	0,90
[37,1, 37,3)	37,2	1	0,05	18	0,95
[37,3, 37,5)	37,4	1	0,05	19	1
		$N = 19$	1		

b) Intervalo modal: [36,7, 36,9)

c) Media: $\bar{x} = \frac{36,6 \cdot 2 + 36,8 \cdot 9 + 37 \cdot 6 + 37,2 \cdot 1 + 37,4 \cdot 1}{19} = \frac{701}{19} = 36,89$

Intervalo mediano: [36,7, 36,9)

d) Recorrido: $37,4 - 36,6 = 0,8$

46. ¿Cuál es la media de todos los números de cinco cifras que pueden formarse con los dígitos 1, 3, 5, 7 y 8 utilizando en cada número los cinco dígitos?

- A. 48 000 B. 53 332,8 C. 49 999,5 D. 55 555

En primer lugar, la cantidad de números de cinco cifras que podemos formar con los dígitos 1, 3, 5, 7 y 8 es $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$. Nos debemos centrar ahora en determinar cuánto vale la suma de todos ellos.

Los números de cinco cifras que podemos formar son de la forma $abcde$, donde a, b, c, d y e pueden tomar los valores 1, 3, 5, 7 y 8. El valor de cada número de la forma $abcde$ es:

$$abcde = a \cdot 10000 + b \cdot 1000 + c \cdot 100 + d \cdot 10 + e \cdot 1$$

Además, cada coeficiente (a, b, c, d y e) toma el valor de cada uno de los dígitos (1, 3, 5, 7 y 8) en 24 ocasiones. Por tanto, la suma de todos los números de cinco cifras que podemos formar con los dígitos 1, 3, 5, 7 y 8 es:

$$24 \cdot (1+3+5+7+8) \cdot (10000 + 1000 + 100 + 10 + 1) = 24 \cdot 24 \cdot 11111 = 24^2 \cdot 11111$$

Finalmente, la media de todos estos números es:

$$\bar{x} = \frac{24^2 \cdot 11111}{120} = \frac{24^2 \cdot 11111}{24 \cdot 5} = \frac{24 \cdot 11111}{5} = 53332,8$$

La respuesta correcta es la B.

47. Si la media aritmética de 10 enteros positivos es 5, ¿cuál es el mayor valor posible que puede tomar alguno de ellos?

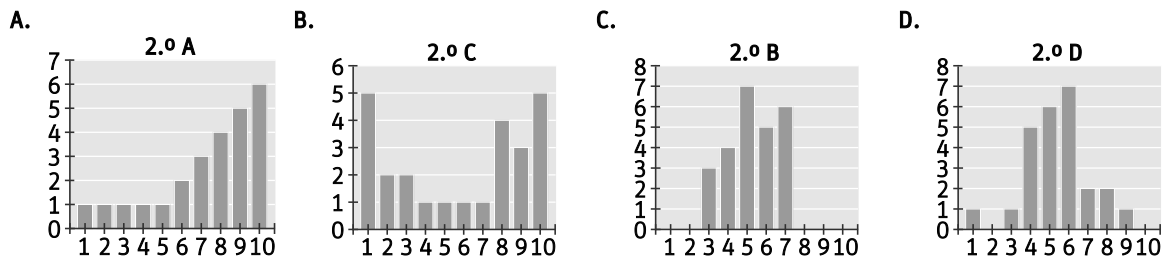
A. 50 B. 5 C. 41 D. 25

Para obtener el mayor número posible que puede tomar alguno de los números que verifican las condiciones dadas, los otros nueve números deben tomar el menor número posible. Por tanto, como tienen que ser números enteros y positivos, vamos a suponer que nueve de ellos valen 1 y calculemos el valor del décimo número para que la media de todos ellos sea 5. Llamaremos a este número n . Así,

$$\frac{9 \cdot 1 + n}{10} = 5 \Rightarrow 9 + n = 5 \cdot 10 \Rightarrow n = 50 - 9 = 41$$

La solución es la C.

48. Las notas de cuatro grupos diferentes de 2.º de ESO vienen representadas en cuatro gráficos distintos. ¿En qué grupo la media y la mediana se diferencian menos?



En todos los grupos de 2.º de ESO hay 25 alumnos, por lo que para calcular la mediana nos debemos fijar en la nota del alumno 13.

2.º A:

Mediana: 8

$$\text{Media: } \bar{x} = \frac{1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 2 + 7 \cdot 3 + 8 \cdot 4 + 9 \cdot 5 + 10 \cdot 6}{25} = \frac{185}{25} = 7,4$$

2.º B:

Mediana: 5

$$\text{Media: } \bar{x} = \frac{1 \cdot 0 + 2 \cdot 0 + 3 \cdot 3 + 4 \cdot 4 + 5 \cdot 7 + 6 \cdot 5 + 7 \cdot 6 + 8 \cdot 0 + 9 \cdot 0 + 10 \cdot 0}{25} = \frac{132}{25} = 5,28$$

2.º C:

Mediana: 7

$$\text{Media: } \bar{x} = \frac{1 \cdot 5 + 2 \cdot 2 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 + 7 \cdot 1 + 8 \cdot 4 + 9 \cdot 3 + 10 \cdot 5}{25} = \frac{146}{25} = 5,84$$

2.º D:

Mediana: 5

$$\text{Media: } \bar{x} = \frac{1 \cdot 1 + 2 \cdot 0 + 3 \cdot 1 + 4 \cdot 5 + 5 \cdot 6 + 6 \cdot 7 + 7 \cdot 2 + 8 \cdot 2 + 9 \cdot 1 + 10 \cdot 0}{25} = \frac{135}{25} = 5,4$$

El grupo en el que menos se diferencian la media y la mediana es 2.º B.

49. En un conjunto finito de números enteros positivos no todos iguales, ¿qué afirmación de las siguientes es siempre verdadera?

- A. La media es un número entero.
- B. El rango es positivo.
- C. La media es un número natural.
- D. La mediana es un número entero.

A. FALSA

La media de 1 y 2 es 1,5, que no es un número entero.

B. VERDADERA

Supongamos que b es el mayor entero positivo del conjunto y que a es el menor entero positivo. Entonces se tiene que $b - a > 0$, es decir, el rango es siempre positivo.

C. FALSA

La media de 1 y 2 es 1,5, que no es un número natural.

D. FALSA

Consideremos el conjunto $\{1, 2, 3, 4\}$, formado por cuatro números enteros positivos. Como hay un número par de datos, la mediana es la media de los datos que ocupan las posiciones centrales, es decir, la media de 2 y 3. Por tanto, la mediana es 2,5, que no es un número entero.

50. Ali tiene que calcular la media de estas 10 alturas:

170, 165, 178, 175, 160, 165, 166, 172, 162 y 165 cm.

Para ello, utiliza la siguiente estrategia:

1. Busca la moda de las estaturas: **165.**

2. Resta 165 a todas las alturas y obtiene:

5, 0, 13, 10, -5, 0, 1, 7, -3, 0

3. Halla la media de estos números:

Media: 2,8

4. Suma este valor a 165:

165 + 2,8 = 167,8 cm

Si la media obtenida es correcta, ¿crees que funcionará siempre? ¿Por qué?

Veamos, en primer lugar, si esta media coincide con la media aritmética habitual.

$$\bar{x} = \frac{170 + 165 + 178 + 175 + 160 + 165 + 166 + 172 + 162 + 165}{10} = \frac{1678}{10} = 167,8 \text{ cm}$$

En este caso particular, ambas medias coinciden. Comprobemos que es cierto en general.

Sean $\{x_1, x_2, \dots, x_n\}$ un conjunto de n datos cuya moda es M_0 . La media aritmética de estos datos es:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Veamos ahora que la media del conjunto de datos $\{x_1 - M_0, x_2 - M_0, \dots, x_n - M_0\}$ coincide con la diferencia entre la media de los datos $\{x_1, x_2, \dots, x_n\}$ y la moda.

$$\frac{(x_1 - M_0) + (x_2 - M_0) + \dots + (x_n - M_0)}{n} = \frac{x_1 + x_2 + \dots + x_n - nM_0}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} - \frac{nM_0}{n} = \bar{x} - M_0$$

Así, podemos afirmar que el método descrito en este ejercicio para calcular la media de un conjunto de datos funciona siempre.

PONTE A PRUEBA

Las pirámides de población

Problema resuelto

Respaldo al presidente

En Estadística se realizaron varios sondeos de opinión para conocer el nivel de respaldo al presidente en las próximas elecciones. Cuatro periódicos hicieron sondeos por separado en toda la nación. Los resultados de los sondeos de los cuatro periódicos se muestran a continuación:

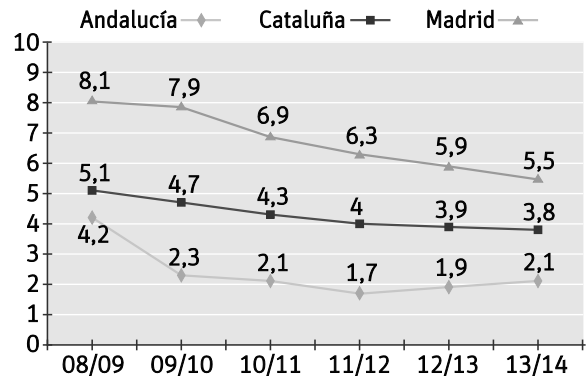
Periódico 1:	Periódico 2:	Periódico 3:	Periódico 4:
36,5 %	41,0 %	39,0 %	44,5 %
Fecha: 6 de enero	Fecha: 20 de enero	Fecha: 20 de enero	Fecha: 20 de enero
Muestra: 500 ciudadanos elegidos al azar y con derecho a voto.	Muestra: 500 ciudadanos elegidos al azar y con derecho a voto.	Muestra: 1000 ciudadanos elegidos al azar y con derecho a voto	Muestra: 1000 lectores que llamaron por teléfono para votar

Si las elecciones se celebraran el 25 de enero, ¿cuál de los resultados de los periódicos sería la mejor predicción del nivel de apoyo al presidente? Da dos razones que justifiquen tu respuesta.

Los resultados del periódico 3. Se encuentra entre los tres periódicos con la fecha de sondeo más próxima a las elecciones (20 de enero), utiliza la mayor cantidad de personas para la muestra junto al periódico 4 (1000 personas) y además es la muestra más representativa por tratarse de ciudadanos con derecho a voto elegidos al azar.

Disponibilidad de ordenadores

Según datos del INE (Instituto Nacional de Estadística), el número de alumnos por ordenador destinado a tareas de enseñanza-aprendizaje en las comunidades de Andalucía, Cataluña y Madrid desde el curso 2008/09 al 2013/14 viene dado por este gráfico.



1. ¿Cuántos alumnos por ordenador había en Andalucía en el curso 2010/11? ¿Cuántos en Madrid?

Andalucía (2010/11) = 2,1

Madrid (2010/11) = 6,9

2. ¿Cuánto ha disminuido el número de alumnos por ordenador en estos años en Cataluña?

$$5,1 - 3,8 = 1,3$$

En Cataluña han pasado de 5,1 alumnos por ordenador en el curso 2008/09 a 3,8 en el curso 2013/14.

3. En el curso 2013/14, ¿qué comunidad tenía una posición más favorable?

Andalucía, puesto que era la comunidad con menos alumnos por ordenador.

4. ¿Qué comunidad ha tenido una mejor evolución relativa?

$$\text{Madrid: } \frac{8,1 - 5,5}{8,1} = \frac{2,6}{8,1} = 0,32$$

$$\text{Andalucía: } \frac{4,2 - 2,1}{4,2} = \frac{2,1}{4,2} = 0,5$$

$$\text{Cataluña: } \frac{5,1 - 3,8}{5,1} = \frac{1,3}{5,1} = 0,25$$

Por tanto, la comunidad con mejor evolución relativa es Andalucía.

AUTOEVALUACIÓN

1. En un equipo de fútbol, el número de tarjetas amarillas acumuladas por cada uno de los jugadores que componen el equipo es:

1 1 0 0 0 2 3 1 4 2
3 2 2 1 1 0 0 3 1 2

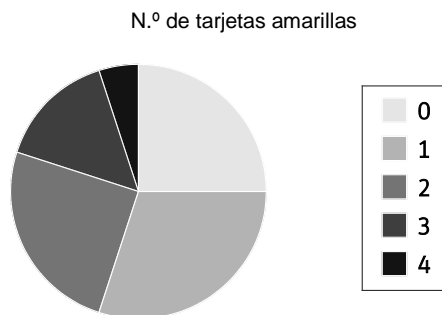
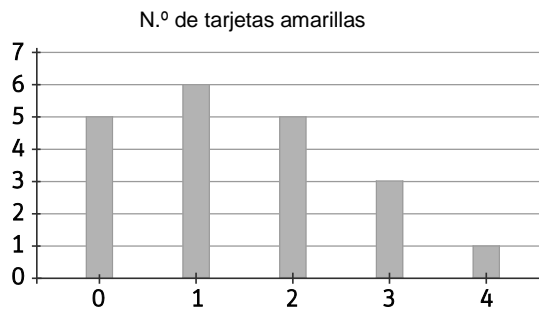
a) Escribe la tabla de frecuencias absolutas y relativas.

b) Dibuja el diagrama de barras y el de sectores.

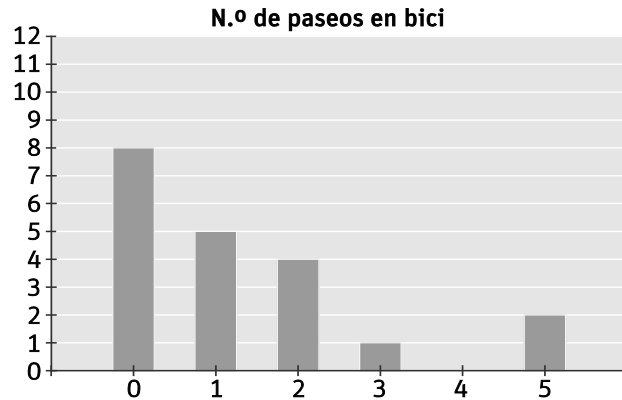
a)

Tarjetas amarillas	f_i	h_i	F_i	H_i	Amplitud del sector
0	5	0,25	5	0,25	90°
1	6	0,30	11	0,55	108°
2	5	0,25	16	0,80	90°
3	3	0,15	19	0,95	54°
4	1	0,05	20	1	18°
$N = 20$		1			

b)



2. En el gráfico siguiente se recoge el número de veces que un grupo de alumnos ha montado en bici en el último mes.



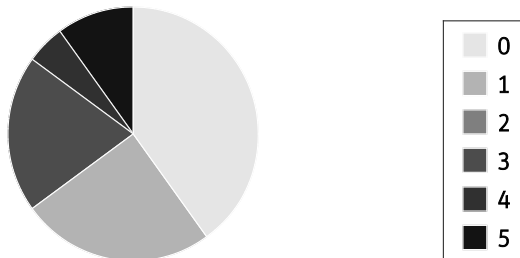
- Haz una tabla de frecuencias con estos datos.
- Construye el diagrama de sectores.
- Calcula la media.
- Halla la mediana y la moda.

a)

Paseos en bici	f_i	h_i	F_i	H_i	Amplitud del sector
0	8	0,40	8	0,40	144°
1	5	0,25	13	0,65	90°
2	4	0,20	17	0,85	72°
3	1	0,05	18	0,90	18°
4	0	0	18	0,90	0°
5	2	0,10	20	1	36°
	$N = 20$	1			

b)

N.º de paseos en bici



c) Media: $\bar{x} = \frac{0 \cdot 8 + 1 \cdot 5 + 2 \cdot 4 + 3 \cdot 1 + 4 \cdot 0 + 5 \cdot 2}{20} = \frac{26}{20} = 1,3$

- d) Puesto que tenemos un número par de datos, la mediana coincide con la media de los datos que ocupan las dos posiciones centrales, es decir, de los datos que ocupan las posiciones 10 y 11. En ambos casos, el dato toma el valor 1, por lo que la mediana es $M = 1$.

La moda coincide con el dato de mayor frecuencia absoluta. Por tanto, $M_0 = 0$.

3. Estas son las marcas, en segundos, de los atletas que han participado en una carrera de 400 m:

48,6 49,2 50,1 50,1 50,1 50,4 51
 51,8 51,8 51,9 52 52,2 52,6 53,28
 49,2 50,1 50,3 50,6 49,8 49,7 48,7
 49 50,2 50,5 50,7 51,1 49,2 49,6

- a) Agrupa los datos en intervalos de amplitud 0,5 y elabora la tabla de frecuencias.
- b) Calcula el intervalo modal y la media.
- c) Representa los datos en un histograma y en un polígono de frecuencias.

a)

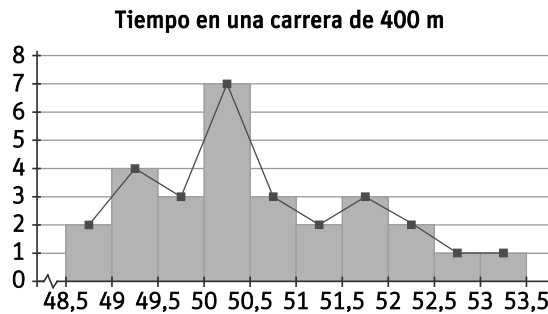
Tiempo (segundos)	Marca de clase, x_i	f_i	h_i	F_i	H_i
[48,5, 49)	48,75	2	0,071	2	0,071
[49, 49,5)	49,25	4	0,142	6	0,213
[49,5, 50)	49,75	3	0,1065	9	0,3195
[50, 50,5)	50,25	7	0,25	16	0,5695
[50,5, 51)	50,75	3	0,1065	19	0,676
[51, 51,5)	51,25	2	0,071	21	0,747
[51,5, 52)	51,75	3	0,1065	24	0,8535
[52, 52,5)	52,25	2	0,071	26	0,9245
[52,5, 53)	52,75	1	0,03775	27	0,96225
[53, 53,5)	53,25	1	0,03775	28	1
		$N = 28$	1		

b) El intervalo modal es el intervalo con mayor frecuencia absoluta. En este caso, [50, 50,5).

Para calcular la media, vamos a tener en cuenta la marca de clase de cada intervalo.

$$\bar{x} = \frac{48,75 \cdot 2 + 49,25 \cdot 4 + 49,75 \cdot 3 + 50,25 \cdot 7 + 50,75 \cdot 3 + 51,25 \cdot 2 + 51,75 \cdot 3 + 52,25 \cdot 2 + 52,75 \cdot 1 + 53,25 \cdot 1}{28} = \frac{1416}{28} = 50,57$$

c)



4. En una empresa, la antigüedad en años de la plantilla de 50 trabajadores viene dada por la tabla:

Años	1	2	3	4	5	6	7	8
Empleados	7	6	4	7	8	9	5	4

- a) Construye la tabla de frecuencias.
- b) Calcula la mediana y la moda.
- c) Calcula la media.

a)

Años	f_i	h_i	F_i	H_i
1	7	0,14	7	0,14
2	6	0,12	13	0,26
3	4	0,08	17	0,34
4	7	0,14	24	0,48
5	8	0,16	32	0,64
6	9	0,18	41	0,82
7	5	0,10	46	0,92
8	4	0,08	50	1
$N = 50$		1		

b) Como tenemos un número par de datos, la mediana se calcula como la media de los datos que ocupan las posiciones 25 y 26. En este caso, ambos datos valen 5, por lo que la mediana es $M = 5$ años.

La moda se corresponde con el dato de mayor frecuencia. Por tanto, $M_0 = 6$ años.

c) Media: $\bar{x} = \frac{1 \cdot 7 + 2 \cdot 6 + 3 \cdot 4 + 4 \cdot 7 + 5 \cdot 8 + 6 \cdot 9 + 7 \cdot 5 + 8 \cdot 4}{50} = \frac{220}{50} = 4,4$ años

5. Los minutos que hablan al día por teléfono los miembros de dos familias son:

Familia 1	75	65	71	56	59	63	70
Familia 2	93	90	70	69	68	72	71

Calcula la desviación media y la desviación típica de cada conjunto de datos.

Familia 1

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
75	9,4	9,4	88,36
65	-0,6	0,6	0,36
71	5,4	5,4	29,16
56	-9,6	9,6	92,16
59	-6,6	6,6	43,56
63	-2,6	2,6	6,76
70	4,4	4,4	19,36
		38,6	279,72

Media:

$$\bar{x} = \frac{75 + 65 + 71 + 56 + 59 + 63 + 70}{7} = \frac{459}{7} = 65,6$$

Desviación media: $\frac{38,6}{7} = 5,5$

Desviación típica: $s = \sqrt{\frac{279,72}{7}} = \sqrt{39,96} = 6,3$

Familia 2

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
93	16,9	16,9	285,61
90	13,9	13,9	193,21
70	-6,1	6,1	37,21
69	-7,1	7,1	50,41
68	-8,1	8,1	65,61
72	-4,1	4,1	16,81
71	-5,1	5,1	26,01
		61,3	674,87

Media:

$$\bar{x} = \frac{93 + 90 + 70 + 69 + 68 + 72 + 71}{7} = \frac{533}{7} = 76,1$$

Desviación media: $\frac{61,3}{7} = 8,8$

Desviación típica: $s = \sqrt{\frac{674,87}{7}} = \sqrt{96,41} = 9,8$