

TEMA 7: Estadística

Activitats

Estadística

L' estadística s'ocupa de recollir i ordenar dades referides a diversos fenòmens per després analitzar-les, interpretar-les i, en alguns casos, fer previsions

Fases:

- a) Planificació de l'estudi i de les fases
- b) Recollida de dades
- c) Expressió de les dades en taules i gràfics
- d) Anàlisi i interpretació a través del càlcul de paràmetres matemàtics (constants numèriques que ens donen informació de les característiques de l'estudi)
- e) Previsió de resultats en relació a la característica indicada

Elements:

- Població: conjunt format per tots els elements de l'estudi
- Mostra: part de la població que forma l'estudi
- Individu: cada element individual que forma part de l'estudi
- Mida o grandària de la mostra (N): nombre d'individus que formen la mostra
- Variable: característica estudiada

EXEMPLE:

Es vol fer un estudi estadístic en relació al pes dels alumnes matriculats a 4t d'ESO a Catalunya. Per realitzar l'estudi s'han seleccionat 300 alumnes de diferents centres escolars.

Elements:

Població: tots els alumnes matriculats a 4t d'ESO a Catalunya

Mostra: els 300 alumnes als que es pesarà

Individu: cada alumne que forma part de la mostra

Mida: 300

Variable estadística: pes

Fases:

D'una correcta planificació i recollida de dades dependrà la fiabilitat dels resultats.

En la planificació caldrà determinar en quins instituts es fa l'estudi, com s'esculleixen, és important que corresponguin a localitats petites o grans ciutats? s'ha de pesar al mateix nombre de noies que de nois? pot alguna d'aquestes coses modificar els resultats?.

En la recollida de dades haurem de discutir com es recullen les dades: es pregunta als alumne el seu pes o es pren el pes als alumnes?, si és així: qui ho fa?, a quina hora?, amb roba o sense?, amb diferents bàscules?, en quina unitat?, fins a quina xifra decimal es pren la mesura?, ...

1. En un estudi sobre l'edat (en mesos) a la que els nens comencen a caminar, hem escollit 50 nens de la nostra població. Determineu: població, mostra, individus, mida i variable estadística.
2. Digueu en quin cas és més convenient estudiar la població o una mostra:
 - a) l'altura dels turistes que visiten Catalunya en un any
 - b) el color d'ulls dels jugadors de bàsquet de l'equip del nostre poble
 - c) el nombre de fills que tenen les dones europees entre 30 i 40 anys
 - d) el nombre de bombetes defectuoses que fabrica una màquina de forma contínua durant tot el mes.
3. En una ciutat de 100.000 persones s'ha passat una enquesta a 500 per saber quins són els seus esports favorits. Digueu quina és la població i quina és la mostra.
4. En una gran població A s'agafa una mostra de 500 persones per fer un estudi i a una altra població més petita B també s'agafa una mostra de 500 persones pel mateix estudi. Els resultats de l'estudi seran igualment de fiables?. Raoneu la resposta.

Tipus de variables

- a) Qualitatives. Els valors de les variables no es pot associar amb un nombre. Exemple: sexe (home o dona)
- b) Quantitatives. Els valors s'associen amb un nombre. Hi ha de dos tipus:
 - Discretes. Només poden prendre determinats valors. Exemple: el nombre de fills pot ser 0, 1, 2, ... però no 3,27
 - Contínues. Entre dos nombres la variable pot prendre infinits valors. Exemple: l'alçada entre 1,70 m i 1,80 m potser 1,74m, 1,654 m, ...

5. Completa la taula amb la població i les característiques dels següents estudis estadístics:
 - a) Alçada dels alumnes de 4t d'ESO del centre.
 - b) Raça de gos preferida pels menors de 20 anys de la teva ciutat.
 - c) Color d'ulls dels alumnes de la teva classe.
 - d) Nombre de telèfons mòbils per família a les poblacions de la teva comarca.

	Població	Variable estadística	Qualitativa	Quantitativa	
				Discreta	Contínua
A					
B					
C					
D					

6. Els alumnes de 4t d'ESO de l'institut fan diverses enquestes als seus companys per saber els seus gustos sobre diferents qüestions. Digues, en cada cas, si es tracta d'una variable qualitativa o una variable quantitativa.
- tipus de pel·lícules que els agrada més
 - hores diàries que juguen amb videojoc
 - nombre de banys que hi ha a casa seva
 - color preferit
7. Inventa't 2 variables quantitatives discretes i 2 variables quantitatives contínues. Han de ser diferents a les dels exercicis anteriors.

Taules de freqüències per variables qualitatives o quantitatives discretes

Les dades recollides es recompten i es recullen els resultats en taules.

A les taules tenim:

- freqüència absoluta (n_i). És el nombre de vegades que apareix un determinat valor o tipus de la característica estudiada
- freqüència relativa (f_i). És el resultat de la divisió entre la freqüència absoluta i

la mida de la mostra. $f_i = \frac{n_i}{N}$

La suma de les freqüències relatives ha de ser 1.

- percentatge (%). És el resultat de multiplicar la freqüència relativa per 100.
 $\% = f_i \cdot 100$

La suma de percentatges ha de ser 100.

En el cas de variables quantitatives a més hi ha

- freqüència absoluta acumulada (N_i) de una dada x_i és la suma de les freqüències absolutes dels valors més petits o iguals a aquesta dada

$$N_i = n_1 + n_2 + \dots + n_i$$

- freqüència relativa acumulada (F_i) de una dada x_i és la suma de les freqüències relatives dels valors més petits o iguals a aquesta dada

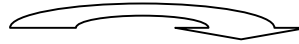
$$F_i = f_1 + f_2 + \dots + f_i$$

EXEMPLE:

Les notes de matemàtiques en un grup de 20 alumnes han estat:

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

X 100



x_i	n_i	N_i	$f_i = \frac{n_i}{N}$	F_i	$\% = f_i \cdot 100$
2	3	3	$0,15 = 3/20$	0,15	15
3	3	6	$0,15 = 3/20$	0,30	15
4	4	10	$0,20 = 4/20$	0,50	20
5	3	13	$0,15 = 3/20$	0,65	15
6	1	14	$0,05 = 1/20$	0,70	5
7	2	16	$0,10 = 2/20$	0,80	10
8	2	18	$0,10 = 2/20$	0,90	10
9	2	20	$0,10 = 2/20$	1,00	10
Suma total	N=20		1,00		100

8. L'edat dels pacients d'un logopeda són:

9 8 11 9 9 10 10 8 9 12

Elaboreu una taula de freqüències

9. Anota el color de cabells dels companys de la teva classe i elabora una taula de freqüències. En aquesta taula apareixeran les columnes de freqüències acumulades?. Justifiqueu la resposta

10. Completa la següent taula de freqüències:

x_i	n_i	N_i	f_i	%
1	4		0'08	
2	4			
3		16	0'16	
4	7		0'14	
5	5	28		
6				
7	7	45	0'14	
8				

Gràfics estadístics per variables qualitatives i quantitatives discretes

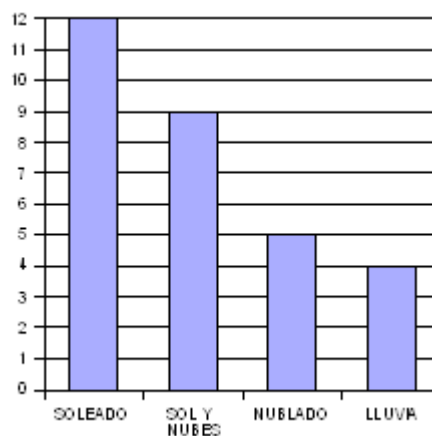
Una altre forma d'ensenyar els resultats de les dades recollides són les representacions gràfiques. Les més importants són:

- a)diagrama de barres
- b)diagrama de sectors
- c)diagrama de línies
- d)pictogrames

a) Diagrama de barres

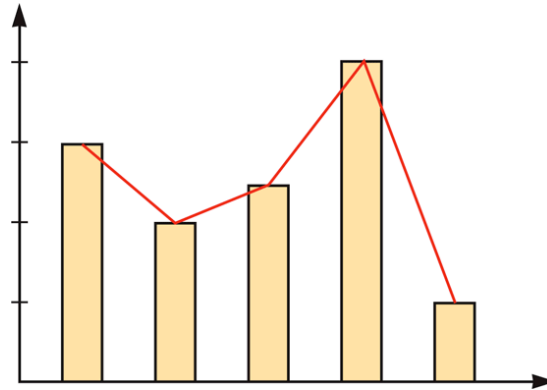
A l'eix horitzontal representem els valors de la variable i a l'eix vertical les freqüències

Ex: Estat del temps a una població durant un mes



Quan la variable aleatòria és quantitativa, podem unir amb línies els extrems superiors de les barres i obtenir un **polígon de freqüències**.

Ex:



b) Diagrama de sectors

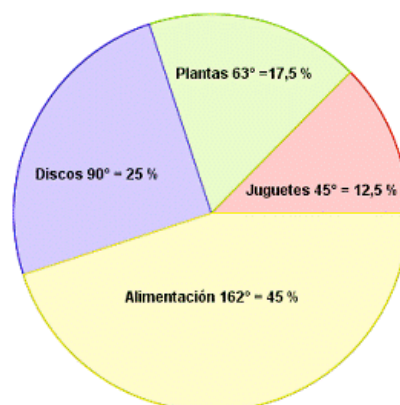
Les dades es representen en un cercle dividit en sectors que representen els valors de la variable. L'angle de cada sector és proporcional a la freqüència de la dada que representa

Per calcular l'angle del sector circular s'aplica la fórmula:

$$\text{Angle} = \frac{n_i}{N} \cdot 360^\circ = f_i \cdot 360^\circ$$

Ex: En una botiga s'han venut joguines per valor de 125 €, plantes per 175€, música per 250€ i alimentació per 450€

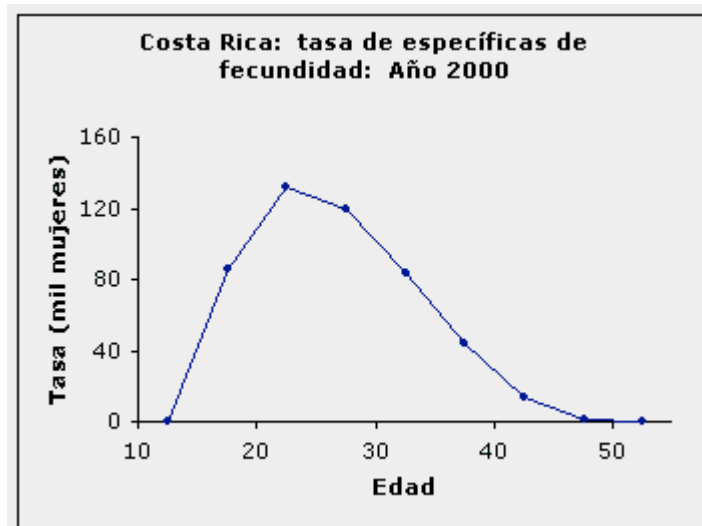
x_i	n_i	$f_i = \frac{n_i}{N}$	Angle = $f_i \cdot 360^\circ$
Joguines	125	0,125	$45^\circ = 0,125 \cdot 360^\circ$
Plantes	175	0,175	$63^\circ = 0,175 \cdot 360^\circ$
Música	250	0,250	$90^\circ = 0,250 \cdot 360^\circ$
Alimentació	450	0,450	$162^\circ = 0,450 \cdot 360^\circ$
	N= 1000	1,000	360°



c) Diagrama de línies

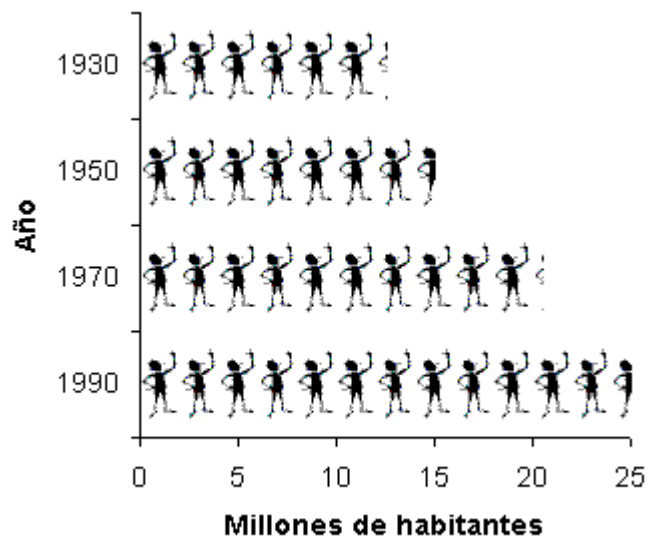
S'utilitza per a variables qualitatives quan volem representar l'evolució d'una variable en relació al temps

Ex:



d) Pictogramas

S'utilitzen figures relacionades amb la variable estudiada.



11. El nombre de fills de totes les famílies dels alumnes d'una aula és:

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

- a) Indica quina variable s'estudia i de quin tipus és.
- b) Quin és el nombre d'individus de la mostra?
- c) Fes una taula estadística.
- d) Representa les dades en un diagrama de barres

12. S'ha preguntat als alumnes d'una classe quants llibres han llegit durant l'estiu. La resposta és:

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

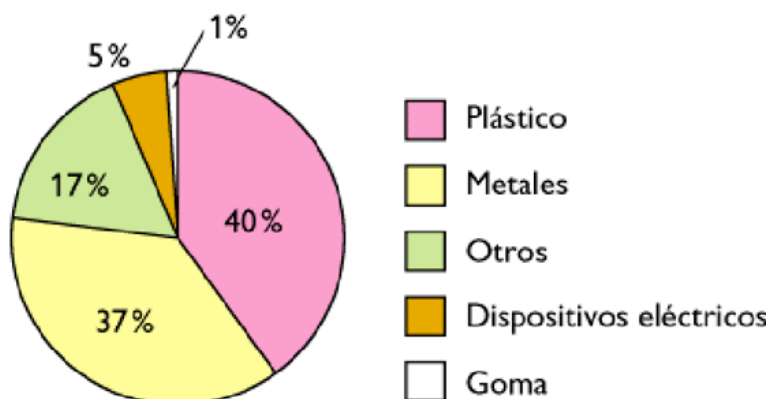
- a) Indica quina variable s'estudia i de quin tipus és.
- b) Quin és el nombre d'individus de la mostra?
- c) Fes una taula estadística.
- d) Representa les dades en un diagrama de barres i en un diagrama de sectors

13. Hem preguntat a un grup d'alumnes que esmorzaven normalment i ens han contestat:

Cereals, fruita, fruita, galetes, cereals, cereals, torrades, galetes, fruita, cereals, galetes, galetes, torrades, galetes, cereals, galetes, galetes, fruita, torrades, galetes

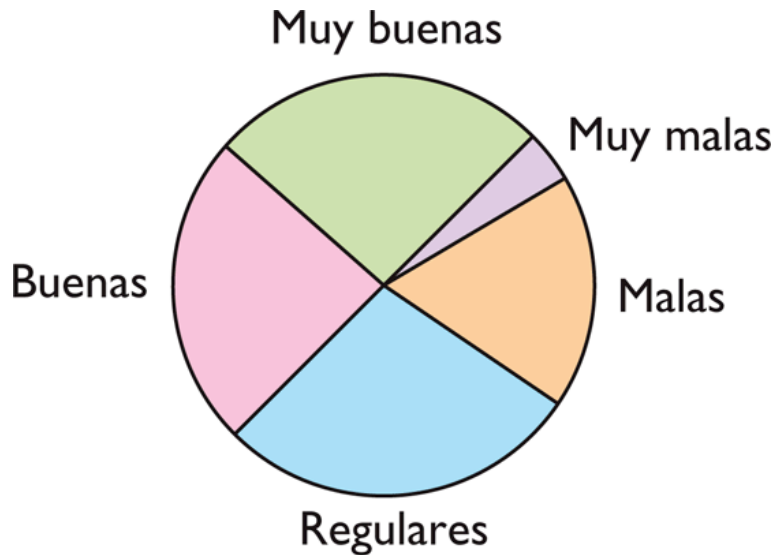
- a) Indica quina variable s'estudia i de quin tipus és.
- b) Quin és el nombre d'individus de la mostra?
- c) Fes una taula estadística.
- e) Fes un diagrama de sectors que representi la informació recollida.

14. El següent diagrama de sectors mostra les despeses de diferents tipus de materials que s'han donat a una empresa per la fabricació d'un producte



Si les despeses totals han estat de 35000 €, quina ha estat la despesa en plàstic? I en metalls?

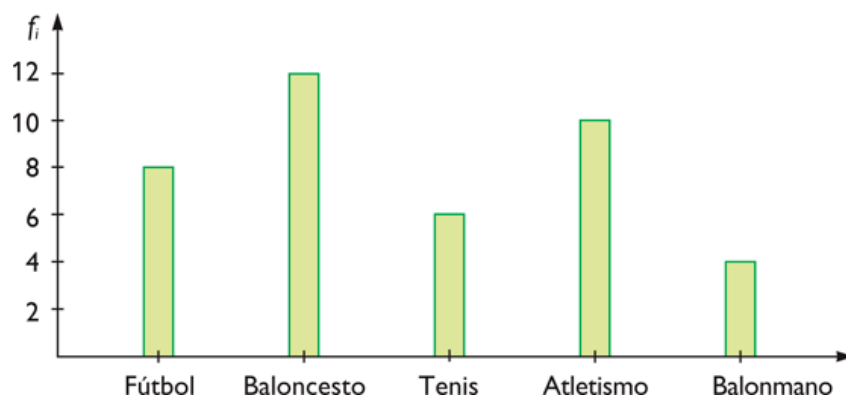
15. En el següent gràfic es mostra els resultats d'una enquesta realitzada en una empresa de 800 empleats en relació a l'opinió dels treballadors respecte a les condicions de la seva feina.



Amb ajut d'un transportador d'angles completeu la taula:

Variable	Angle	$f_i = \frac{Angle}{360^\circ}$	%	Nº de treballadors que tenen aquesta opinió
Molt bones				
Bones				
Regulars				
Dolentes				
Molt dolentes				

16. De l'enquesta a un grup de persones en relació al seu esport favorit els resultats han estat:



- a) A quantes persones s'ha fet l'enquesta?
 b) Quin és l'esport amb més afició? Quin és l'esport amb menys afició?
 c) Completeu la taula:

Esport	n _i	f _i	%
Fútbol			
Bàsquet			
Tenis			
Atletisme			
Balonmano			

17. En un article de "La Vanguardia" es parla de l'adopció de nens a Espanya i a Catalunya



- a) Si l'any 2002 es van adoptar 3625 nens a tota Espanya, quants corresponien a Xina?
 b) I a Rússia?
 c) L'any 2000 el nombre d'adopcions a Catalunya va ser cinc vegades menor que a Espanya?
 d) Estudiem per continents el nombre d'adopcions a Catalunya l'any 2002. Per això utilitzarem el quadre de la dreta i descartarem els 15 nens adoptats en països no determinats

Continent	Nº nens adoptats	%
Europa		
Àsia		
Amèrica del sud		
Àfrica		

Taules de freqüències per variables quantitatives contínues

Quan treballem amb variables contínues ens podem trobar amb una gran varietat de resultats que podem ajuntar en grups sense que les conclusions de l'estudi quedin alterades. Aquests grups de nombres s'anomenen intervals.

Interval

Es tracta de escriure de forma simbòlica tots els nombres que hi ha entre altres determinats. Per exemple, si volem escriure tots els nombres entre 1 i 10 veurem que és impossible:

$$\begin{array}{c} 1 \\ 1,0000 \dots 01 \\ 1,0000 \dots 02 \\ \dots \\ 9,9999 \dots 99 \\ 10 \end{array}$$

de fet, ni tant sols podem escriure el segon nombre ja que té infinites xifres decimals. Per això si volem escriure tots els nombres entre 1 i 10 posem

$$1, 10$$

però, inclòs l'1? I el 10?. El signe [o] al costat del nombre implica que si està inclòs mentre que el signe (o) implica que està exclòs. D'aquesta forma

$$\begin{array}{l} [1, 10] \quad \rightarrow \quad 1 \\ \quad \quad \quad \quad 1,0000 \dots 01 \\ \quad \quad \quad \quad \dots \\ \quad \quad \quad \quad 9,9999 \dots 99 \\ \quad \quad \quad \quad 10 \end{array}$$
$$\begin{array}{l} (1, 10] \quad \rightarrow \quad 1,0000 \dots 01 \\ \quad \quad \quad \quad \dots \\ \quad \quad \quad \quad 9,9999 \dots 99 \\ \quad \quad \quad \quad 10 \end{array}$$
$$\begin{array}{l} (1, 10) \quad \rightarrow \quad 1,0000 \dots 01 \\ \quad \quad \quad \quad \dots \\ \quad \quad \quad \quad 9,9999 \dots 99 \end{array}$$

Els nombres que delimiten l'interval, en aquest cas 1 i 10, són els extrems i l'amplada de l'interval és la diferència entre els extrems, al nostre exemple és 9 (10-1)

EXEMPLE:

Estem fent un estudi en relació al temps que necessiten els alumnes del nostre centre en desplaçar-se a peu fins a l'institut des de casa seva. Per això demanem que, durant una setmana, mesurin el temps amb ajut d'un rellotge i ens donin el promig dels resultats obtinguts en minuts amb una xifra decimal. Així un company ens dona la dada 2,5 minuts que resulta ser la més petita, mentre altre ens dona com a temps promig 18,2 minuts que és la dada més gran. Podem preveure que la quantitat de resultats diferents que tindrem serà enorme i que el que més ens convé es establir intervals:

De 0 a 5 minuts justos	→	[0 , 5]
de 5 a 10 minuts justos	→	(5 , 10]
de 10 a 15 minuts justos	→	(10 , 15]
de 15 a 20 minuts	→	(15 , 20)

que ens faran més fàcil la recollida de dades.

- Per ser rigorosos el nombre d'intervals s'estableix amb la fórmula:

$$N^{\circ} \text{ d'intervals} = \sqrt{\text{nombre de dades}}$$

Taules de freqüències

En el cas de fer estudis amb dades agrupades en intervals cal trobar la **marca de classe** que serà el nombre a partir del qual farem els càlculs.

$$\text{Marca de classe} = \frac{\text{suma d'extrems de l'interval}}{2}$$

EXEMPLE:

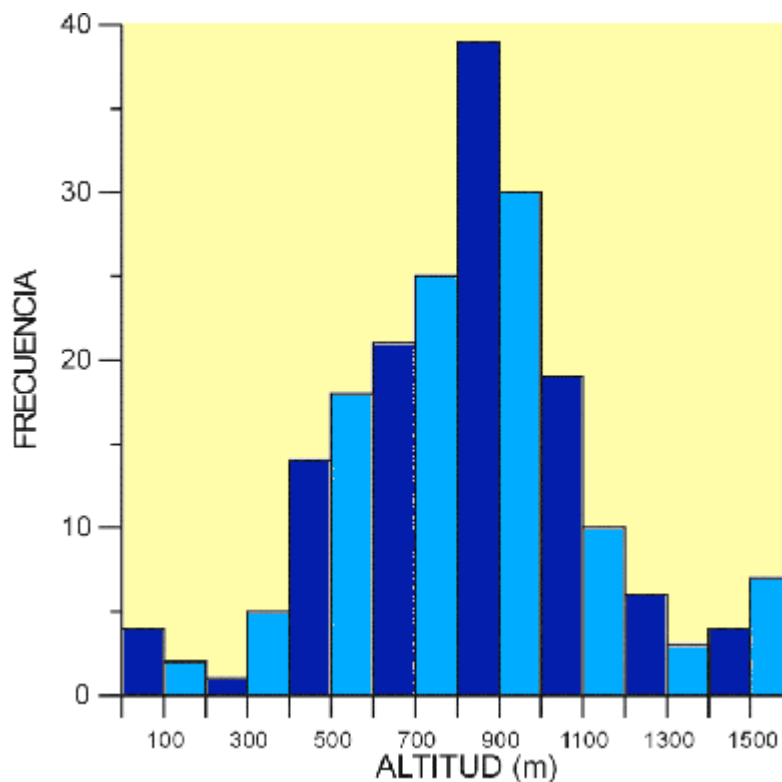
x_i	marca de classe	n_i	N_i	$f_i = \frac{n_i}{N}$	F_i	$\% = f_i \cdot 100$
[0,5]	2,5	4	4	0,20	0,20	20
(5,10]	7,5	8	12	0,40	0,60	40
(10,15]	12,5	5	17	0,25	0,85	25
(15,20]	17,5	3	20	0,15	1,00	15
Suma total				1,00		100

18. Donat l'interval $[3,6)$,
- Quins són els extrems?
 - Quina és l'amplitud de l'interval?
 - Està inclòs el 3? I el 6?. Justifica la resposta
 - Anomena dos nombres inclosos en l'interval i dos que no ho estiguin
19. Expressen amb un interval tots els nombres entre -2 i 5, inclosos ambdós.

Gràfics estadístics per variables quantitatives contínues

La representació gràfica d'aquestes variables és l' **histograma** on l'àrea del rectangle és proporcional a la freqüència de la dada.

Ex: Nombre d'escarbats *Pseudolucanus barbarossa* (Fabricius, 1801) en la Península Ibèrica segons l'altitud (metres)



20. El número de les sabates esportives que utilitzen a la classe d'educació física els 25 alumnes és:

37, 40, 39, 38, 40, 40, 38, 41, 41, 39, 39, 42, 40,
40, 41, 42, 42, 43, 38, 43, 37, 41, 39, 41, 40.

- Indica quina variable s'estudia i de quin tipus és.
- Quin és el nombre d'individus de la mostra?
- Fes una taula estadística.
- Representa les dades en un diagrama de barres

21. Una enquesta sobre la despesa diària en transport ha donat:

Despesa en Euros	Nombre de persones
De 0,50 a 0,65	24
De 0,65 a 0,80	43
De 0,80 a 0,95	22
De 0,95 a 1,10	10
De 1,10 a 1,25	11
De 1,25 a 1,40	12

- Troba les marques de classe.
- Fes una taula estadística.
- Quin percentatge gasta més de 1,1 € ?

22. Donada la sèrie estadística:

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

- Fes una taula estadística agrupant les dades en 3 intervals : $[0,3)$, $[3,6)$, $[6,9)$.
- Representeu les dades mitjançant un histograma.

23. S'ha cronometrat la durada de 25 pel·lícules. Els resultats en minuts són:

89, 103, 92, 117, 95, 100, 109, 123, 110, 88, 99, 95,
95, 98, 102, 98, 114, 117, 120, 95, 91, 92, 105, 119, 127.

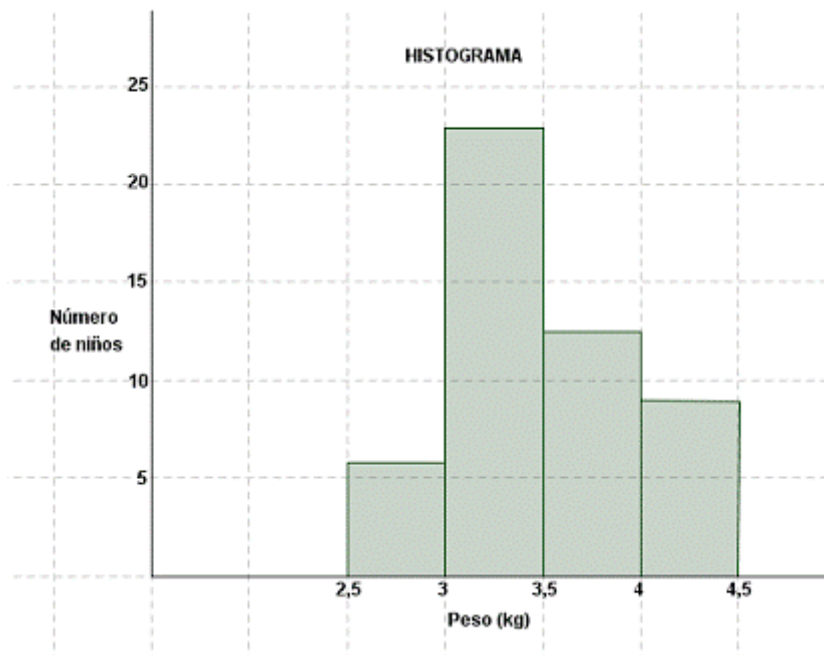
- Ordena les dades en ordre ascendent.
- Reparteix les dades en 5 intervals.
- Calcula el representat de cada interval (marca de classe).
- Fes la taula estadística
- Dibuixa d'histograma.

24. El nombre de CDs de música que tenen un grup de joves és:

Nombre de CDs	Nombre de joves
0 a 12	2
13 a 25	5
26 a 38	10
39 a 51	15
52 a 64	17
65 a 77	9
78 a 90	6
91 a 103	1

- Fes la taula estadística
- Dibuixa d'histograma.
- Mirant la taula i sense calcular res contesta:
Quants joves tenen 38 CDs o menys?
Quants joves tenen 64 CDs o menys?
Quants joves tenen 85 CDs o menys?

25. Donat el gràfic:



- Quina és la variable contínua estudiada.
- Fes la taula estadística amb els intervals corresponents

Paràmetres estadístics

Després de recollir les dades i expressar els resultats obtinguts en taules o gràfiques resulta interessant calcular una sèrie de paràmetres o de nombres que ens expliquen més coses de la característica estudiada i que ens deixa comparar els resultats de dos estudis semblants.

Aquests nombres característics o paràmetres poden ser:

- De centralització. Ens indiquen al voltant de quin valor central es distribueixen les dades.
- De dispersió. Ens diuen com estan de concentrats o dispersos al voltant d'un valor central les dades obtingudes.
- De posició. Divideixen el grup de dades en grups amb el mateix nombre de dades.

Estudiarem els **paràmetres de centralització**:

- a) Mitjana aritmètica (\bar{x}). És el valor promig de les dades.

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \dots + x_n \cdot f_n}{N}$$

- b) Mediana (Me). És el valor que ocupa la posició central quan ordenem les dades quantitatives de menor a major.
- c) Moda (Mo). És el valor que té una major freqüència

EXEMPLE:

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

x_i	n_i	$f_i = \frac{n_i}{N}$	$x_i \cdot f_i$
0	4	$0,20 = 4/20$	$0 = 0 \cdot 4$
1	8	$0,40 = 8/20$	$8 = 1 \cdot 8$
2	6	$0,30 = 6/20$	$12 = 2 \cdot 6$
3	2	$0,10 = 2/20$	$6 = 3 \cdot 2$
Suma total	N = 20	1,00	26

$$\bar{x} = \frac{26}{20} = 1,3 \quad \text{Me} = 1 \quad \text{Mo} = 1$$

- En el cas de treballar amb intervals utilitzarem com x_i la marca de classe

26. Calcula la mitjana aritmètica, la mediana i el mode amb les dades:

a) 3, 9, 14, 14, 22, 31, 33, 60.

b) 8, 11, 4, 3, 2, 5, 10, 6, 4, 1, 10, 8, 12, 6, 5, 7.

c)

53	51	58	54	59	60	52	52
55	49	51	53	55	60	58	51
57	56	50	53	58	59	59	57

27. Calcula la mitjana aritmètica de la següent sèrie de dades:

0, 2, 1, 2, 3, 1, 4, 3, 2, 2, 0, 1, 2, 3, 0, 2, 0, 2, 3, 2, 3, 4, 2, 6, 2

28. Calcula la mitjana aritmètica de la sèrie següent:

17 °C, 18 °C, 21 °C, 28 °C, 17 °C, 18 °C, 20 °C, 22 °C, 20 °C

29. Les 10 alumnes d'una classe calcen el següent nombre de sabata:

35, 38, 39, 36, 38, 37, 39, 35, 37, 36

Calcula el nombre mitjà de sabata que calcen les alumnes d'aquesta classe.

30. Entre un grup d'amics hi ha: 5 que tenen 15 anys, 9 que en tenen 16, 7 que en tenen 17, 2 que en tenen 18 i 2 que en tenen 19.

Calcula la mitjana aritmètica de les edats.

31. Hem preguntat l'edat a un grup de persones que estaven esperant el bus i les respostes són:

14, 29, 35, 67, 69, 52, 47, 12, 36, 47, 56, 81, 60,
19, 18, 21, 34, 55, 64, 47, 15, 47, 9, 47, 32

a) calcula la mitjana aritmètica corresponent a l'edat.

b) calcula la moda.

c) calcula la mediana.

32. Els alumnes d'una classe de 4t d'ESO diuen que durant l'estiu han vist, entre TV i cinema el següent nombre de pel·lícules:

12, 19, 10, 25, 36, 42, 16, 31, 52, 11, 14, 37, 25, 46, 24, 52, 13, 14, 9, 12.

- a) calcula la mitjana aritmètica del nombre de pel·lícules vistes.
- b) calcula la moda
- c) calcula la mediana.

33. Hem preguntat a un grup de veïns quina era la superfície del seu pis en metres quadrats. Les respostes són:

70, 79, 80, 90, 93, 96, 110, 115, 62, 69, 82, 87, 88,
89, 97, 97, 100, 102, 104, 90, 89, 75, 74, 68, 66.

- a) calcula la mitjana aritmètica de la superfície del pis.
- b) calcula la moda
- c) calcula la mediana.

34. A continuació tens el nombre de llibres que tenen en diferents famílies del barri:

Nombre de llibres	Nombre de famílies
0 a 12	2
13 a 25	5
26 a 38	10
39 a 51	15
52 a 64	17
65 a 77	9
78 a 90	6
91 a 103	1

Calcula la mitjana aritmètica del nombre de llibres que hi ha a cada família.