

●
cruïlla
●

10. Estadística

1 Mostres. Representativitat

- La **població** és el conjunt de tots els elements sobre els quals es vol realitzar un estudi estadístic.
- Quan la població és molt gran, se n'agafa una **mostra**, que és un petit grup d'elements de tota la població.
- És molt important escollir bé la mostra perquè les dades obtingudes siguin **significatives**.

245 *En un institut hi ha 125 alumnes de Batxillerat (50 nois i 75 noies) i 375 alumnes de Secundària (175 nois i 200 noies). Si es vol agafar una mostra de 40 alumnes per fer una enquesta, quants alumnes (nois i noies) caldrà escollir de Batxillerat i quants de Secundària?*

En aquesta mostra de 40, cada individu representa $\frac{500}{40} = 12,5$ alumnes.

Per tant, per tal que la mostra sigui representativa s'ha d'escollir:

$$\frac{125}{12,5} = 10 \text{ alumnes de Batxillerat, dels quals, } \frac{50}{12,5} = 4 \text{ nois i } \frac{75}{12,5} = 6 \text{ noies.}$$

$$\frac{375}{12,5} = 30 \text{ alumnes de Secundària, dels quals, } \frac{175}{12,5} = 14 \text{ nois i } \frac{200}{12,5} = 16 \text{ noies.}$$

246 Digues si en els estudis estadístics següents cal agafar una mostra o tota la població.

Exemple *La durada de les bombetes produïdes per una empresa \Rightarrow MOSTRA (si provéssim la durada de totes les bombetes, no se'n podria vendre cap).*

- a) L'alçada dels joves de la teva comunitat autònoma.
- b) L'edat mitjana de les persones que viuen a casa teva.
- c) La resistència dels motors que produeix una fàbrica.
- d) La puresa de la llet que es recull en un dia en una central lletera.
- e) La nota mitjana de Matemàtiques dels teus companys de classe.
- f) El gruix dels arbres d'un bosc.
- g) La nota mitjana de les teves notes de la primera avaluació.
- h) Els quilòmetres diaris recorreguts per un autobús durant el mes d'agost.

247 En una ciutat hi ha 8.000 joves, dels quals 3.600 són nois, i la resta, noies. Quants se n'ha de triar de cada sexe per realitzar una mostra representativa de 400 persones?

248 Als instituts de Manresa es vol saber quines són les assignatures que més agraden als alumnes de Secundària. Per fer-ho s'agafa una mostra de 400 alumnes. Si saps que en aquella població hi ha el nombre d'alumnes següent:

1r ESO	2n ESO	3r ESO	4t ESO
1.080	1.060	970	890

Calcula com s'hauria de prendre la mostra per tal que fos representativa.

249 A Viladecans, l'ajuntament vol construir un poliesportiu. Per fer-ho escull 9 homes i 11 dones perquè aportin les seves idees. Si saps que al poble hi ha 15.000 persones i que la mostra és representativa, quantes dones viuen a Viladecans?

250 Per conèixer els gustos sobre el lloc favorit per anar de vacances –platja, muntanya o ciutat– s'escull una mostra entre els estiuejants que hi ha a Lloret, Cadaqués i Torredembarra. Creus que l'enquesta serà fiable? Per què?

2 Taules de freqüències

- **Freqüència absoluta (f_i)** és el nombre de vegades que apareix un determinat valor.
- **Freqüència relativa (n_i)** és el quocient entre la seva freqüència absoluta i el nombre total de dades.
- **Freqüència acumulada (F_i)** és la suma de les freqüències absolutes de tots els valors inferiors o iguals que el valor considerat. També es pot calcular com la suma de la freqüència acumulada del valor anterior més la freqüència absoluta d'aquest valor.

251 Agafada una mostra de 20 persones entre els alumnes de Secundària d'un institut, les seves edats són: 12, 14, 15, 13, 16, 14, 14, 13, 12, 15, 12, 15, 14, 14, 13, 13, 16, 15, 13, 14. Calcula'n la taula de freqüències.

Primer se n'ha de fer el recompte i després construir la taula. Hi ha 3 alumnes amb 12 anys, 5 amb 13, 6 amb 14, 4 amb 15 i 2 amb 16. Per tant, la taula és la següent.

Edat (x_i)	f_i	n_i	F_i	F_i
12	3	$\frac{3}{20} = 0,15$	3	3
13	5	$\frac{5}{20} = 0,25$	$3 + 5 = 8$	$3 + 5 = 8$
14	6	$\frac{6}{20} = 0,3$	$8 + 6 = 14$	$3 + 5 + 6 = 14$
15	4	$\frac{4}{20} = 0,2$	$14 + 4 = 18$	$3 + 5 + 6 + 4 = 18$
16	2	$\frac{2}{20} = 0,1$	$18 + 2 = 20$	$3 + 5 + 6 + 4 + 2 = 20$

Es pot observar que amb les dues maneres de calcular la freqüència acumulada s'obté el mateix, i per això s'usa la primera, que és més ràpida.

252 En un grup de persones hi ha 28 noies i 22 nois. Calcula'n les freqüències relatives.

253 En una classe de 4t d'ESO, el nombre d'assignatures suspeses de cada alumnes és:

4, 1, 0, 0, 4, 0, 2, 2, 3, 2, 2, 0, 0, 4, 3, 5, 1, 0, 1, 1, 0, 3, 2, 0, 2

Elabora'n la taula de freqüències.

254 Omple la taula en què es mostren les freqüències acumulades de les edats dels alumnes d'una escola infantil.

<i>Edat</i>	F_i	f_i	n_i
0	8		
1	18		
2	30		
3	45		

255 En un estudi realitzat a 200 persones sobre si estiuegen a Espanya o a l'estranger, s'han obtingut les freqüències relatives següents: n (Espanya) = 0,8 ; n (estranger) = 0,2. Calcula quantes persones estiuegen a Espanya i quantes a l'estranger.

256 En un estudi realitzat a 100 nens i nenes sobre el nombre de vasos de llet que beuen al dia, s'ha obtingut la taula següent. Calcula'n les dades que hi falten (a , b i c).

N. de vasos	f_i	n_i
1	20	0,2
2	25	0,25
3	a	b
4	10	0,1
5	c	0,05

257 Completa les dades que falten en aquesta taula.

x_i	f_i	F_i	n_i
1	8		0,16
2		20	
3		30	
4	6		
5		50	

3 Gràfics estadístics. Gràfics múltiples. Diagrames de caixes

- Segons l'estudi estadístic que es vulgui fer, hi ha molts tipus de **gràfics estadístics**. Els més habituals són:

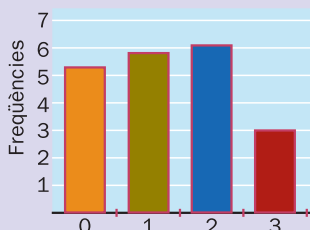


Diagrama de barres



Polígon de freqüències

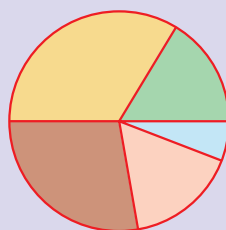


Diagrama de sectors

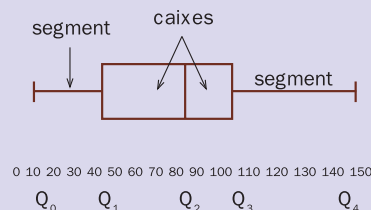


Diagrama de caixes

- De vegades, en un únic dibuix hi ha més d'un gràfic. En aquest cas s'anomenen **gràfics múltiples**.

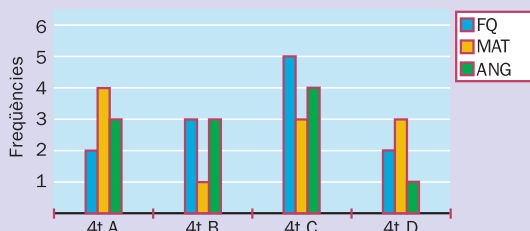


Diagrama de barres múltiple

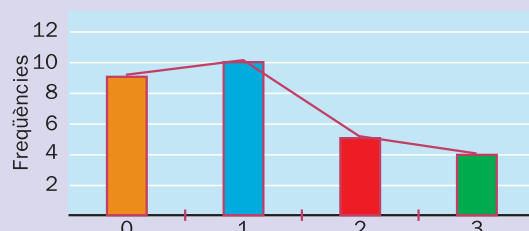


Diagrama de barres i polígon de freqüència

258 Les notes (ordenades de més baixa a més alta) d'un examen de matemàtiques van ser:

0, 1, 2, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 7, 7, 7, 7, 8, 8, 8, 8, 8, 9, 9, 9, 10

Dibuixa'n el diagrama de caixes.

Per poder dibuixar-lo dividim les dades en quatre parts iguals.

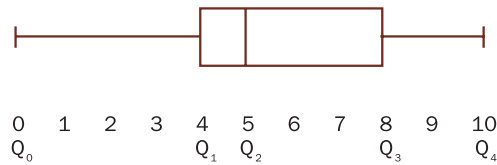
0, 1, 2, 3, 3, 4, 4, 4 4, 4, 5, 5, 5, 5, 5, 5 5, 6, 6, 7, 7, 7, 7, 8 8, 8, 8, 8, 9, 9, 9, 10

Els dos blocs centrals es representen en caixes, i els laterals, en segments.

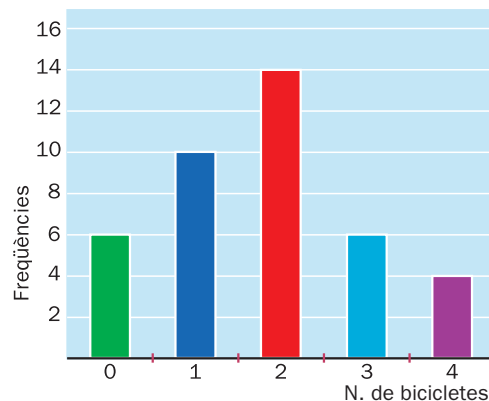
Q_1 és el primer valor del segon grup; per davant seu hi ha un 25% de les dades. És un quartil (una quarta part).

Q_2 és la mitjana de la distribució.

Q_3 és l'últim valor del tercer grup; per davant seu hi ha un 75% de les dades.



259 En una mostra de 40 llars s'ha preguntat pel nombre de bicicletes que hi ha a la casa i se n'ha obtingut el gràfic adjunt. Elabora una taula de freqüències (absolutes, relatives i acumulades) amb aquestes dades.



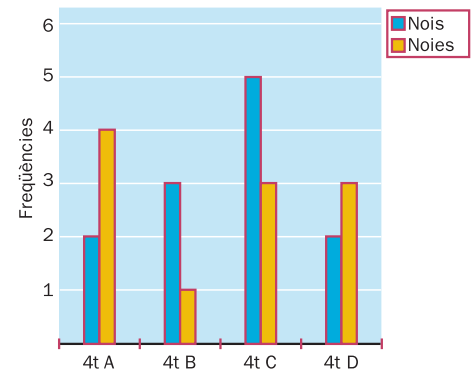
260 Aquesta taula indica el nombre d'avis que tenen els alumnes d'una classe.

Avis	0	1	2	3	4
Alumnes	2	5	7	12	4

Dibuixa el diagrama de barres corresponent.

261 El gràfic següent mostra el nombre d'excel·lents aconseguits pels alumnes que hi ha en els quatre grups de 4t d'ESO d'un institut. Omple la taula de freqüències amb els valors que dona el gràfic.

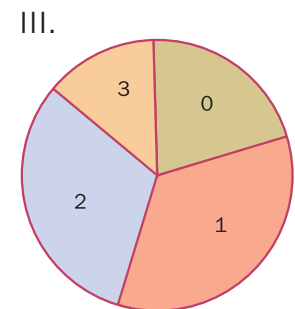
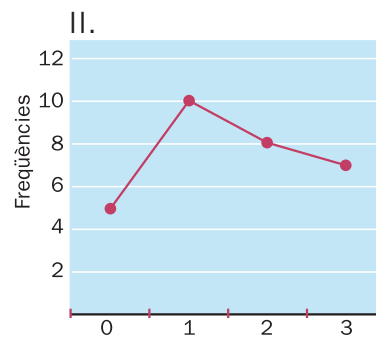
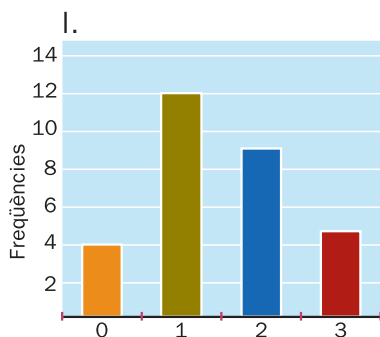
	4t A	4t B	4t C	4t D
Nois				
Noies				



262 Preguntats sobre el nombre de llibres que han llegit el darrer mes els alumnes de tres grups de Secundària, s'han obtingut aquestes dades.

	0	1	2	3
Grup A	4	12	9	5
Grup B	6	10	10	4
Grup C	5	10	8	7

Associa a cada gràfic el grup corresponent.



4 Taules i gràfics quotidians. Fal·làcies

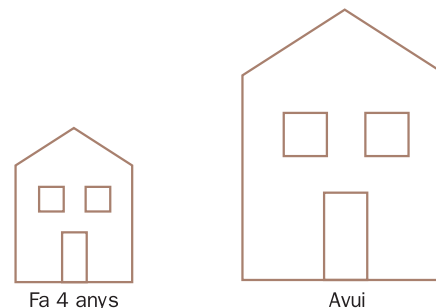
- En llibres, diaris i altres mitjans de comunicació es mostren moltes taules i gràfics estadístics perquè el lector sigui capaç de visualitzar la informació.
- Els més usats solen ser els anomenats temporals, en què el valor de la variable x representa el temps (dies, anys...). No obstant això, s'ha d'anar amb molt de compte a l'hora d'observar aquests gràfics perquè, intencionadament o no, poden provocar interpretacions errònies.

263 *Suposem que segons alguns estudis recents, en els darrers 4 anys s'ha duplicat el preu de l'habitatge. En una determinada revista, el gràfic presentat amb aquestes dades ha estat el de la imatge.*

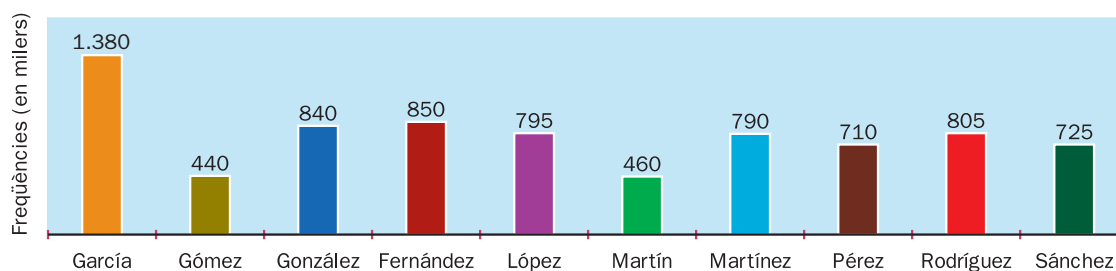
Per què creus que el gràfic no és correcte?

S'ha multiplicat per totes les longituds de la figura. I per això l'àrea de la nova casa és 4 vegades més gran que l'àrea de l'anterior.

En veure el dibuix sembla que hagi augmentat molt més que el doble.



264 En aquest gràfic es mostren els 10 cognoms més comuns a Espanya.



- Ordena els cognoms de més a menys freqüència.
- Quin és el cognom més freqüent? Quantes persones tenen aquest cognom?
- Quanta gent en total té aquests cognoms?
- Pot provocar el gràfic interpretacions errònies?

265 Fa temps, un anunci de pasta de dents deia que dos de cada tres dentistes preguntats recomanaven aquella pasta de dents.

a) Per què creus que amb aquesta informació no cal animar-se gaire a comprar aquell producte?

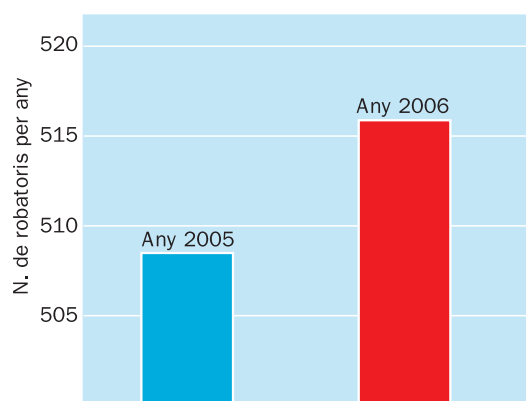
b) Què pensaries si sabessis que només s'ha entrevistat tres dentistes?

c) I si els entrevistats fossin 3.000 dentistes?

d) En realitat, van ser-ne 67. Què n'opines ara, de l'anunci?

266 Un presentador va ensenyar el gràfic següent i va dir: «El gràfic mostra que hi ha un augment enorme del nombre de robatoris si es comparen els anys 2005 i 2006.»

Creus que aquesta afirmació és una interpretació raonable del gràfic? Per què?



5 Paràmetres de centralització: mitjana aritmètica, moda i mediana

Els principals paràmetres de centralització són:

- La **mitjana aritmètica** (\bar{x}), que s'obté aplicant la fórmula: $\bar{x} = \frac{\sum f_i \cdot x_i}{N}$.
- La **moda** (**Mo**), que és el valor (o valors) que té freqüència absoluta més gran.
- La **mediana** (**Me**), que és el valor que ocupa el lloc central de totes les dades en ordenar-les de més petites a més grans. Per calcular-la es divideix el nombre de dades (N) entre 2. Es busca a la taula de freqüències la primera freqüència acumulada (F_i) més gran que aquell quocient. El valor corresponent a aquella freqüència acumulada és la mediana.

267 En un estudi basat en 250 nadons sobre l'edat en què comencen a caminar, s'ha obtingut la taula següent.

Calcula'n els paràmetres centrals.

Mesos	N. de nadons
11	30
12	87
13	80
14	34
15	19

Primer completem la taula de freqüències que ens permetrà obtenir la mediana i la mitjana.

x_i	f_i	F_i	$f_i \cdot x_i$
11	30	30	330
12	87	117	1.044
13	80	197	1.040
14	34	231	476
15	19	250	285
TOTAL	250		3.175

Moda: 12 (la freqüència absoluta més gran és 87, que correspon al valor 12) \Rightarrow **Mo = 12**

Mediana: $\frac{N}{2} = \frac{250}{2} = 125 \Rightarrow F_i > 125$ és 197 \Rightarrow **Me = 13** (el valor F_i del qual és 197)

Mitjana aritmètica: $\bar{x} = \frac{\sum f_i \cdot x_i}{N} = \frac{3.175}{250} = 12,7 \Rightarrow$ **$\bar{x} = 12,7$**

268 Un dau trucat es tira 200 vegades i s'obtenen els resultats següents.

Cares	1	2	3	4	5	6
f_i	12	32	54	18	24	60

Calcula'n la mitjana, la mediana i la moda.

269 El nombre de fills de 25 famílies és:

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

Calcula'n la mitjana, la mediana i la moda.

270 En la sèrie de dades següent: 1, 2, 3, 3, 4, 4, 4, 5, 5, 6, es canvia el 6 pel 10. Quin paràmetre central canvia? Quins queden igual?

271 *Calcula per quin nombre s'ha de substituir la dada a perquè sigui 6 la mitjana aritmètica del conjunt de nombres següent.*

5 7 a 6 4

Hi ha 5 dades la suma de les quals és $5 + 7 + 6 + 4 + a = 22 + a$.

Perquè la mitjana sigui 6, ha de passar que la suma sigui $\left(6 = \frac{30}{5}\right)$.

Per tant, $22 + a = 30 \Rightarrow a = 8$.

272 Per posar la nota, un professor fa quatre exàmens. En els tres primers, un alumne ha aconseguit 5, 3, 4. Quina nota ha de treure com a mínim per aprovar?

273 Les notes dels 10 millors alumnes d'una classe han estat 6, 8, 6, 5, 7, 9, 8, 7, 6, x. Quant ha de valer x perquè la mitjana sigui 7?

6 Paràmetres de dispersió: rang, variància, desviació típica i coeficient de variació

Els paràmetres de dispersió més usuals són:

- El **rang o recorregut** d'una distribució estadística és la diferència entre la dada més gran i la més petita d'aquesta distribució.
- La **variància** (S^2) s'obté mitjançant la fórmula $S^2 = \frac{\sum f_i \cdot (x_i - \bar{x})^2}{N}$ o la fórmula $S^2 = \frac{\sum f_i \cdot x_i^2}{N} - \bar{x}^2$.
- L'arrel quadrada positiva de la variància s'anomena **desviació típica** (S).
- Per comparar diferents desviacions típiques s'usa el **coeficient de variació** (CV): $CV = \frac{S}{\bar{x}}$.

274 En un estudi basat en 250 nadons sobre l'edat en què comencen a caminar s'ha obtingut la taula següent.

Calcula'n els paràmetres de dispersió.

Primer elaborem la taula que ens fa falta per obtenir la variància (observa que la mitjana aritmètica s'ha obtingut en l'apartat anterior i és 12,7).

Mesos	N. d'infants
11	30
12	87
13	80
14	34
15	19

x_i	f_i	$f_i \cdot x_i^2$
11	30	3.630
12	87	12.528
13	80	13.520
14	34	6.664
15	19	4.275
TOTAL	250	40.617

$$\text{Rang: } 15 - 11 = 4$$

$$\text{Variància: } S^2 = \frac{\sum f_i \cdot x_i^2}{N} - \bar{x}^2 = \frac{40.617}{250} - 12,7^2 = 1,178$$

$$\text{Desviació típica: } S = \sqrt{1,178} = 1,085$$

$$\text{Coeficient de variació: } CV = \frac{S}{\bar{x}} = \frac{1,085}{12,7} = 0,085$$

275 Indica en cada cas el recorregut d'aquestes dades.

Exemple 8, 5, 4, 9, 6, 5, 7, 7, 8, 4 \Rightarrow La dada més gran és 9, i la més petita, 4; per tant, $9 - 4 = 5$.

- 5, 2, 6, 2, 4, 6, 4, 3, 7, 2, 6, 5
- 2, 4, 4, 5, 5, 5, 7, 7, 8, 12
- 17, 21, 42, 38, 12, 29, 18, 37, 31, 40, 22
- 8, 8, 8, 8, 8, 8, 8, 8

276 En cada una de les nits d'un mes, el nombre d'urgències realitzat pels diferents metges d'un gran hospital van ser:

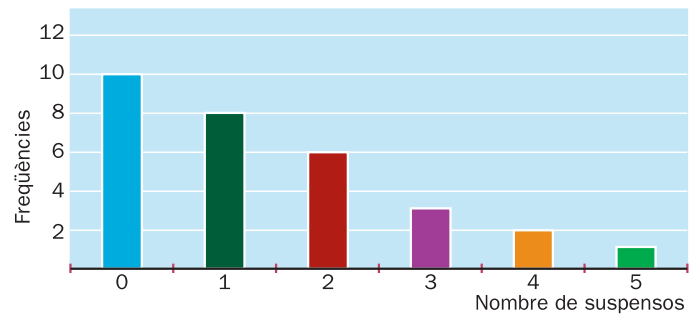
17, 15, 15, 16, 14, 15, 15, 16, 15, 14, 14, 15, 17, 15, 15, 16, 14, 15, 15, 14, 18, 15, 16, 14, 15

Calcula'n la variància i el coeficient de variació.

277 En una classe amb 30 alumnes, el nombre de suspensos en la primera avaluació ve donat pel diagrama de barres de la dreta.

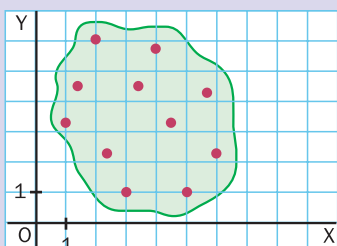
Calcula:

- a) La seva variància
- b) El seu coeficient de variació

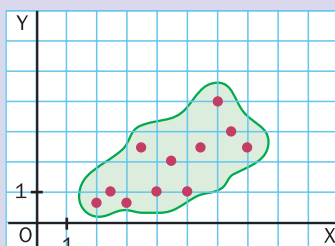


7 Distribució bidimensional. Núvol de punts

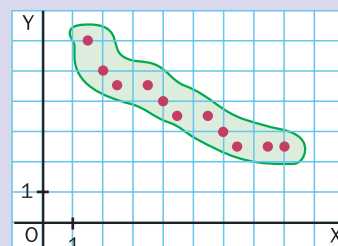
- L'**estadística bidimensional** mostra la relació entre dues variables estadístiques. Se solen representar per X i Y . Les dades vénen donades per parells de valors (x_i, y_i) .
- Un **núvol de punts** és la representació gràfica dels parells de valors de les variables.
- Els núvols de punts indiquen el grau de dependència entre les variables.



Núvol de punts sense dependència



Dependència dèbil

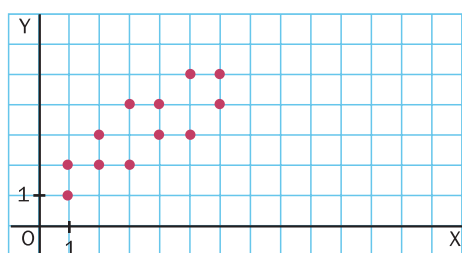


Dependència forta

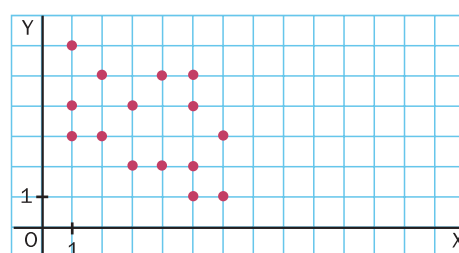
278 Cinc nenes de 2, 3, 5, 7 i 8 anys d'edat pesen, respectivament, 14, 20, 30, 42 i 44 quilos. Fes el núvol de punts i indica si la dependència entre les variables és dèbil o forta.

279 Digues si els núvols de punts següents presenten una dependència forta o dèbil.

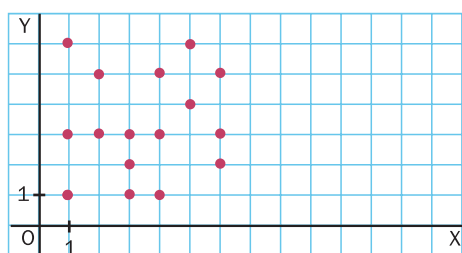
a)



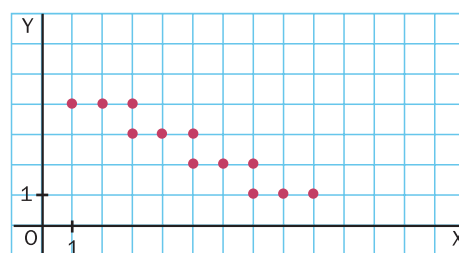
c)



b)



d)



► COMPROVA EL QUE HAS APRÈS

10. Estadística

- 1 Preguntades 25 persones sobre el nombre d'hores que miren la televisió cada dia, s'han obtingut aquestes dades: 3, 2, 0, 1, 5, 5, 3, 2, 3, 3, 4, 4, 2, 1, 1, 3, 2, 2, 4, 1, 2, 3, 3, 2, 5. Omple aquesta taula.

x_i	f_i	F_i	n_i	$f_i \cdot x_i$	$f_i \cdot x_i^2$
0					
1					
2					
3					
4					
5					
TOTAL					

- 2 Calcula'n la mitjana aritmètica i la desviació típica de l'activitat anterior.

- 3 La taula mostra la distribució per pesos (en quilograms) dels 24 alumnes homes d'una classe. Completa la taula i calcula la mitjana i la desviació típica.

	x_i	f_i	$f_i \cdot x_i$	$f_i \cdot x_i^2$
56-60		1		
60-64		6		
64-68		10		
68-72		5		
72-76		2		

- 4 En una botiga de cotxes, la mitjana dels preus d'un grup de cotxes és de 16.500 €, i la desviació típica és de 2.970 €. En una rellotgeria, els preus d'un grup de rellotges tenen una mitjana de 360 €, amb una desviació típica de 90 €. Quins preus són més dispars, els dels cotxes o els dels rellotges?