

# Estadística

Àmbit de les Matemàtiques, de la Ciència  
i de la Tecnologia

G<sub>2</sub> R<sub>1</sub> A<sub>1</sub> D<sub>2</sub> U<sub>1</sub> ï<sub>1</sub> ' S<sub>1</sub>

ARA EN SECUNDÀRIA



# Estadística

**Àmbit de les Matemàtiques, de la Ciència  
i de la Tecnologia**



Col·lecció **GRADUÏ'S. ARA EN SECUNDÀRIA**

**Mòdul opcional 19**  
**Estadística**

Àmbit de les Matemàtiques, de la Ciència i de la Tecnologia

© Generalitat de Catalunya  
Departament d'Educació i Universitats

Edició [Servei de Difusió i Publicacions](#)

Coordinació [Subdirecció General de Formació de Persones Adultes](#)

Equip de redacció [Annabel Fernández](#)  
[Fermí Porta](#)

Il·lustració [Jordi Alba](#)

Muntatge [TIMIT 2003](#)

Recerca i cessió dels drets d'imatge [Ciència Activa](#)

Impressió [Gràfiques Cuscó, SA](#)

Es prohibeix la reproducció total o parcial d'aquesta obra, per qualsevol mitjà o procediment, compresos la reprografia i el tractament informàtic, sense autorització prèvia i per escrit del titular del «Copyright».

ISBN: ISBN-84-393-7233-7  
2a edició: setembre de 2006  
Tiratge: 1.000 exemplars

Dipòsit legal: B-46839-2006

## SUMARI

<b>ORGANITZACIÓ DELS MÒDULS I LES UNITATS</b>	5
<b>INTRODUCCIÓ</b>	7
<b>PUNT DE PARTIDA</b>	9
<b>UNITAT 1    RECOLLIM DADES</b>	11
QUÈ TREBALLARÀS?	12
CONTINGUTS	13
ACTIVITATS D'APRENTATGE	23
ACTIVITATS COMPLEMENTÀRIES	27
ACTIVITATS D'AVUACIÓ	28
SOLUCIONS DE LES ACTIVITATS D'APRENTATGE	30
SOLUCIONS DE LES ACTIVITATS D'AVUACIÓ	35
QUÈ HAS TREBALLAT?	37
COM HO PORTO?	38
<b>UNITAT 2    LA REPRESENTACIÓ GRÀFICA DE LES DADES</b>	39
QUÈ TREBALLARÀS?	40
CONTINGUTS	41
ACTIVITATS D'APRENTATGE	52
ACTIVITATS COMPLEMENTÀRIES	66
ACTIVITATS D'AVUACIÓ	68
SOLUCIONS DE LES ACTIVITATS D'APRENTATGE	71
SOLUCIONS DE LES ACTIVITATS D'AVUACIÓ	86
QUÈ HAS TREBALLAT?	89
COM HO PORTO?	90
<b>UNITAT 3    TERMES MITJANS</b>	91
QUÈ TREBALLARÀS?	92
CONTINGUTS	93
ACTIVITATS D'APRENTATGE	105
ACTIVITATS COMPLEMENTÀRIES	116
ACTIVITATS D'AVUACIÓ	117
SOLUCIONS DE LES ACTIVITATS D'APRENTATGE	119
SOLUCIONS DE LES ACTIVITATS D'AVUACIÓ	133
QUÈ HAS TREBALLAT?	136
COM HO PORTO?	137
<b>PUNT D'ARRIBADA. ACTIVITATS D'AVUACIÓ DEL MÒDUL</b>	138
<b>SOLUCIONS DE LES ACTIVITATS D'AVUACIÓ DEL MÒDUL</b>	142

A l'inici del mòdul trobaràs sempre dos apartats:

## Introducció del mòdul:

És la presentació del mòdul i s'hi explica en quin nivell es troba, si és comú o opcional i en quines unitats es divideix. En aquest apartat et suggerim quins passos pots seguir per treballar les unitats del mòdul.

## Punt de partida:

Et presenta els continguts del mòdul organitzats. Et donarà una primera visió del que treballaràs abans de començar les unitats.

Cada unitat didàctica està estructurada en:

## Què treballaràs?:

Presenta les idees principals que veuràs a la unitat i els objectius que es pretenen i que al final hauràs d'haver assolit.

## Bloc de continguts

Presenta el desenvolupament dels blocs temàtics de cada unitat amb alguns suggeriments que et faciliten el seu estudi.

## Bloc d'activitats

**ACTIVITATS D'APRENTATGE:** Contenen activitats per tal que vagis construint i consolidant els aprenentatges així com assolint els objectius que es pretenen.

**ACTIVITATS COMPLEMENTÀRIES:** Són activitats que complementen les activitats d'aprenentatge amb l'objectiu d'ampliar i reforçar l'aprenentatge dels continguts i objectius del mòdul. No tenen bloc de solució perquè són propostes de treball que poden servir com activitats per a l'avaluació formativa del mòdul: parla'n amb el teu tutor.

**ACTIVITATS D'AVALUACIÓ:** Pretenen informar-te del nivell d'assoliment dels objectius proposats al començament en l'apartat "Què treballaràs?". Aquesta informació et servirà per identificar els continguts fonamentals i els objectius que no has assolit i que hauries de reforçar.

## Bloc de solucions

**SOLUCIONS DE LES ACTIVITATS D'APRENTATGE:** Són les respostes de les activitats d'aprenentatge. No les hauries de consultar fins haver fet totes les activitats d'aprenentatge proposades.

**SOLUCIONS DE LES ACTIVITATS D'AVALUACIÓ:** Són les respostes de les activitats d'avaluació. No les hauries de consultar fins haver fet totes les activitats tant d'aprenentatge com d'avaluació.

## 6

**Què has treballat?:**

És una proposta d'esquema o de mapa conceptual que relaciona o resumeix els continguts treballats en la unitat. És una eina per facilitar-te la comprensió, la síntesi i el repàs dels continguts de la unitat.

**Com ho porto?:**

Presenta un quadre d'autoavaluació per comprovar si has assolit els objectius proposats a l'inici de la unitat o quins són els que hauries de tornar a repassar.

Al final del mòdul trobaràs un últim apartat:

**Punt d'arribada:**

Facilita l'autoavaluació dels continguts fonamentals treballats en el mòdul. Conté:

**ACTIVITATS D'AVALUACIÓ DEL MÒDUL:** Inclou les activitats que permeten autoavaluar els continguts fonamentals del mòdul.

**SOLUCIONS DE LES ACTIVITATS D'AVALUACIÓ DEL MÒDUL:** Són les respostes a les activitats d'avaluació del mòdul.

El mòdul Estadística és un mòdul opcional per treballar en el segon o tercer nivell de l'àmbit de les matemàtiques, de la ciència i de la tecnologia del Graduat en Educació Secundària per a adults. Abans de cursar-lo, és convenient tenir assolits els objectius dels mòduls Economia domèstica i Taules i gràfics.

MÒDULS OPCIONALS	
11. Geometria plana	Nivell 1 o 2
12. Taules i gràfics	
13. Models matemàtics de fenòmens	
14. Operacions numèriques	
15. Aliments i nutrició	
16. Planeta Terra	
17. Ecosistemes	
18. El món del treball	
<b>19. Estadística</b>	Nivell 2 o 3
20. Creixement i decreixement	
21. Organismes i entorn	
22. Força i moviment	
23. El canvi de la matèria	
24. Instruments i tècniques	

Aquest mòdul està format per tres unitats didàctiques:

#### Unitat 1 Recollim dades

Estudiaràs com s'ordenen les dades amb les quals volem fer un estudi estadístic, per tal de facilitar els càlculs numèrics.

#### Unitat 2 La representació gràfica de les dades

Treballaràs com es construeixen i s'interpreten els principals diagrames en els estudis estadístics.

#### Unitat 3 Termes mitjans

Coneixeràs com es calculen i s'interpreten els valors numèrics que permeten resumir el comportament d'un conjunt de dades estadístiques.

En aquest mòdul aprendràs les principals aplicacions de l'estadística en el món de la ciència. Intentarem fer aquesta aproximació d'una manera senzilla i directa.

Ens centrarem en conèixer com l'**estadística descriptiva** ens permet expressar i interpretar la informació continguda en grans quantitats de dades. Aplicada a les ciències socials i a les ciències biològiques, ens permet entendre els mecanismes que defineixen fets històrics, conductes, relacions ecològiques, etc. Les eines que utilitza l'estadística ens permetran descriure, organitzar i presentar les dades d'una manera més clara.

En aquest mòdul també et familiaritzaràs amb la utilització de gràfics. En els mitjans de comunicació, tot sovint trobem diagrames de barres, diagrames de sectors, pictogrames i polígons de freqüències, per descriure determinats fets. El coneixement de l'estadística et permetrà interpretar i construir tots aquests tipus de gràfics.

Esperem que, un cop finalitzat el mòdul, diguis: «malgrat tot, l'estadística no és pas tan avorrida!».

L'assoliment dels objectius d'aquest mòdul facilitaran l'estudi del mòdul de Genètica.



Posa la data d'avui en la primera columna.  
Per respondre les preguntes posa un número de l'1 al 3 en funció del que en sàpigues.

1. No en sé res.
2. En sé alguna cosa.
3. Ho sé bé.

Quan acabis d'estudiar el mòdul emplena la segona columna. Així podràs veure què has après.

	Data:	Data:
Saps com s'organitzen les dades recollides per tal de fer un estudi estadístic?		
Saps dibuixar diagrames estadístics a partir d'una col·lecció de dades?		
Saps interpretar els diagrames estadístics que apareixen en els mitjans de comunicació?		
Saps com es calcula la probabilitat d'un determinat esdeveniment?		
Coneixes els paràmetres numèrics que s'utilitzen per resumir la informació continguda en una col·lecció de dades?		
Saps de quina manera tendeixen a distribuir-se, en general, les dades numèriques recollides en un estudi estadístic?		



# Unitat 1

## RECOLLIM DADES

11

UNITAT 1 RECOLLIM DADES

19. ESTADÍSTICA

Matemàtiques, Ciència i Tecnologia

# què treballaràs?

En aquesta unitat aprendràs a organitzar les dades recollides per fer un estudi estadístic i coneixeràs les relacions entre estadística i probabilitat.

En acabar aquesta unitat has de ser capaç de:

- Explicar i reconèixer els conceptes estadístics de població, mostra i variable.
- Distingir entre les variables qualitatives i les variables quantitatives.
- Reconèixer les variables quantitatives discretes i les variables quantitatives contínues.
- Precisar els diferents tipus de freqüències estadístiques que poden ésser estudiades a partir d'un recull de dades.
- Agrupar les dades en taules de freqüències.
- Calcular la freqüència absoluta, la freqüència relativa, la freqüència absoluta acumulada i la freqüència relativa acumulada.
- Distingir entre probabilitat teòrica i probabilitat experimental.
- Calcular la probabilitat experimental a partir de la població sencera i a partir d'una mostra de la població.

## 1. Població, mostra i variable

Abans d'endinsar-nos en el tractament estadístic de dades, és necessari que ens familiaritzem amb alguns termes estadístics.

### Població

Una població? Com Terrassa, Manresa o Berga? Sí, això són poblacions en la nostra manera de parlar habitual, però, a l'hora de fer estudis estadístics, ens caldrà redefinir aquest terme.

En estadística, anomenem **població** la totalitat de persones, animals o coses que són objecte d'un estudi estadístic.

Alguns exemples de poblacions són:

- Els estudiants que hi ha en una classe
- Els nens d'entre cinc i deu anys que viuen a Manresa
- Els ordinadors d'una escola
- Els arbres d'una plaça de la ciutat
- Els gats de la ciutat

### Mostra

Moltes vegades no és factible estudiar tota la població sencera perquè, atesa la seva grandària, comportaria massa despesa i hauríem de dedicar-hi molt de temps. Aquest és el cas de les enquestes sobre la intenció de vot. Entrevistar tothom seria molt costós i, per aquest motiu, es treballa amb una mostra.

Una **mostra** és una part de la població.

Així doncs, els investigadors seleccionen sovint un subgrup de la població, anomenat mostra, per tal de fer els seus estudis. Una mostra mai no és perfecta, ja que es tracta d'una part de la població. Malgrat tot, amb una bona mostra, agafada a l'atzar però de forma representativa, disposarem d'una bona aproximació a la realitat de la població sencera.

Què significa que una mostra sigui representativa? Continuem amb l'exemple de les enquestes sobre la intenció de vot. A l'hora de triar la mostra, caldrà que el percentatge d'homes i dones en la mostra sigui el mateix que el que hi ha a la població sencera. Serà necessari considerar el mateix pel que fa al nivell d'estudis, la situació econòmica, etc. D'aquesta manera la mostra s'ajustarà millor a la realitat de la població sencera i els resultats de l'enquesta seran molt més fiables.

La mida de la mostra dependrà de la mida de la població i del tipus de variable que estudiem.

### Variable

Anomenem **variable** allò que estudiem en la població.

Les variables poden ésser qualitatives o quantitatives.

## 14 Variables qualitatives

Són les que ens proporcionen estudis sobre qualitats. Els valors d'aquest tipus de variables són noms o nivells. Fem una ullada a alguns exemples:

Variable	Valors
Professió del pare	Electricista, professor, secretari, escriptor, arquitecte, polític, etc.
Estat civil	Solter/a, casat/ada, vidu/dua, divorciat/ada, separat/ada
Color dels ulls	Blau, negre, verd, gris, marró
Tipus de matalàs	D'aire, d'aigua, de plomes, futon, de làtex, etc.
Llengua materna	Anglès, català, espanyol, àrab, francès, etc.

## Variables quantitatives

Són les que proporcionen estudis numèrics, de quantitats. N'hi ha de dos tipus: discretes i contínues.

Les **variables discretes** són les que tenen valors concrets i periòdics, sense valors intermedis. Vegem-ne alguns exemples:

Variable	Valors
Treballadors d'una fàbrica	1, 2, 3, 4, 5, 6,...
Nombre de germanes	0, 1, 2, 3, 4, 5, 6,...
Diaris venuts al quiosc	0, 1, 2, 3, 4, 5, 6,...
Gent que ha assistit al cinema	0, 1, 2, 3, 4, 5, 6,...
Nombre de visitants al museu Dalí, avui	0, 1, 2, 3, 4, 5, 6,...

Si considerem el primer exemple, és a dir, els treballadors d'una fàbrica, fixa't que no hi ha valors intermedis. No hi poden haver 3,4 treballadors.

Les **variables contínues** són les que tenen valors sense discontinuïtats i, per tant, infinits. Vegem-ne alguns exemples:

Variable	Valors
Pressió sanguínia	Infinits
Quantitat de nicotina a les cigarretes	Infinits
Alçada dels ximpanzés	Infinits
Pes dels nounats	Infinits
Consum d'aigua per persona	Infinits

Entens per què els valors són infinits? Si considerem l'alçada dels ximpanzés, l'alçada d'un ximpanzé podria ésser 123,2 cm i la d'un altre podria ésser 123,3 cm. Malgrat la poca diferència de mida entre els dos ximpanzés, podríem trobar ximpanzés d'una alçada entre la d'aquests dos: 123,221 cm, 123,2475 cm, etc.

Tanmateix, en alguns estudis, les variables contínues es converteixen en discretes. Com? Si considerem el mateix exemple de l'alçada dels ximpanzés, podem redefinir la variable i considerar l'alçada dels ximpanzés en centímetres sense decimals, aproximant la part decimal al valor més proper sense decimals. Així, per exemple, faríem:

Alçada real	Alçada arrodonida sense decimals
119,2 cm	119 cm
121,6 cm	122 cm
116,4 cm	116 cm

- Fes les activitats d'aprenentatge 1, 2 i 3

## 2. Una eina molt útil: les taules de freqüències

Quan hem de manejar un gran nombre de dades amb les quals volem fer un estudi estadístic, és molt útil ordenar-les per tal de facilitar els càlculs numèrics i disminuir la possibilitat de cometre errades. Amb aquesta finalitat, tot seguit aprendràs a construir taules de freqüències, les quals poden ésser de diferents tipus segons la manera de representar les dades.

A partir d'ara haurem de realitzar moltes operacions amb nombres decimals. És per aquest motiu que cal establir uns criteris a l'hora d'expressar aquestes quantitats decimals:

- En general, deixarem dues xifres decimals significatives. Això vol dir que cal posar dues xifres decimals a partir de la primera que no sigui un zero.  
Exemples: 0,58 0,034 0,0032
- Abans d'escriure el segon decimal significatiu cal fixar-se en el tercer. Si el tercer és 5 o més de 5, apujarem un dígit el segon decimal. Si és 4 o menys de 4, deixarem el segon decimal tal com està. D'aquesta manera minimitzarem l'error produït per escriure només dos decimals significatius.

Vegem uns exemples:

El nombre 0,8245 el deixarem així: 0,82

El nombre 0,03271 el deixarem així: 0,033

El nombre 0,01924 el deixarem així: 0,019

El nombre 0,2963 el deixarem així: 0,3

El nombre 0,5555 el deixarem així: 0,56

- Si un nombre decimal només té tres xifres decimals significatives, les posarem totes tres. Així doncs, si dividim 1 entre 8 o 1 entre 16:

$1 : 8 = 0,125$

Deixarem les tres xifres i no posarem 0,13

$1 : 16 = 0,0625$

Deixarem les tres xifres i no posarem 0,063

## 16 Freqüència absoluta

Una taula de freqüències bàsica mostra els valors de la variable que estudiem i el nombre de vegades que succeeix cada valor, és a dir, la seva **freqüència absoluta** ( $n_i$ ). És una manera d'organitzar les dades, fent un llistat de tots els valors possibles de la variable en una columna i la freqüència en què es produeixen en una altra columna.

Per tal d'entendre millor com fer taules, considerarem un exemple.

S'analitzen 20 peixeres i s'estudia quantes espècies d'algues unicel·lulars conté cada una. Els resultats obtinguts són els següents:

3	5	2	3	6	3	5	2	6	1
1	2	7	5	5	3	6	6	3	0

Per construir una taula de freqüències, primer cal fer una columna a l'esquerra amb els valors de la variable enregistrats (nombre d'espècies d'algues unicel·lulars), des del més petit fins al més gran. Després farem una altra columna a la dreta, on escriurem la freqüència amb què es produeix cada un dels valors.

Nombre d'espècies d'algues unicel·lulars	$n_i$
0	1
1	2
2	3
3	5
4	0
5	4
6	4
7	1
	$N = 20$

Fixa't que hem escrit tots els valors de la variable, des del més petit fins al més gran, fins i tot els que no succeeixen.

La suma de la columna de les freqüències absolutes ha d'ésser igual al nombre de dades recollides ( $N$ ), en el cas de l'exemple, el nombre de peixeres analitzades.

## Freqüència relativa

Malgrat que la freqüència absoluta ens dona una bona informació sobre la població estudiada, el valor de la freqüència absoluta no ofereix molta informació tot sol. Imagina't que, després d'haver analitzat un determinat nombre de peixeres, dic que hi ha 5 peixeres que contenen 3 espècies d'algues unicel·lulars. Això són moltes peixeres si el total de peixeres analitzades és de 20, però no en són gaires si el total de peixeres analitzades és de 1.800. És per això que ens cal un altre paràmetre: la freqüència relativa.

La **freqüència relativa** ( $f_i$ ) d'un valor de la variable s'obté dividint la freqüència absoluta del valor de la variable entre el total de la població  $N$ , en el nostre exemple el nombre de peixeres analitzades.

Nombre d'espècies d'algues unicel·lulars	$n_i$	$f_i$
0	1	$1 : 20 = 0,05$
1	2	$2 : 20 = 0,1$
2	3	$3 : 20 = 0,15$
3	5	$5 : 20 = 0,25$
4	0	$0 : 20 = 0$
5	4	$4 : 20 = 0,2$
6	4	$4 : 20 = 0,2$
7	1	$1 : 20 = 0,05$
	$N = 20$	Total $f_i = 1$

Fixa't que la suma de les freqüències relatives ha d'ésser sempre 1. Tanmateix, si les freqüències relatives d'alguns valors de la variable no són exactes, la suma de totes les freqüències relatives pot ésser lleugerament superior o inferior a 1.

Ara, si dic que la freqüència relativa del valor 3 espècies d'algues unicel·lulars és 0,25, queda molt clara la proporció del total de peixeres analitzades que contenen 3 espècies, sense que ens calgui saber si el total de peixeres analitzades és 20, 800 o 20.000.

Si multipliquem per 100 la freqüència relativa d'un valor de la variable, n'obtenim el seu percentatge, que és la forma popular d'expressar la freqüència relativa:

Nombre d'espècies d'algues unicel·lulars	$n_i$	$f_i$	Percentatge
0	1	0,05	5 %
1	2	0,1	10 %
2	3	0,15	15 %
3	5	0,25	25 %
4	0	0	0 %
5	4	0,2	20 %
6	4	0,2	20 %
7	1	0,05	5 %
	$N = 20$	Total $f_i = 1$	Total % = 100

Fixa't que la suma dels percentatges de tots els valors de la variable ha d'ésser 100. Malgrat que, en general, no utilitzarem els percentatges en l'elaboració de les taules, sí que els utilitzarem de vegades a l'hora de fer diagrames. L'operació de calcular els percentatges a partir de les freqüències relatives la farem directament quan elaborem el diagrama on vulguem que apareguin. Només caldrà multiplicar per cent les freqüències relatives o, el que és el mateix, passar la coma dos llocs cap a la dreta.

## 18 Freqüència acumulada

La freqüència acumulada expressa la freqüència d'un valor determinat de la variable sumada a la de tots els valors inferiors. Podem calcular la freqüència absoluta acumulada i la freqüència relativa acumulada.

La **freqüència absoluta acumulada** ( $N_i$ ) d'un valor de la variable és el nombre de vegades que succeeix el valor en consideració o un d'inferior. Es calcula sumant les freqüències absolutes dels valors igual o inferiors al que estem considerant. També es pot calcular sumant a la freqüència absoluta del valor en consideració, la freqüència absoluta acumulada del valor immediatament inferior.

Podem calcular la **freqüència relativa acumulada** ( $F_i$ ) seguint els mateixos procediments, però amb les freqüències relatives. Però hi ha una tercera manera: dividir la freqüència absoluta acumulada entre el total de dades ( $N$ ).

Així doncs, la taula que inclou totes les freqüències estudiades és:

Nombre d'espècies d'algues unicel·lulars	$n_i$	$N_i$	$f_i$	$F_i$
0	1	1	0,05	0,05
1	2	3	0,1	0,15
2	3	6	0,15	0,3
3	5	11	0,25	0,55
4	0	11	0	0,55
5	4	15	0,2	0,75
6	4	19	0,2	0,95
7	1	20	0,05	1
		$N = 20$		Total $f_i = 1$

Si agafem, per exemple, el valor de la variable 3 espècies, podem comprovar com les tres maneres d'obtenir la freqüència relativa acumulada són equivalents:

1. Sumar les freqüències relatives dels valors iguals o inferiors al valor de la variable en consideració:

$$F_i = 0,05 + 0,1 + 0,15 + 0,25 = 0,55$$

2. Sumar la freqüència relativa del valor en consideració amb la freqüència relativa acumulada del valor immediatament inferior:

$$F_i = 0,25 + 0,3 = 0,55$$

3. Dividir la freqüència absoluta acumulada del valor en consideració entre el total de dades ( $N$ ):

$$F_i = \frac{11}{20} = 0,55$$

Nosaltres utilitzarem aquest mètode en la resolució de les activitats.



Les freqüències acumulades només es poden obtenir si es treballa amb variables quantitatives, atès que en les qualitatives no té sentit establir-hi una ordenació de menor a major en els valors de la variable.

La utilització de les freqüències acumulades és molt habitual en els estudis temporals com, per exemple, l'increment de l'IPC (Índex de Preus al Consum) durant un any.

Les taules de freqüències ens seran molt útils per dibuixar gràfiques i calcular els termes mitjans. No sempre ens caldran totes les freqüències, per la qual cosa només en calcularem les que ens facin falta, segons l'exercici.

- Fes les activitats d'aprenentatge 4, 5, 6 i 7

### 3. La probabilitat

Malgrat que l'estudi a fons de la probabilitat es duu a terme en el mòdul comú de Genètica, ara tractarem alguns aspectes relacionats amb l'estadística que no es desenvolupen prou en l'esmentat mòdul de Genètica.

#### La probabilitat teòrica

Quan llancem una moneda enlaire, hi ha dos resultats possibles, anomenats **esdeveniments**: cara i creu. Els dos tenen la mateixa probabilitat de produir-se. Diem que cara i creu són **equiprobables**. En casos com aquest en què els esdeveniments possibles d'un experiment aleatori tenen la mateixa probabilitat  $p$  de produir-se, aquesta probabilitat es pot calcular a partir de la fórmula:

$$p = \frac{\text{Nombre casos favorables}}{\text{Nombre casos possibles}}$$

En el cas de la moneda, la probabilitat que surti creu és:

$$p(\text{creu}) = \frac{1}{2} = 0,5$$

Si tirem un dau, la probabilitat que surti un sis és:

$$p(\text{sis}) = \frac{1}{6} = 0,17$$

Això és així perquè només una de les sis cares és un sis.

Si al tirar el mateix dau volem saber la probabilitat que surti un nombre parell:

$$p(\text{parell}) = \frac{3}{6} = 0,5$$

Això és així perquè, de les sis cares del dau, tres tenen un nombre parell (2, 4 i 6) i, per tant, són favorables per tal que es produeixi l'esdeveniment.

Fixa't que la probabilitat d'un esdeveniment ha d'ésser sempre un valor entre 0 i 1. Mai més gran que 1, ja que el nombre de casos favorables no pot sobrepassar mai el nombre de casos possibles. Malgrat que la probabilitat també es pot expressar com a fracció o en forma de percentatge, nosaltres l'expressarem sempre com a valor entre 0 i 1. Això farà que la probabilitat s'expressi amb un nombre

decimal, excepte en els casos de probabilitat 0 (esdeveniments impossibles) o de probabilitat 1 (esdeveniments segurs).

## La probabilitat experimental

Hi ha moltes situacions en què la probabilitat teòrica no pot ésser determinada, ja que els esdeveniments possibles no són equiprobables ni se'n pot determinar la seva probabilitat de manera teòrica. Aleshores, l'única manera de conèixer la probabilitat és fent un estudi estadístic per determinar-la. Es tracta de determinar la freqüència relativa ( $f_i$ ) de l'esdeveniment del qual en volem conèixer la probabilitat.

### Càlcul de la probabilitat a partir de la població sencera

Vegem-ho a través de dos exemples.

#### Exemple 1

Es volia conèixer la probabilitat que, entrevistant a l'atzar un dels 660 treballadors d'una empresa, aquest tingués estudis universitaris.

Per fer-ho, es va fer un estudi amb els 660 treballadors de l'empresa, que revelà que 99 treballadors tenien estudis universitaris. Aleshores es va poder calcular la probabilitat que entrevistant-ne un a l'atzar en tingués, calculant-ne la seva freqüència relativa:

$$p(\text{Estudis Universitaris}) = f_i(\text{Estudis Universitaris}) = \frac{n_i}{N} = \frac{99}{660} = 0,15$$

#### Exemple 2

Un estudi sobre un model de cotxe que ha sortit al mercat recentment indica que, de les 5.000 unitats venudes fins ara, 375 han patit una avaria electrònica durant el primer mes de funcionament.

Podem calcular la probabilitat que un comprador d'aquests cotxes hagi patit un problema electrònic durant el primer mes:

$$p(\text{avaría}) = f_i(\text{avaría}) = \frac{n_i}{N} = \frac{375}{5.000} = 0,075$$

En aquests dos exemples, l'estudi agafa la totalitat de la població, però en altres casos això no és possible atesa la grandària de la població. En aquests casos cal treballar amb mostres.

### Càlcul de la probabilitat a partir d'una mostra

Vegem-ho també a través de dos exemples.

#### Exemple 1

Al Canadà es volia saber l'eficiència d'un nou antibiòtic per combatre una malaltia. Atès que no es podia fer un estudi que agafés totes les persones del país, es va fer una prova amb 800 pacients. Els resultats demostraren que només en 44 pacients l'antibiòtic fou ineficaç. La resta de pacients van superar la malaltia

gràcies a l'antibiòtic. Amb aquestes dades es determinà que la probabilitat que un malalt que arribés a un hospital es guarís seguint un tractament amb el nou antibiòtic era:

$$p_o(\text{Guariment}) = f_i(\text{Guariment}) = \frac{n_i}{n} = \frac{756}{800} = 0,945$$

Anomenem  $p_o$  la probabilitat calculada a partir d'una mostra de la població i  $n$  al nombre d'individus que hi ha a la mostra.

Cal entendre, però, que, en aquest cas, la fiabilitat de la probabilitat obtinguda depèn de la mostra que s'hagi agafat. Cal que sigui una **mostra representativa** de tota la població.

D'una banda cal que a la mostra hi siguin representats en la proporció que correspongui, els diferents grups específics que hi ha a la població i que poden comportar-se de manera diferent respecte a la variable. En el cas de l'exemple, caldria tenir en compte l'alimentació, l'edat, etc.

D'altra banda cal que la mostra tingui una mida prou gran. Si en comptes de fer l'estudi amb 800 malalts, s'hagués fet amb 200, el resultat de la probabilitat de guariment hagués estat menys fiable. Si l'estudi s'hagués fet amb 2.000 malalts, la probabilitat de guariment obtinguda hagués estat més fiable. Sempre hi haurà un **marge d'error**. Cal que sigui el més baix possible sense haver de fer estudis amb mostres excessivament grans, que encaririen molt els estudis.

S'han fet diversos estudis per determinar quina ha d'ésser la mida de la mostra perquè els resultats siguin aplicables al conjunt de la població amb un determinat marge d'error. Fruit d'aquests estudis hom disposa de diverses fórmules matemàtiques per calcular quina ha d'ésser la mida de la mostra. Nosaltres només estudiarem una fórmula molt simplificada:

$$n = \frac{1}{e^2} \quad \begin{array}{l} n = \text{mida de la mostra} \\ e = \text{error en tant per u} \end{array}$$

Per poder aplicar aquesta fórmula cal que es compleixin aquestes dues condicions:

- Que la població sigui molt gran.
- Que s'accepti un nivell de confiança de 0,95 (95 %). Això significa que hi ha un 95 % de possibilitats que la probabilitat real (la que en resultaria d'entrevistar tota la població sencera) estigui dins del marge d'error, és a dir, dins de l'interval entre  $p_o - e$  i  $p_o + e$ .

Vegem un exemple de càlcul de la mida d'una mostra per entendre millor el funcionament d'aquesta fórmula.

### Exemple 2

Es vol calcular la mida que ha de tenir una mostra de la població de Catalunya per poder determinar, en una entrevista a l'atzar, la probabilitat que l'entrevistat practiqui un esport, amb un error inferior al 6 % i amb un nivell de confiança de 0,95. També es vol calcular quina hauria d'ésser la mida de la mostra si només s'accepta un error inferior al 4 %.

Apliquem la fórmula per a un error inferior al 6 %:

$$n = \frac{1}{0,06^2} = \frac{1}{0,0036} = 277 \text{ persones}$$

Apliquem ara la fórmula per a un error inferior al 4 %:

$$n = \frac{1}{0,04^2} = \frac{1}{0,0016} = 625 \text{ persones}$$

Fixa't que si volem que l'error sigui més petit, cal que la mostra sigui més gran. En tots dos casos (i sempre que s'utilitzi aquesta fórmula) hi ha un nivell de confiança de 0,95. Això vol dir que hi ha un 95 % de possibilitats que la probabilitat real  $p$ , la que obtindríem entrevistant a tota la població de Catalunya estigui dins de l'interval comprès entre la probabilitat obtinguda a partir de la mostra  $p_0$  i el marge d'error  $e$  acceptat (cap amunt i cap avall). Complementàriament hi ha un 5 % de possibilitats que  $p$  s'allunyi de  $p_0$  més enllà del marge d'error  $e$  acceptat. Sempre que treballem amb mostres cal acceptar la possibilitat que  $p$  s'allunyi de  $p_0$  més enllà del marge d'error  $e$  acceptat. Aquesta probabilitat es pot disminuir augmentant la mida de la mostra, però no es pot reduir a zero. Reduir-la a zero implicaria treballar amb tota la població i no amb una mostra.

Hi ha diferents fórmules per calcular la mida de la mostra, segons les circumstàncies d'estudi: si hi ha estudis previs, si la població és petita, etc. Algunes d'aquestes fórmules són força complexes. Nosaltres només hem intentat fer una primera aproximació al treball amb mostres.

- Fes les activitats d'aprenentatge 8, 9 i 10

**Activitat 1**

Digues sis poblacions que puguin ésser objecte d'un estudi estadístic dins d'una escola i una variable que es pugués estudiar en cada una d'elles.

Població	Variable

**Activitat 2**

Què cal tenir en compte a l'hora de triar una mostra perquè sigui representativa de la població de votants en unes eleccions municipals a Manresa?

---



---



---



---



---

**Activitat 3**

Classifica les variables següents:

Pes del cotxe, pel·lícula favorita, nombre de telèfons a casa, nombre de germans, consum diari d'aigua, superfície de l'habitació, color del pijama dut ahir a la nit, llengua materna, nombre de llibres llegits l'any passat.

Variabls qualitatives	Variabls quantitatives discretes	Variabls quantitatives contínues

**Activitat 4**

La Maria fa un estudi sobre el mitjà de transport utilitzat pels seus companys de feina per anar a treballar. Aquí hi ha els resultats:

Metro	Autobús	Cotxe particular	Bicicleta
Autobús	Moto	Moto	Metro
Bicicleta	A peu	Metro	Tren
A peu	Cotxe particular	Tren	Metro
Metro	Moto	Metro	Autobús
A peu	Autobús	Cotxe particular	Taxi

Fes una taula que inclogui les freqüències absoluta i relativa.

**Activitat 5**

La Marta va cronometrar, durant el novembre passat, el temps que trigava per anar de casa fins a l'estació de tren. Aquests són els resultats enregistrats:

5	4	6	7	8	6	5	4	7	9
4	4	4	7	10	5	8	9	5	6
8	4	8	5	12	10	9	6	5	4

Construeix una taula on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .



**Activitat 8**

Un estudi de mercat fet a Sabadell a partir de 2.000 persones entrevistades revela que només 450 d'aquestes persones estaven disposades a contestar una enquesta sobre la intenció de vot en unes eleccions.

- A. Quina és la probabilitat que a Sabadell una persona agafada a l'atzar no vulgui contestar una enquesta sobre intenció de vot?

- B. La probabilitat obtinguda és teòrica o experimental? És real o estimada?

---

---

---

---

---

**Activitat 9**

Volem fer un estudi estadístic per determinar la probabilitat que una dona catalana d'entre 35 i 37 anys escollida a l'atzar tingui la intenció d'ésser mare biològica abans dels 40 anys.

Quina mida ha de tenir la mostra si s'accepta un error de inferior al 3 % i un nivell de confiança del 95 %?

**Activitat 10**

Si per determinar la probabilitat que en unes eleccions un determinat candidat sigui votat pels electors, amb un nivell de confiança del 95 %, s'han entrevistat 494 persones, quin error s'ha acceptat en aquest estudi?



**Activitat 1**

Agafa dades sobre algun fet quotidià (persones a l'interior de cada cotxe, gent que puja a l'autobús a cada parada, etc.) i construeix una taula de freqüències on apareguin totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

Fes la taula tal com has après a fer-ne. Tingues en compte el tractament dels nombres decimals. Suma totes les freqüències absolutes i totes les freqüències relatives per tal de comprovar que no hi ha errades. Recorda que la suma de les freqüències absolutes ha d'ésser igual al total de dades, i la suma de les freqüències relatives ha d'ésser 1.

--

**Activitat 2**

Calcula la probabilitat estimada d'algun fet quotidià com, per exemple, que en arribar a casa hi hagi l'ascensor a la planta baixa preparat per agafar-lo, acceptant un error de inferior al 15 % i un nivell de confiança del 95 %?

Abans de fer l'estudi has de calcular la mida de la mostra a partir de la fórmula

$$n = \frac{1}{e^2}$$

Un cop recollides les dades, per calcular la probabilitat, hauràs d'aplicar la fórmula

$$p_o = \frac{n_i}{n}$$

--

**Activitat 1**

Classifica les variables següents:

Diàmetre de monedes, programa de TV vist ahir a les 19 hores, cançó més escoltada l'any passat, nombre de títols universitaris, consum de benzina en cent quilòmetres, nombre d'àpats fora de casa, beguda preferida, pes dels bolígrafs, nombre de radiadors a casa.

Variables qualitatives	Variables quantitatives discretes	Variables quantitatives contínues

**Activitat 2**

En Cristian va sembrar un grapat de llavors de julivert i van néixer 25 plantes. Quaranta dies després de sembrar, va comptar les fulles que tenia cada planta i va recollir aquestes dades:

4 7 2 8 6 3 5 8 11 8  
 5 2 8 5 6 5 6 2 3 7  
 9 8 8 7 5

Fes una taula on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

--	--

**Activitat 3**

Un estudi fet a partir de 400 entrevistes a persones usuàries de la línia 1 del metro de Barcelona en un dia laborable de 9 a 10 del matí indica que 120 de les persones entrevistades no resideixen a Barcelona.

D'acord amb aquestes dades, digues:

- A. Quina és la probabilitat que una persona que utilitza el metro en les condicions esmentades no resideixi a Barcelona?

- B. Explica si la probabilitat obtinguda és real o estimada.

---

---

---

- C. Quantes persones s'haurien d'haver entrevistat, acceptant un error inferior al 3,6 % amb un nivell de confiança del 95 %?

**Activitat 1**

Digues sis poblacions que puguin ésser objecte d'un estudi estadístic dins d'una escola i una variable que es pogués estudiar en cada una d'elles.

Respostes possibles:

Població	Variable
alumnes	color d'ulls
guixos	llargada
professors	edat
ordinadors	memòria RAM
taules	alçada
aules	nombre de llums

**Activitat 2**

Què cal tenir en compte a l'hora de triar una mostra perquè sigui representativa de la població de votants en unes eleccions municipals a Manresa?

Cal que la mostra mantingui la proporció que hi ha a la població sencera pel que fa als paràmetres econòmics i socials més importants: habitants de cada barri, proporció de sexes, titulació acadèmica, professió, llengua d'ús habitual, etc.

**Activitat 3**

Classifica les variables següents:

Pes del cotxe, pel·lícula favorita, nombre de telèfons a casa, nombre de germans, consum diari d'aigua, superfície de l'habitació, color del pijama dut ahir a la nit, llengua materna, nombre de llibres llegits l'any passat.

Variabls qualitatives	Variabls quantitatives discretes	Variabls quantitatives contínues
Pel·lícula favorita	Nombre de telèfons a casa	Pes del cotxe
Color del pijama dut ahir a la nit	Nombre de germans	Consum diari d'aigua
Llengua materna	Nombre de llibres llegits l'any passat	Superfície de l'habitació

**Activitat 4**

La Maria fa un estudi sobre el mitjà de transport utilitzat pels seus companys de feina per anar a treballar. Aquí estan els resultats:

La Maria fa un estudi sobre el mitjà de transport utilitzat pels seus companys de feina per anar a treballar. Aquí hi ha els resultats:

Metro	Autobús	Cotxe particular	Bicicleta
Autobús	Moto	Moto	Metro
Bicicleta	A peu	Metro	Tren
A peu	Cotxe particular	Tren	Metro
Metro	Moto	Metro	Autobús
A peu	Autobús	Cotxe particular	Taxi

Fes una taula que inclogui les freqüències absoluta i relativa.

Mitjà de transport	$n_i$	$f_i$
Metro	6	$6 : 24 = 0,25$
Autobús	4	$4 : 24 = 0,17$
Cotxe particular	3	$3 : 24 = 0,125$
Tren	2	$2 : 24 = 0,083$
Moto	3	$3 : 24 = 0,125$
A peu	3	$3 : 24 = 0,125$
Bicicleta	2	$2 : 24 = 0,083$
Taxi	1	$1 : 24 = 0,042$
	$N = 24$	Total $f_i = 1,003$

Atès que les freqüències relatives d'alguns dels valors de la variable no són exactes, la suma de totes les freqüències relatives no dona exactament 1.

**Activitat 5**

La Marta va cronometrar, durant el novembre passat, el temps que trigava per anar de casa fins a l'estació de tren. Aquests són els resultats enregistrats:

5 4 6 7 8 6 5 4 7 9  
 4 4 4 7 10 5 8 9 5 6  
 8 4 8 5 12 10 9 6 5 4

Construeix una taula on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ . Així doncs, la taula que inclou totes les freqüències estudiades és:

Temps de casa a l'estació en minuts	$n_i$	$N_i$	$f_i$	$F_i$
4	7	7	0,23	0,23
5	6	13	0,2	0,43
6	4	17	0,13	0,57
7	3	20	0,1	0,67
8	4	24	0,13	0,8
9	3	27	0,1	0,9
10	2	29	0,067	0,97
11	0	29	0	0,97
12	1	30	0,033	1
	N = 30		Total $f_i = 0,99$	

**Activitat 6**

La Joana fa un estudi sobre el nombre d'hectàrees de vinya conreades pels integrants d'una cooperativa i recull aquestes dades:

9 11 10 8 10 7 9 11 10 8  
 9 11 12 11 12 9 10 10 10 11  
 7 7 10 11 8 10 9 10 12 7

Construeix una taula on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

Nombre d'hectàrees	$n_i$	$N_i$	$f_i$	$F_i$
7	4	4	0,13	0,13
8	3	7	0,1	0,23
9	5	12	0,17	0,4
10	9	21	0,3	0,7
11	6	27	0,2	0,9
12	3	30	0,1	1
	N = 30		Total $f_i = 1$	

**Activitat 7**

En Jaume està fent un estudi sobre el nombre de vianants que creuen cada dia, de 9 a 9:05 h, pel pas de vianants que hi ha enfront de casa. Fins ara ha recollit les dades següents:

4    7    11    12    8    2    6    18    19    23  
1    2    0    13    24    16    17    3    1    6

Construeix una taula on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ . Agrupa el nombre de vianants en classes de 5 valors.

Nombre de vianants	$n_i$	$N_i$	$f_i$	$F_i$
0 a 4	7	7	0,35	0,35
5 a 9	4	11	0,2	0,55
10 a 14	3	14	0,15	0,7
15 a 19	4	18	0,2	0,9
20 a 24	2	20	0,1	1
	N = 20		Total $f_i = 1$	

**Activitat 8**

Un estudi de mercat fet a Sabadell a partir de 2.000 persones entrevistades revela que només 450 d'aquestes persones estaven disposades a contestar una enquesta sobre la intenció de vot en unes eleccions.

- A. Quina és la probabilitat que a Sabadell una persona agafada a l'atzar no vulgui contestar una enquesta sobre la intenció de vot?

La probabilitat que una persona de Sabadell agafada a l'atzar no vulgui contestar una enquesta sobre intenció de vot és:

$$p_o(\text{No contestar}) = f_i(\text{No contestar}) = \frac{n_i}{n} = \frac{1.550}{2.000} = 0,775$$

- B. La probabilitat obtinguda és teòrica o experimental? És real o estimada?  
Aquesta probabilitat obtinguda és experimental, ja que ens ha fet falta realitzar un estudi empíric per tal de determinar-la.

És una probabilitat estimada, ja que no s'ha obtingut fent l'estudi de mercat amb tota la població de Sabadell, sinó amb una mostra de 2.000 persones.

**Activitat 9**

Volem fer un estudi estadístic per determinar la probabilitat que una dona catalana d'entre 35 i 37 anys escollida a l'atzar tingui la intenció d'ésser mare biològica abans dels 40 anys.

Quina mida ha de tenir la mostra si s'accepta un error de inferior al 3 % i un nivell de confiança del 95 %?

Aplicant la fórmula per calcular la mida de la mostra amb un nivell de confiança del 95 % i un error de 0,03:

$$n = \frac{1}{e^2} = \frac{1}{0,03^2} = \frac{1}{0,0009} = 1.111 \text{ persones}$$

**Activitat 10**

Si per determinar la probabilitat que en unes eleccions un determinat candidat sigui votat pels electors, amb un nivell de confiança del 95 %, s'han entrevistat 494 persones, quin error s'ha acceptat en aquest estudi?

Si partim de la fórmula  $n = \frac{1}{e^2}$  i aïllem  $e^2$ , obtenim que  $e^2 = \frac{1}{n}$ .

Ara substituïm  $n$  pel seu valor i fem l'arrel quadrada per obtenir l'error  $e$ :

$$e^2 = \frac{1}{n} = \frac{1}{494} \qquad e = \sqrt{\frac{1}{494}} = \frac{1}{22,23} = 0,045$$

Així doncs, s'accepta un error del 4,5 %.



## Activitat 1

Classifica les variables següents:

Diàmetre de monedes, programa de TV vist ahir a les 19 hores, cançó més escoltada l'any passat, nombre de títols universitaris, consum de benzina en cent quilòmetres, nombre d'àpats fora de casa, beguda preferida, pes dels bolígrafs, nombre de radiadors a casa.

Variables qualitatives	Variables quantitatives discretes	Variables quantitatives contínues
Programa de TV vist ahir a les 19 hores	Nombre de títols universitaris	Diàmetre de monedes
Cançó més escoltada l'any passat	Nombre d'àpats fora de casa	Consum de benzina en cent quilòmetres
Beguda preferida	Nombre de radiadors a casa	Pes dels bolígrafs

## Activitat 2

En Cristian va sembrar un grapat de llavors de julivert i van néixer 25 plantes. Quaranta dies després de sembrar, va comptar les fulles que tenia cada planta i va recollir aquestes dades:

4 7 2 8 6 3 5 8 11 8  
 5 2 8 5 6 5 6 2 3 7  
 9 8 8 7 5

Fes una taula on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

Nombre de fulles	$n_i$	$N_i$	$f_i$	$F_i$
2	3	3	0,12	0,12
3	2	5	0,08	0,2
4	1	6	0,04	0,24
5	5	11	0,2	0,44
6	3	14	0,12	0,56
7	3	17	0,12	0,68
8	6	23	0,24	0,92
9	1	24	0,04	0,96
10	0	24	0	0,96
11	1	25	0,04	1
	<b>N = 25</b>		<b>Total <math>f_i = 1</math></b>	

**Activitat 3**

Un estudi fet a partir de 400 entrevistes a persones usuàries de la línia 1 del metro de Barcelona en un dia laborable de 9 a 10 h del matí indica que 120 de les persones entrevistades no resideixen a Barcelona.

D'acord amb aquestes dades, digues:

- A. Quina és la probabilitat que una persona que utilitza el metro en les condicions esmentades no resideixi a Barcelona?

La probabilitat que una persona, en les condicions esmentades, no resideixi a Barcelona és:

$$p_o = \frac{120}{400} = 0,3$$

- B. Explica si la probabilitat obtinguda és real o estimada.

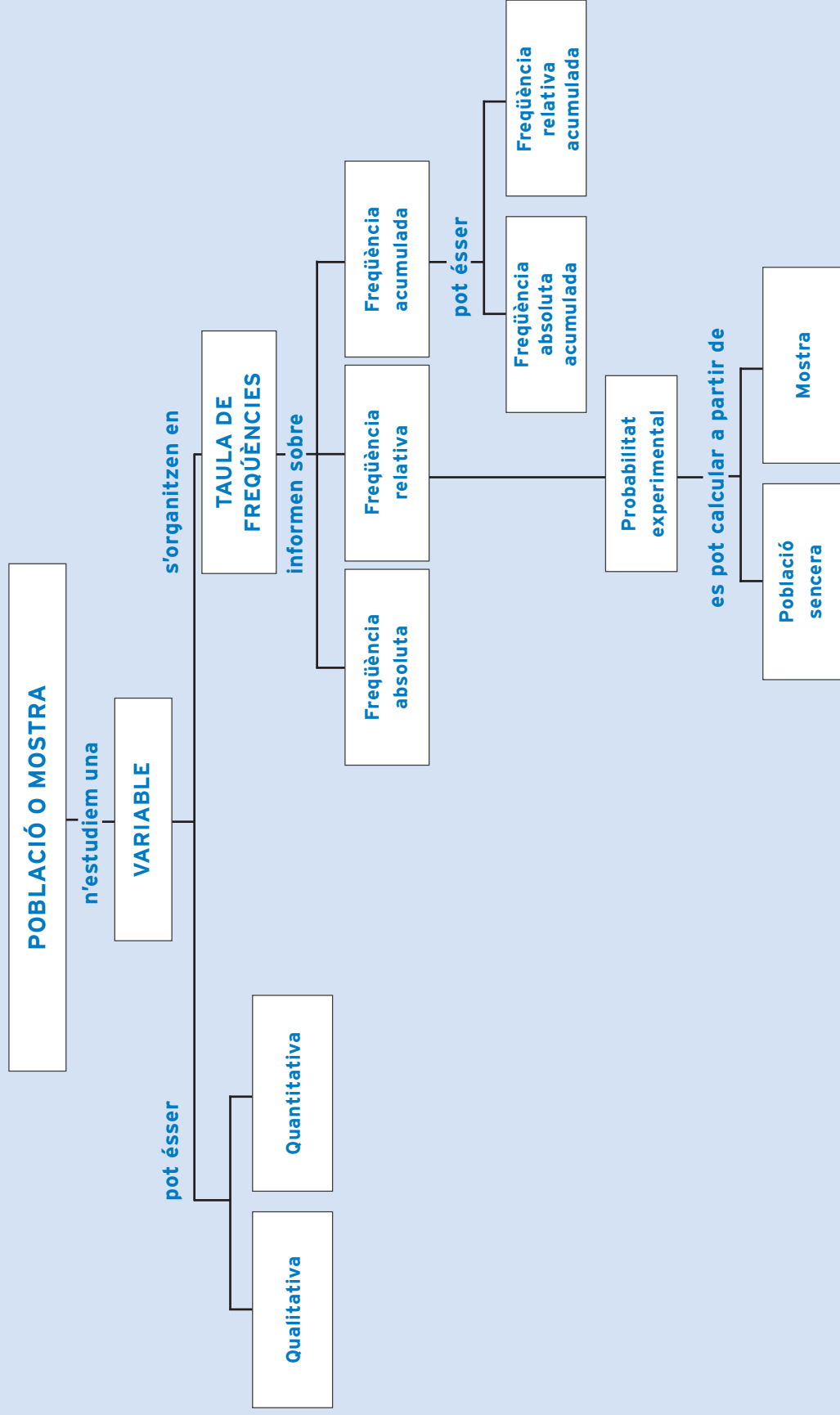
La probabilitat obtinguda és estimada, ja que no s'han entrevistat totes les persones usuàries de la línia 1 del metro de Barcelona en dies laborables de 9 a 10 h del matí.

- C. Quantes persones s'haurien d'haver entrevistat acceptant un error inferior al 3,6 % amb un nivell de confiança del 95 %?

Per esbrinar quantes persones s'haurien d'haver entrevistat, acceptant un error inferior al 3,6 % amb un nivell de confiança del 95 %, hem d'aplicar la fórmula de la mida d'una mostra:

$$n = \frac{1}{e^2} = \frac{1}{0,036^2} = \frac{1}{0,001296} = 772$$

S'haurien d'haver entrevistat 772 persones.



# com ho porto?

Omple la graella següent posant una creu on correspongui:

En acabar la unitat sóc capaç de:	Bé	Regular	Malament
Distingir els termes <i>població</i> , <i>mostra</i> i <i>variable</i> .			
Distingir entre les variables qualitatives i les variables quantitatives.			
Reconèixer les variables quantitatives discretes i les variables quantitatives contínues.			
Precisar les característiques dels diferents tipus de freqüències estadístiques que es poden considerar a partir d'un recull de dades.			
Elaborar taules de freqüències a partir del càlcul dels diferents tipus de freqüències estadístiques: $n_i$ , $N_i$ , $f_i$ i $F_i$ .			
Distingir entre la probabilitat teòrica i la probabilitat experimental.			
Calcular la probabilitat experimental a partir de la població sencera i a partir d'una mostra de la població.			

# Unitat 2

39

## LA REPRESENTACIÓ GRÀFICA DE LES DADES

UNITAT 2 LA REPRESENTACIÓ GRÀFICA DE LES DADES

UNITAT 2

19. ESTADÍSTICA

Matemàtiques, Ciència i Tecnologia

# què treballaràs?

En aquesta unitat aprendràs a construir i interpretar diagrames estadístics com: el diagrama de barres, el polígon de freqüències, el pictograma i el diagrama de sectors.

En acabar aquesta unitat has de ser capaç de:

- Construir les principals representacions gràfiques d'estudis estadístics.
- Interpretar les principals representacions gràfiques.
- Construir una taula de freqüències a partir de les dades contingudes en les gràfiques.
- Precisar quines freqüències estadístiques cal utilitzar en l'elaboració de les diferents representacions gràfiques.

Un cop tenim les dades ben ordenades en una taula de freqüències, és molt útil fer-ne una representació gràfica que ens ajudi a copsar la realitat d'un cop d'ull. Hi ha diverses maneres de representar gràficament les dades. Ara només n'estudiarem les més habituals: diagrama de barres, polígon de freqüències, pictograma i diagrama de sectors. Cada una té unes característiques especials que la fan adient per representar determinades dades estadístiques.

Malgrat que en aquest mòdul aprendrem a construir les gràfiques de manera manual, hi ha molts programes informàtics que permeten construir representacions gràfiques de tota mena. Si disposes d'ordinador i d'un programa adient, també t'animem a provar d'elaborar les gràfiques informàticament.

Per entendre millor com es construeixen els diferents diagrames, veurem un exemple que anirem utilitzant en la construcció de cada un. D'aquesta manera, podràs comparar millor els diferents diagrames. Enunciem ara l'exemple i construïm la taula de freqüències corresponent:

Imagina't que preguntem als vint estudiants d'una classe quants rellotges de canell tenen i que les seves respostes són:

1 5 2 2 4 0 1 2 8 1  
1 2 4 5 4 3 1 0 3 0

Per construir un diagrama, primer haurem de fer una taula de freqüències a partir de les dades obtingudes. Farem una taula amb tots els tipus de freqüències estudiades. Quan dibuixem un determinat diagrama només farem servir les freqüències que ens siguin necessàries.

Nombre de rellotges	$n_i$	$N_i$	$f_i$	$F_i$
0	3	3	0,15	0,15
1	5	8	0,25	0,4
2	4	12	0,2	0,6
3	2	14	0,1	0,7
4	3	17	0,15	0,85
5	2	19	0,1	0,95
6	0	19	0	0,95
7	0	19	0	0,95
8	1	20	0,05	1
	$N = 20$		Total $f_i = 1$	

## 1. Diagrama de barres

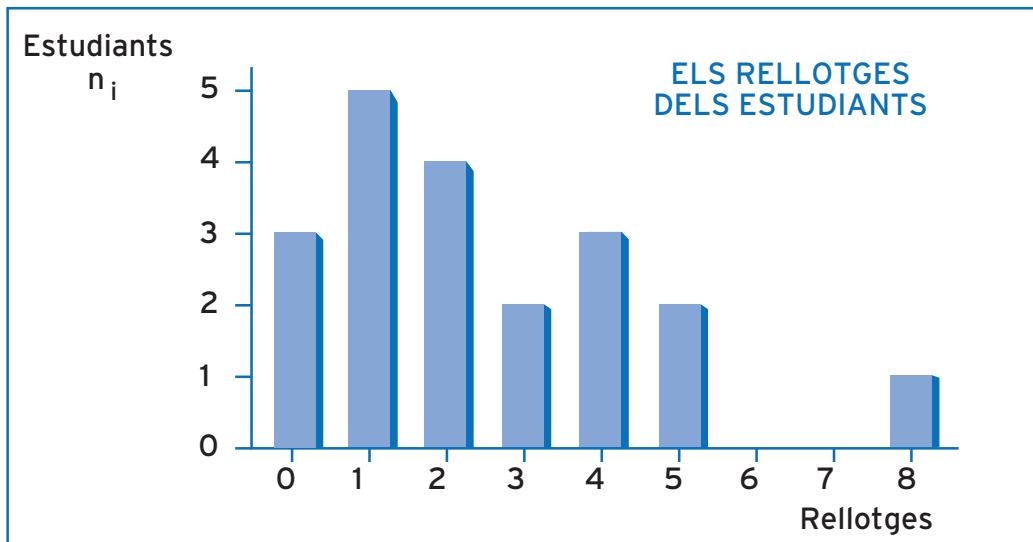
És una de les representacions gràfiques més habituals. Es tracta de dibuixar barres de mida proporcional a la freqüència de cada un dels valors de la variable estadística.

Per dibuixar un diagrama de barres cal que segueixis les instruccions següents:

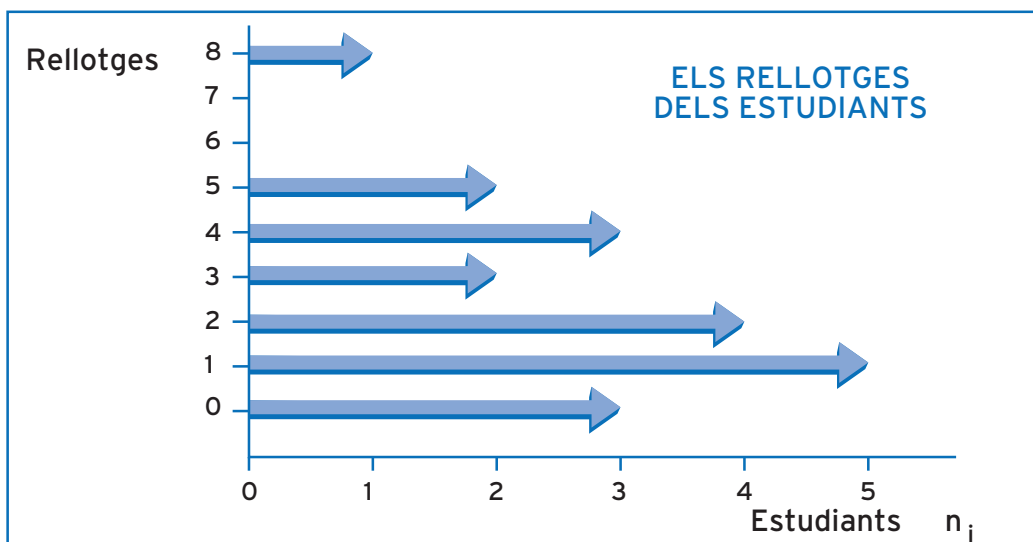
- Agafa un full quadriculat.
- Dibuixa uns eixos de coordenades.

- Posa-li nom a cada un dels eixos. Sobre l'eix d'abscisses (l'horitzontal) representem els valors de la variable i sobre l'eix d'ordenades (el vertical) representem les freqüències.
- Numera l'eix d'ordenades, començant amb el 0. Si cal, utilitza una escala: de dos en dos, de cinc en cinc...
- ESCRIU els noms de les barres sota l'eix d'abscisses.
- Dóna una amplada igual a totes les barres.
- Deixa un espai entre les barres si no treballes amb dades agrupades.
- Per dibuixar una barra comença per fer la línia del topall, que estarà determinada per la freqüència del valor de la variable que representa.
- Posa-li un nom al diagrama de barres, preferentment a la part superior.

Per fer aquest diagrama de barres utilitzarem les freqüències absolutes, tot i que també es poden fer diagrames de barres utilitzant les freqüències relatives, generalment expressades en percentatges. A l'hora de fer un exercici, considerarem les freqüències absolutes si l'enunciat no especifica una altra cosa.



Malgrat que la manera més habitual de dibuixar un diagrama de barres és utilitzar l'eix d'abscisses pels valors de la variable i l'eix d'ordenades per a les freqüències, també es poden utilitzar els eixos de l'inrevés, és a dir, l'eix d'abscisses per a les freqüències i l'eix d'ordenades per als valors de la variable:





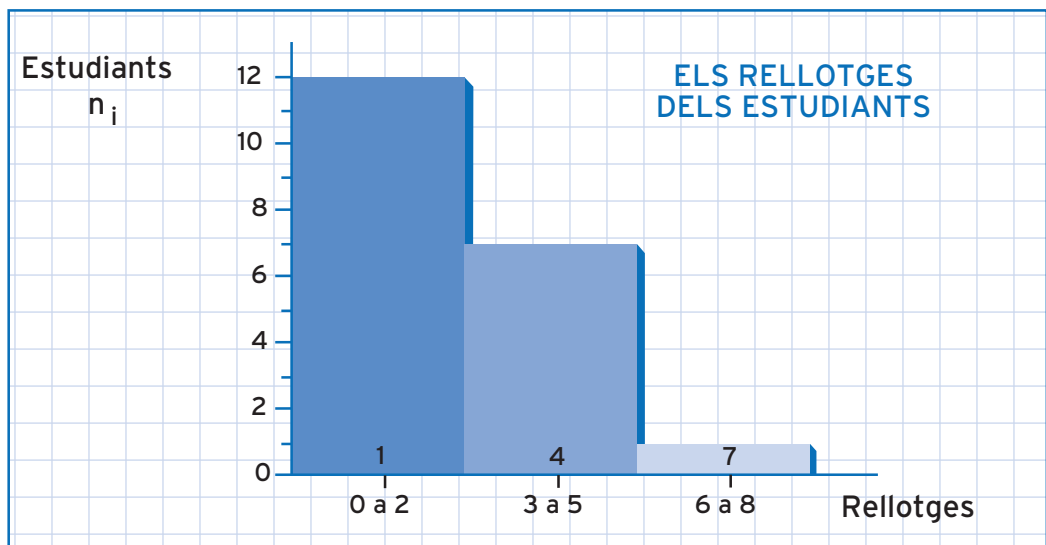
De vegades, agrupem alguns valors de la variable per tal de simplificar i concentrar la informació obtinguda. Les agrupacions s'anomenen classes.

En el nostre exemple, podem fer agrupacions de tres valors de la variable: 0 a 2, 3 a 5 i 6 a 8.

El valor mitjà de cada una de les agrupacions s'anomena marca de classe. En el cas del nostre exemple, les marques de classe són:

Nombre de rellotges	Marca de classe	$n_i$
0 a 2	1	12
3 a 5	4	7
6 a 8	7	1

Quan treballem amb classes no s'ha de deixar cap espai entre les barres de les classes. Les marques de classe es posen a les bases de les barres corresponents. Així doncs, en resultarà aquest diagrama:



Aquest tipus de diagrama de barres també rep el nom d'histograma, i és especialment utilitzat quan es treballa amb variables quantitatives contínues, com el pes o l'altura, ja que la variable pren valors infinits. Malgrat que nosaltres representarem histogrames amb classes que agrupin un mateix nombre de valors de la variable (intervals iguals), també és possible fer classes amb agrupacions diferents (intervals diferents), si això ajuda a interpretar millor les dades. En aquest cas, l'amplada de cada barra ha d'ésser proporcional al nombre de valors agrupats.

Agrupar valors de la variable està especialment recomanat quan n'hi ha molts, perquè clarifica la informació. L'agrupació de dades en classes es pot dur a terme a l'hora de dibuixar qualsevol tipus de gràfic.

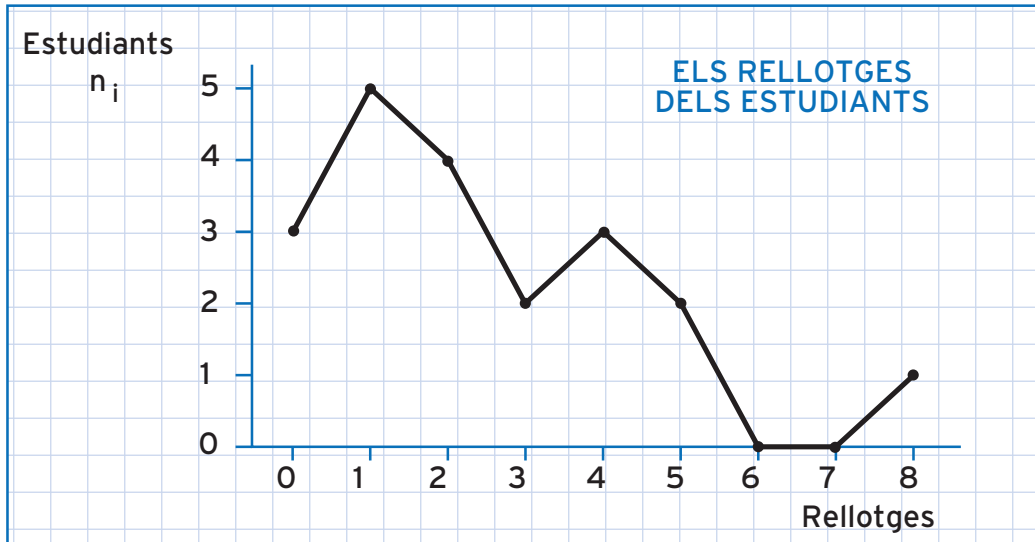
- **Fes les activitats d'aprenentatge 1, 2, 3 i 4**

## 44 2. Polígon de freqüències

Aquesta representació gràfica és molt utilitzada en estudis estadístics on la variable és temporal: l'IPC mensual, les cotitzacions en borsa d'una empresa durant una setmana, els litres de pluja per metre quadrat recollits durant cada un dels dies del mes de març, etc.

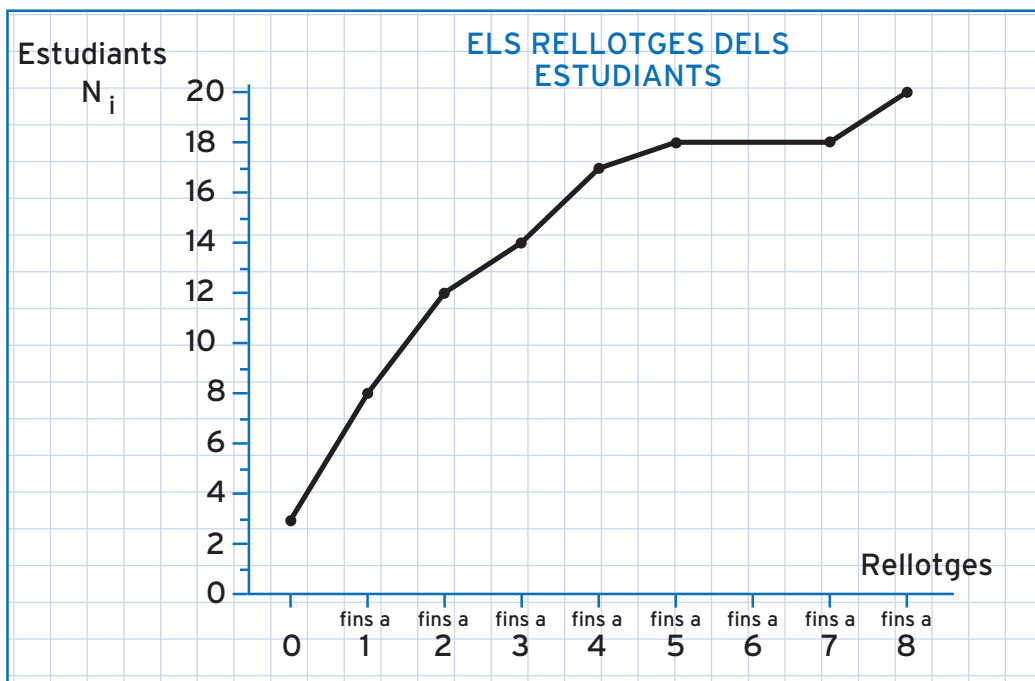
Per obtenir un **polígon de freqüències** cal dibuixar una línia poligonal unint els punts mitjans dels topalls de les barres d'un diagrama de barres.

Seguint amb el mateix exemple dels rellotges dels estudiants, i utilitzant les freqüències absolutes, obtenim:

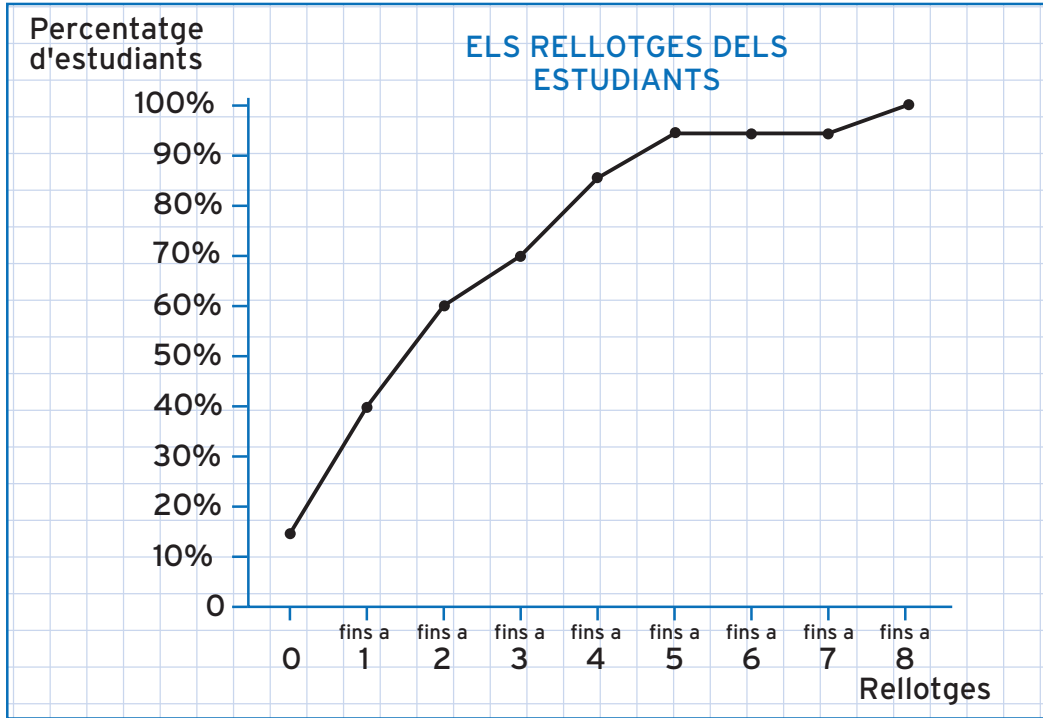


Quan el polígon de freqüències es dibuixa a partir de les freqüències acumulades, el polígon de freqüències resultant sempre és creixent, és a dir que la línia poligonal mai no va enrere, sempre es manté en el mateix nivell o puja de nivell.

Dibuixem ara un **polígon de freqüències acumulades**, seguint amb el nostre exemple:



De la mateixa manera que passava amb els diagrames de barres, es poden dibuixar polígons de freqüències (acumulades o no) a partir de les freqüències absolutes o de les relatives. Els diagrames realitzats a partir de les freqüències relatives són pràcticament idèntics, malgrat la manera diferent d'expressar la informació. Si aquest darrer polígon de freqüències acumulades el dibuixem a partir de les freqüències relatives acumulades, expressades en percentatge, obtenim:



- Fes les activitats d'aprenentatge 5, 6 i 7

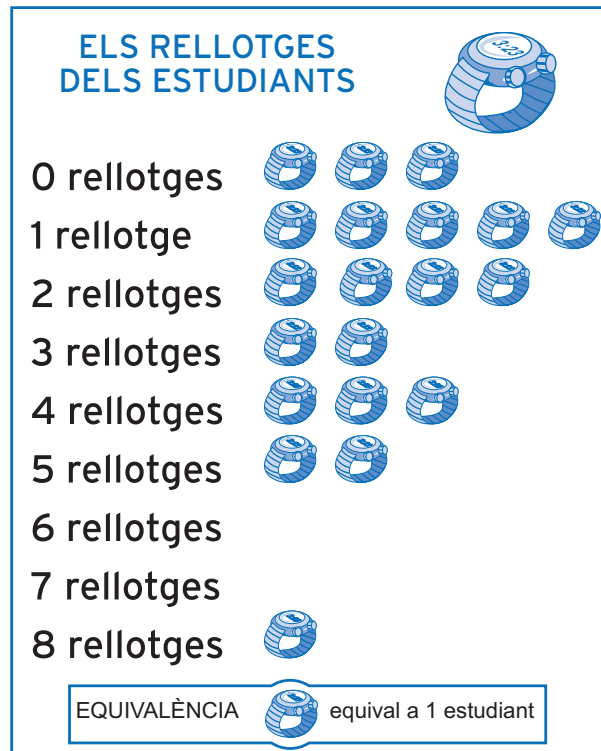
## 46 3. Pictograma

És una representació gràfica que cada cop s'utilitza més en diaris i revistes d'actualitat.

Per dibuixar un pictograma cal que segueixis les instruccions següents:

- Agafa full ratllat o quadriculat.
- Fes els dibuixos de la mateixa mida, forma i color.
- Deixa espais iguals entre els dibuixos.
- Posa l'equivalència entre cada dibuix i la freqüència absoluta que representa.
- Dóna-li un títol al pictograma.

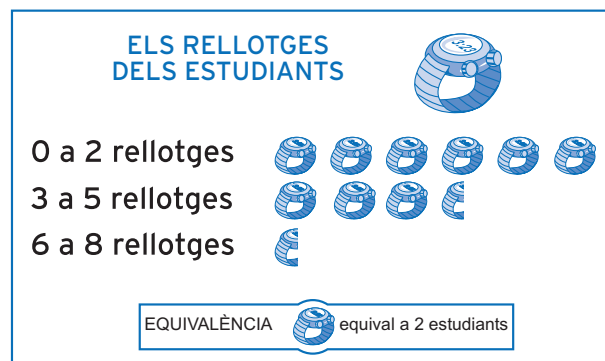
Seguint aquestes passes en el nostre exemple:



En els pictogrames utilitzem sempre les freqüències absolutes.

Podem agrupar alguns valors de la variable per tal de simplificar el pictograma i concentrar la informació obtinguda. Per exemple:

Nombre de rellotges	$n_i$
0, 1, 2	12
3, 4, 5	7
6, 7, 8	1
	$N = 20$



Fixa't que ara cada rellotge equival a dos estudiants, la qual cosa ens estalvia haver de fer llargues fileres de rellotges.

- Fes les activitats d'aprenentatge 8, 9 i 10

## 4. Diagrama de sectors

És un altre gràfic molt utilitzat per representar les dades recollides. Es tracta d'una mena de formatge en porcions, en què la grandària de cada porció és proporcional a la freqüència del valor de la variable que representa.

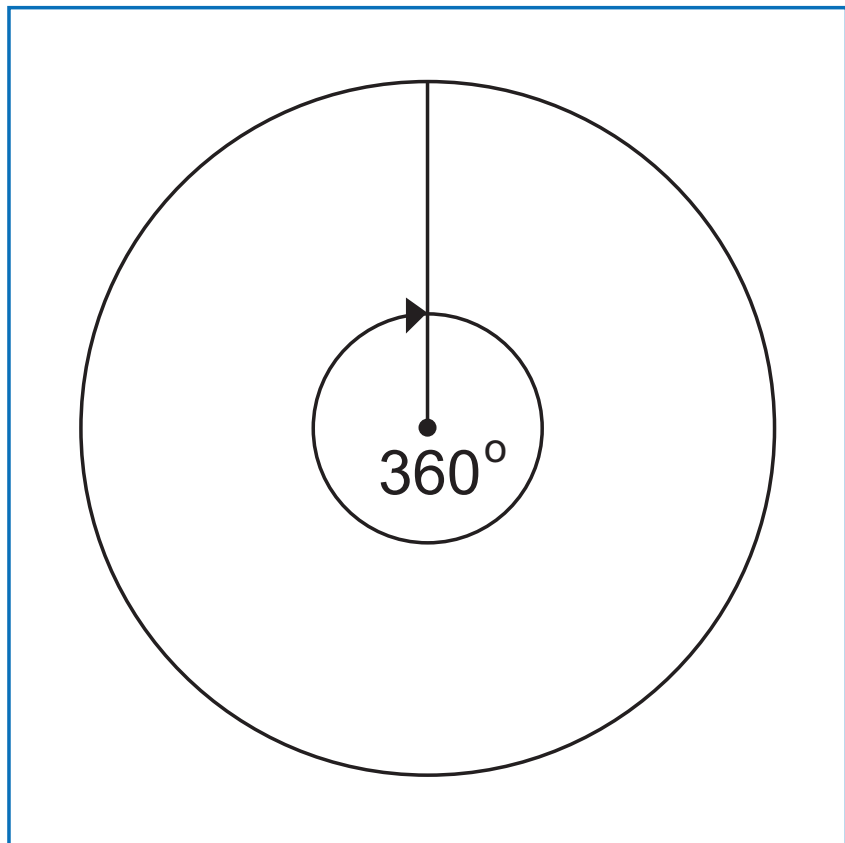
Hi ha dues variants: el diagrama de sectors circular i el diagrama de sectors semicircular. El diagrama de sectors circular es construeix utilitzant la totalitat del cercle. Si utilitzem únicament la meitat del cercle, obtenim un diagrama de sectors semicircular.

Per dibuixar un **diagrama de sectors circular** cal que segueixis els passos següents:

- Dibuixa una circumferència amb un radi d'entre 4 i 6 cm (6 és una bona tria) i marca-hi el centre.
- Dibuixa un radi vertical des del centre cap amunt, fins a la circumferència.
- Calcula l'angle que correspon a cada sector.
- Utilitza un transportador d'angles per mesurar cada sector, en el sentit de les busques del rellotge.
- Pinta cada un dels sectors i posa-li el nom del valor de la variable que representa. També hi podem afegir la seva freqüència absoluta o la freqüència relativa expressada com a percentatge.
- Encapçala el diagrama de sectors circular amb un títol.

Ara seguirem aquests passos en el nostre exemple dels rellotges dels estudiants.

Per començar, dibuixem un cercle, marquem el centre i dibuixem un radi vertical cap amunt:

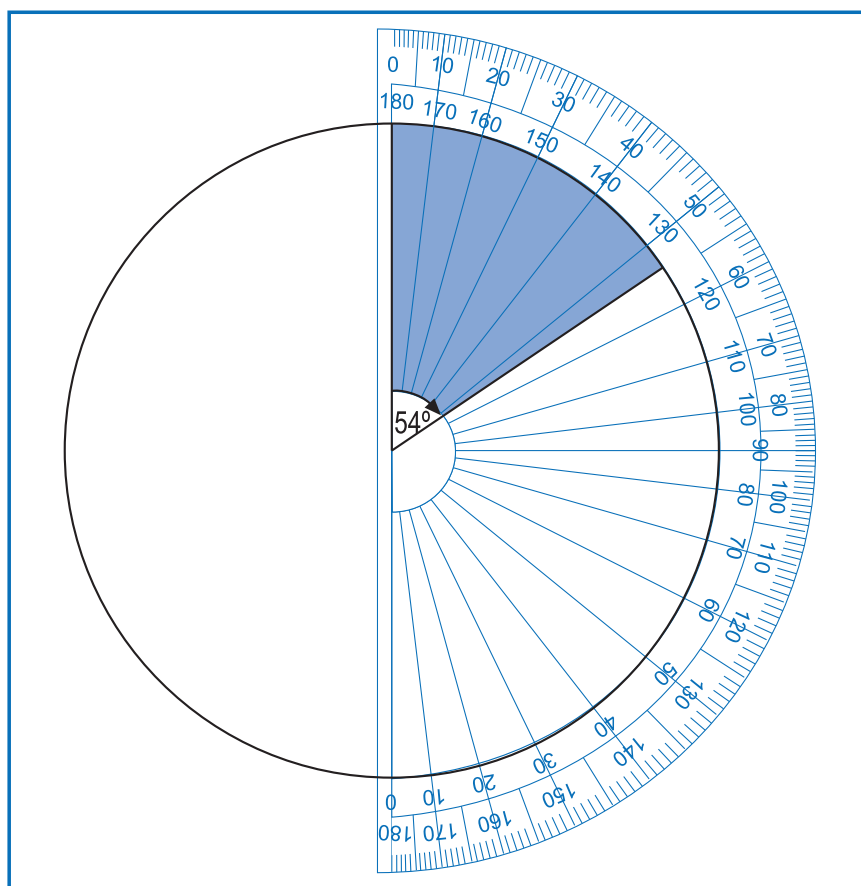


Ara calculem l'angle que correspon a cada sector. Per fer-ho, multipliquem la freqüència absoluta per  $360^\circ$  i dividim el resultat entre  $N$ , el nombre total de dades:

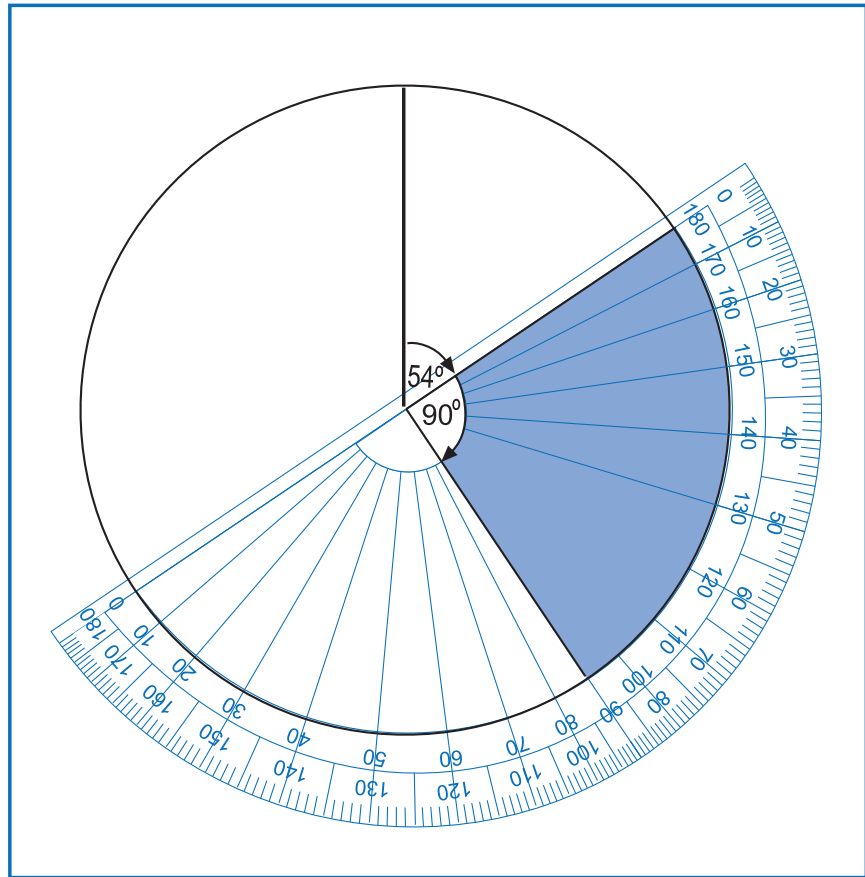
Nombre de rellotges	$n_i$	Angle
0	3	$3 \times 360^\circ : 20 = 54^\circ$
1	5	$5 \times 360^\circ : 20 = 90^\circ$
2	4	$4 \times 360^\circ : 20 = 72^\circ$
3	2	$2 \times 360^\circ : 20 = 36^\circ$
4	3	$3 \times 360^\circ : 20 = 54^\circ$
5	2	$2 \times 360^\circ : 20 = 36^\circ$
6	0	$0 \times 360^\circ : 20 = 0^\circ$
7	0	$0 \times 360^\circ : 20 = 0^\circ$
8	1	$1 \times 360^\circ : 20 = 18^\circ$
	$N = 20$	Total = $360^\circ$

Ara utilitzem el transportador d'angles per mesurar l'angle corresponent a cada sector, en el sentit de les busques del rellotge. Dibuixem un sector per a cada valor de la variable, excepte els que tinguin freqüència absoluta zero ( $n_i = 0$ ), que no tindran sector.

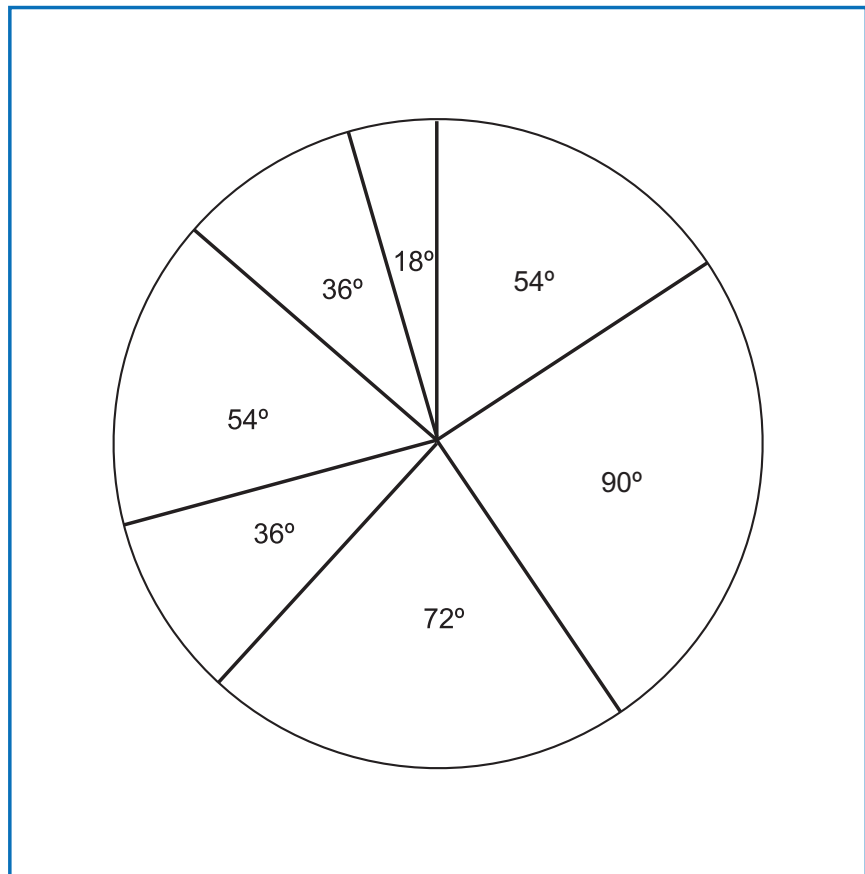
Per dibuixar el sector **0 rellotges** mesurem  $54^\circ$  en el sentit de les busques del rellotge, a partir del radi vertical, i dibuixem un altre radi:



Per dibuixar el sector 1 rellotge mesurem  $90^\circ$  en el sentit de les busques del rellotge, a partir del final del primer sector, i dibuixem un altre radi:



Seguim el mateix procediment per dibuixar la resta de sectors. Finalment obtenim:



Ara només cal pintar cada sector i posar-li el nom corresponent, és a dir, el nom del valor de la variable que representa. També hi farem constar, dins de cada sector, la freqüència relativa en percentatge.

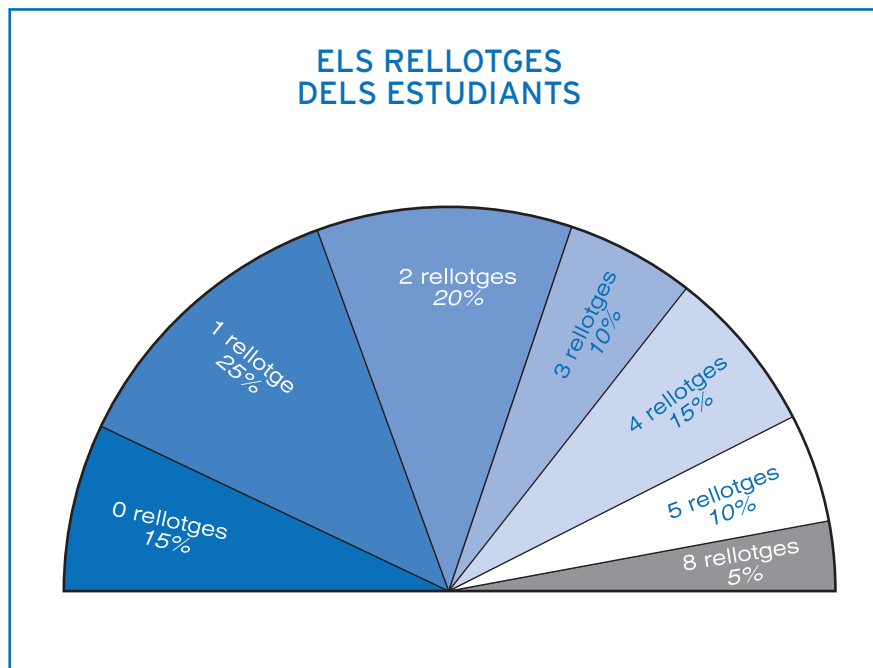
Per dibuixar un **diagrama de sectors semicircular** cal que segueixis les instruccions següents:

- Dibuixa una línia horitzontal que faci entre 8 i 16 cm. Marca el centre de la línia i dibuixa un semicercle amb el compàs. El radi ha de mesurar la meitat que la línia horitzontal.
- Calcula l'angle que li correspon a cada sector, tenint en compte que la quantitat total de graus per repartir és de  $180^\circ$ .
- Utilitza un transportador d'angles per mesurar els angles corresponents a cada sector. Comença des de l'esquerra del semicercle i continua en el sentit de les busques del rellotge.
- Pinta cada un dels sectors i posa-li el nom del valor de la variable que representa. De la mateixa manera que en el circular, també hi podem afegir la freqüència absoluta de cada sector, o la relativa expressada com a percentatge.
- Encapçala el diagrama de sectors semicircular amb un títol.



Seguint aquests passos en el nostre exemple, obtenim aquesta taula i aquest gràfic:

Nombre de rellotges	$n_i$	Angle
0	3	$3 \times 180^\circ : 20 = 27^\circ$
1	5	$5 \times 180^\circ : 20 = 45^\circ$
2	4	$4 \times 180^\circ : 20 = 36^\circ$
3	2	$2 \times 180^\circ : 20 = 18^\circ$
4	3	$3 \times 180^\circ : 20 = 27^\circ$
5	2	$2 \times 180^\circ : 20 = 18^\circ$
6	0	$0 \times 180^\circ : 20 = 0^\circ$
7	0	$0 \times 180^\circ : 20 = 0^\circ$
8	1	$1 \times 180^\circ : 20 = 9^\circ$
	$N = 20$	Total = $180^\circ$



Recorda que en les seccions d'un diagrama de sectors, a més del valor de la variable, es pot posar la seva freqüència absoluta o relativa corresponent. Generalment la freqüència relativa s'expressa en forma de percentatge. En cap cas s'ha d'escriure dins de les seccions el valor de l'angle representat.

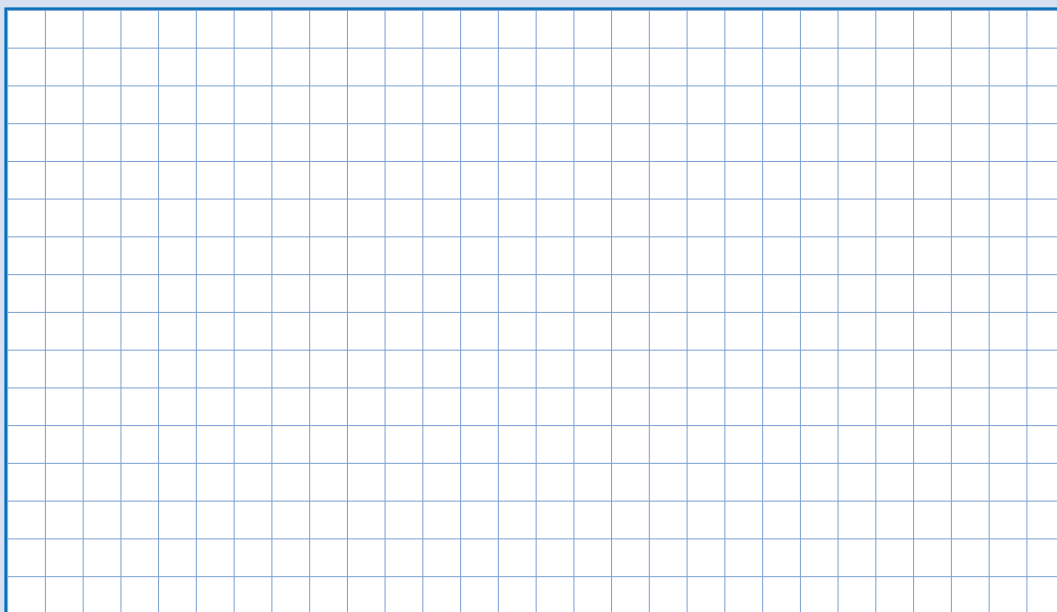
- Fes les activitats d'aprenentatge 11, 12, 13, 14 i 15

**Activitat 1**

En Miquel arrenca totes les males herbes que han envaït la gespa del seu jardí, i, mentre ho fa, va apuntant el nom de cada una de les plantes arrencades:

Ortiga	Ravenissa	Ravenissa	Blet	Fonoll	Ortiga
Ravenissa	Ravenissa	Rosella	Civada	Civada	Fonoll
Civada	Fonoll	Rosella	Ravenissa	Fonoll	Ravenissa
Rosella	Fonoll				

Dibuixa un diagrama de barres amb aquestes dades, utilitzant les freqüències relatives en forma de percentatges.

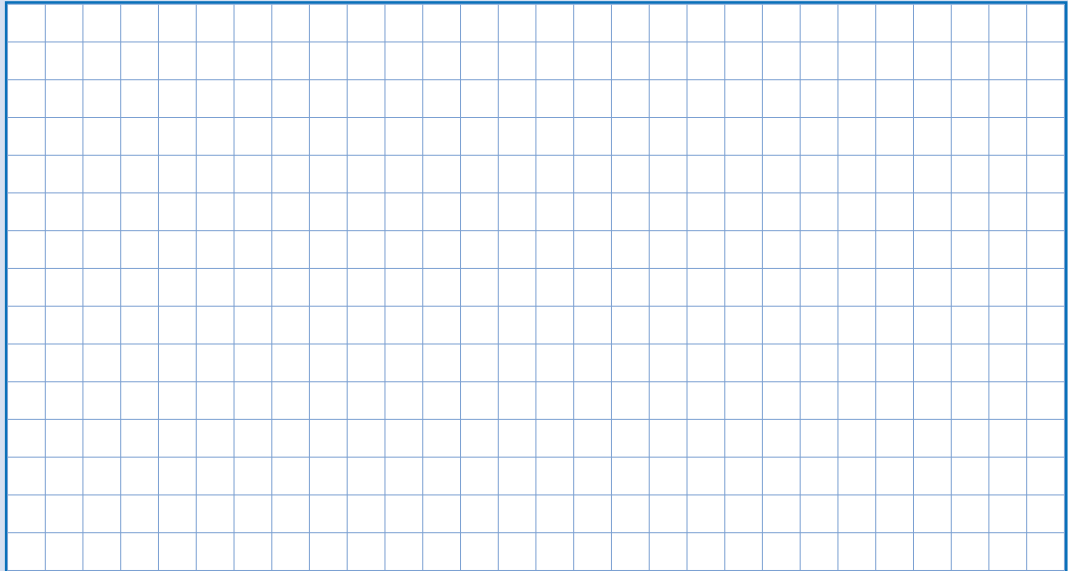


**Activitat 2**

Un estudi sobre l'esport que practiquen els treballadors d'una petita empresa dóna aquests resultats:

Senderisme	Gimnàstica	Ciclisme	Ciclisme	Tennis
Bàsquet	Gimnàstica	Gimnàstica	Tennis	Cap esport
Gimnàstica	Ciclisme	Cap esport	Bàsquet	Gimnàstica

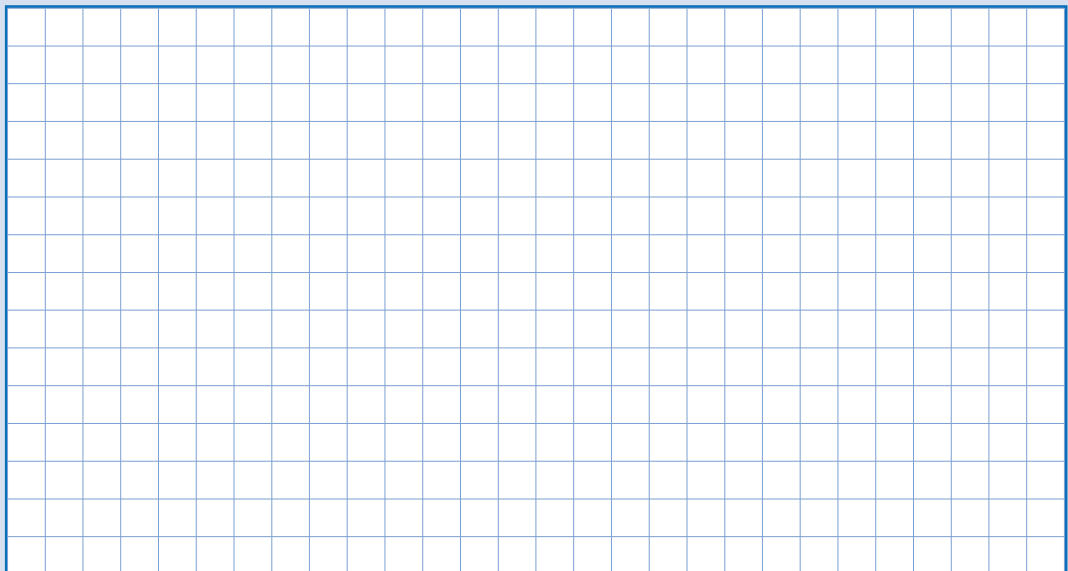
Dibuixa un diagrama de barres amb aquestes dades.

**Activitat 3**

Es va fer un estudi sobre la distància que els estudiants d'un grup de segon curs de filosofia havien de recórrer per arribar a la Universitat. Heus aquí les dades que mostren la distància en quilòmetres que recorria cada un dels estudiants:

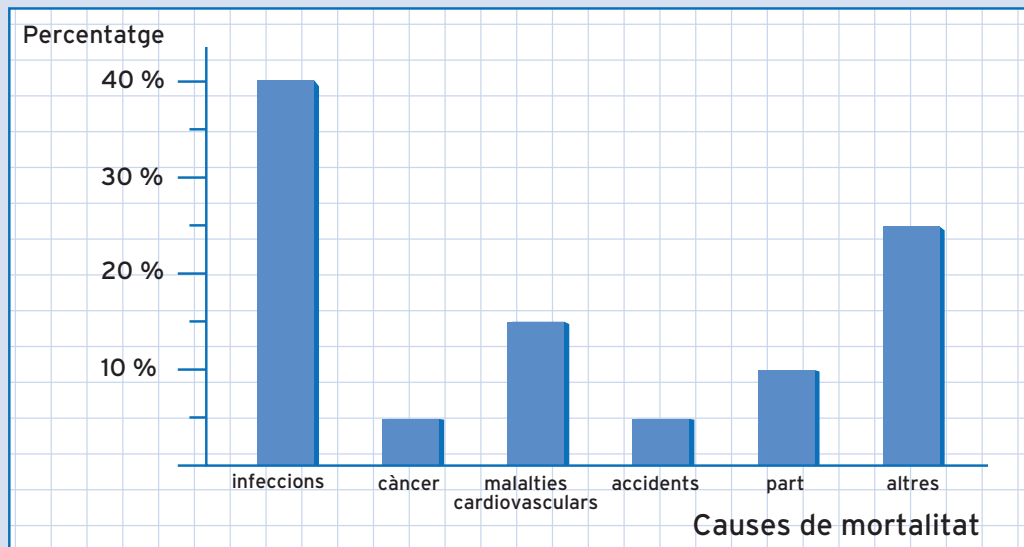
2	1	8	23	19	0	21	2	27	5
6	20	18	14	14	19	5	3	10	9
9	8	7	3	6	27	13	10	19	12
3	21	25	15	5	8	14	6	6	9

Construeix una taula de freqüències, agrupant les distàncies en classes de 5 valors, i dibuixa un histograma, assenyalant les marques de classe.



## Activitat 4

El diagrama de barres adjunt mostra les causes de mortalitat a l'Àfrica central. L'estudi es va fer amb una mostra de 10.000 casos.



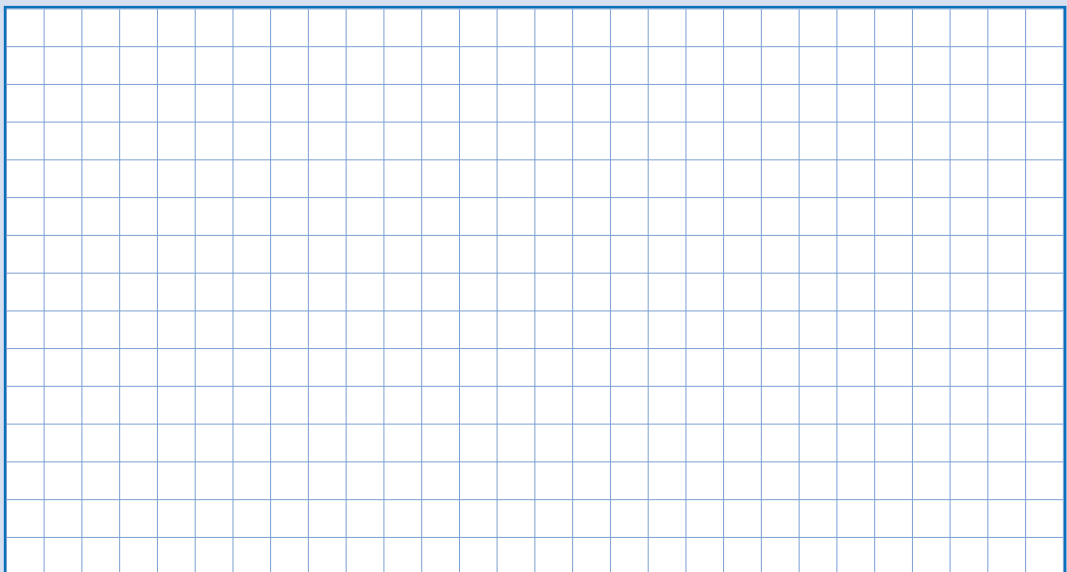
A partir de la informació que dóna aquest diagrama de barres, construeix una taula de freqüències amb totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

**Activitat 5**

Durant un mes de primavera, l'Adela, una noia que viu a pagès, compta el nombre de perdius que veu caminant des de casa fins al poble, i recull aquestes dades:

5	4	6	7	8	6	5	4	7	9
8	4	8	5	12	10	9	6	5	4
4	4	4	7	10	5	8	9	5	6

Amb aquestes dades, dibuixa un polígon de freqüències relatives en percentatges.

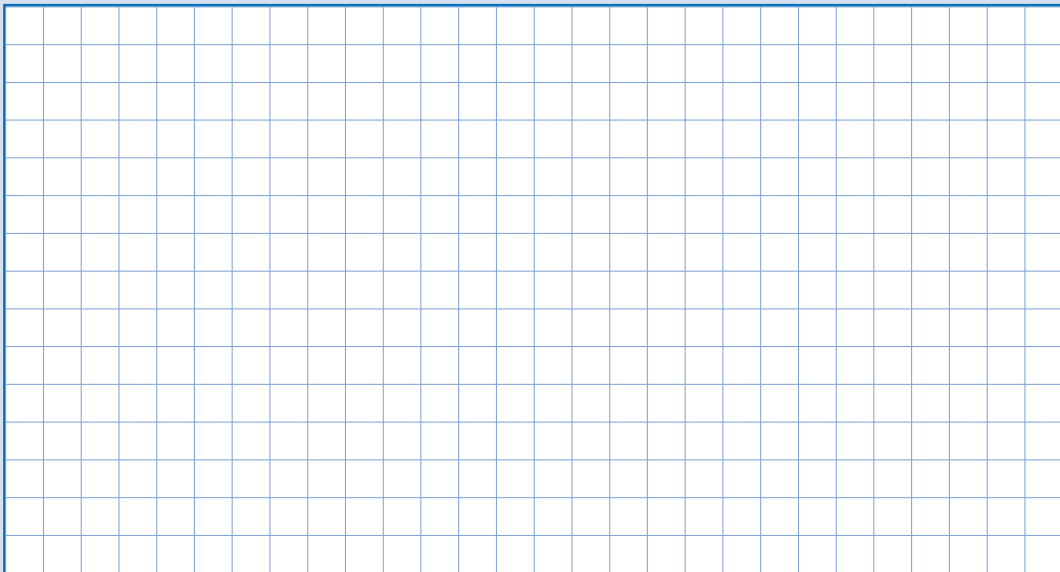


**Activitat 6**

En Rossend estudia el nombre de cries de conill per cada part, en una granja de conills, i recull aquestes dades:

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

A partir d'aquestes dades, dibuixa un polígon de freqüències relatives acumulades en percentatges.

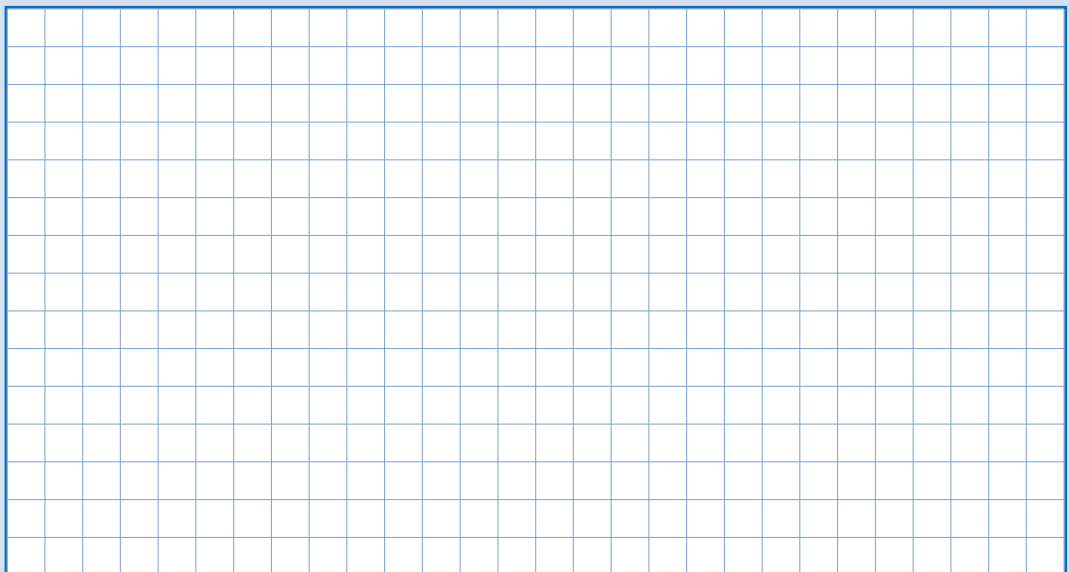
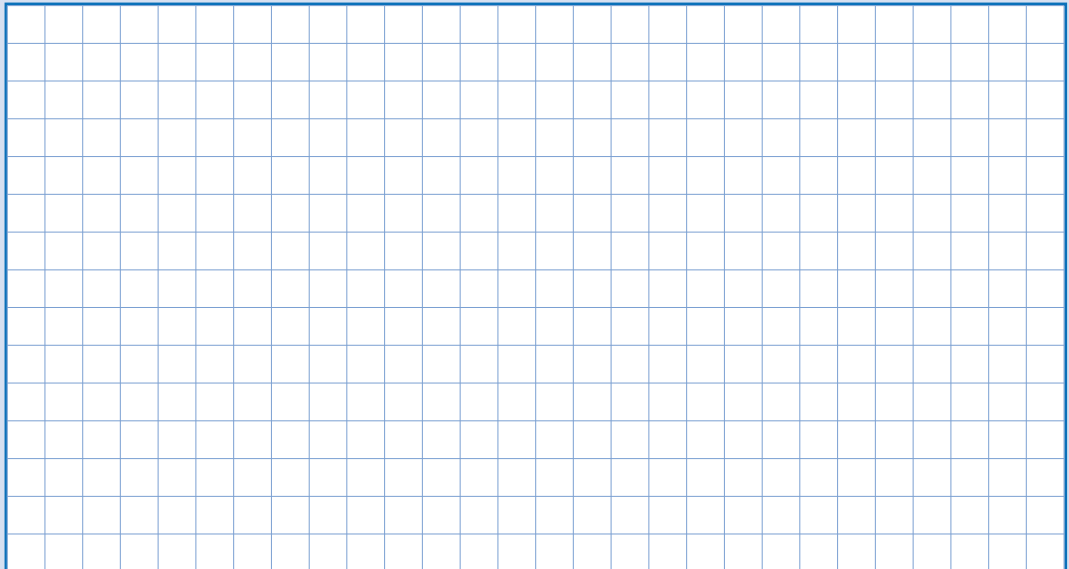


**Activitat 7**

La Laura demana als companys de classe quants cafès van prendre ahir. Heus aquí els resultats:

3	0	5	7	3	2	1	1	1	0
2	1	1	2	4	0	5	8	7	4
3	0	1	1						

Amb aquestes dades, dibuixa un polígon de freqüències i un polígon de freqüències acumulades.

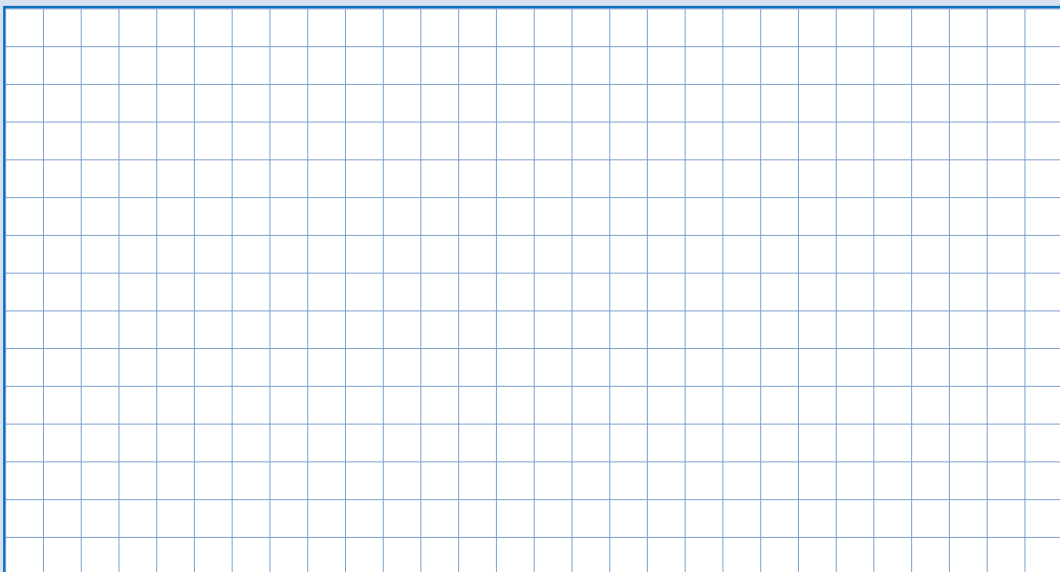
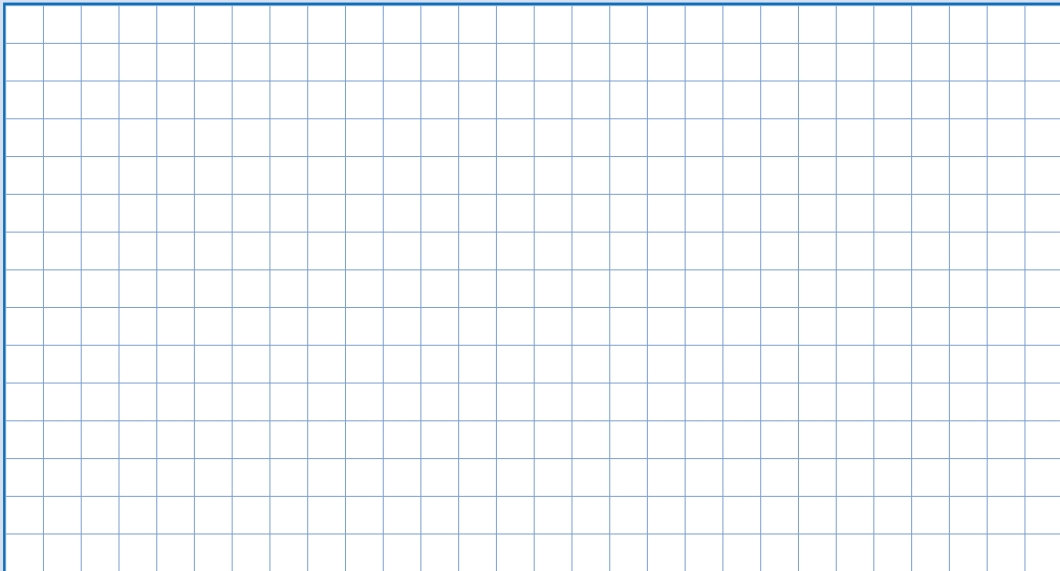


**Activitat 8**

Un investigador recull dades sobre el nombre d'ous de les postes de vint mallerengues. Heus aquí els resultats:

8	10	10	9	8	9	10	11	8	10
9	10	12	9	8	10	10	9	11	7

Amb aquestes dades, agrupades en classes de dos valors de la variable, dibuixa un pictograma i un histograma amb les freqüències relatives en percentatges.



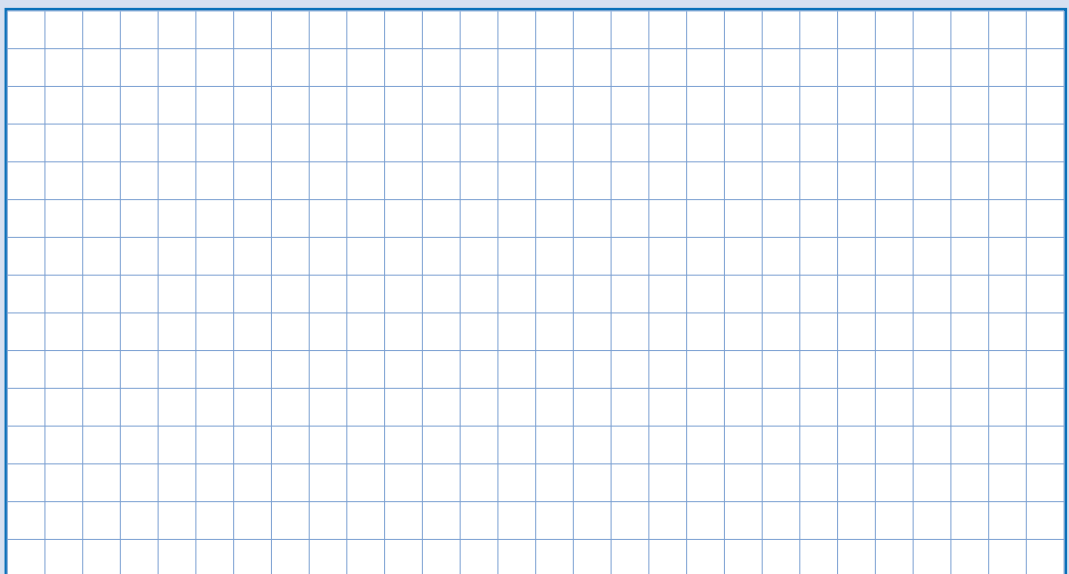
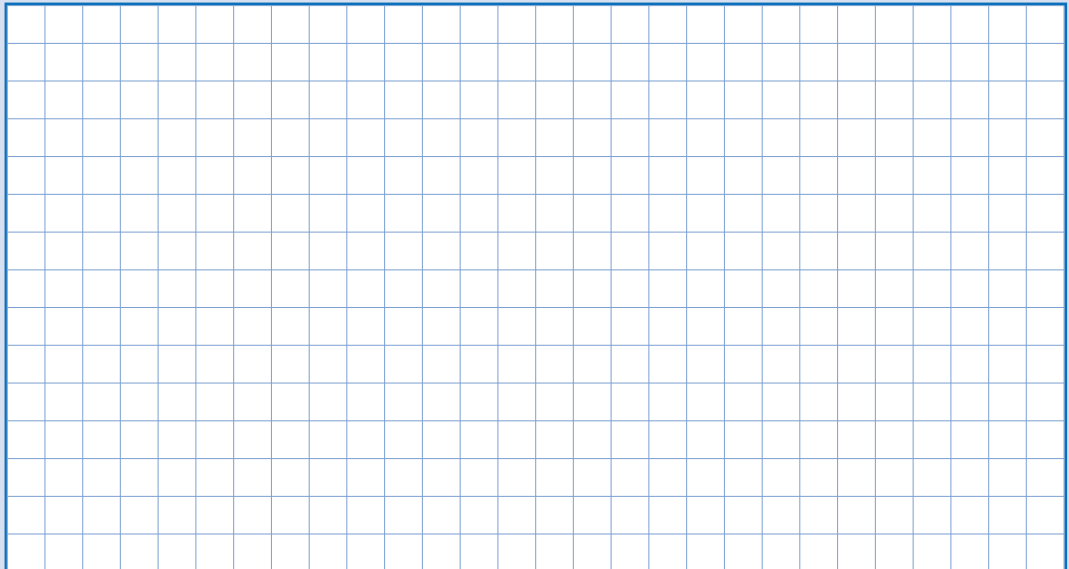


**Activitat 9**

La Maria demana a 30 persones quants viatges amb avió van fer l'any passat. Aquests són els resultats:

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

Dibuixa un pictograma i un polígon de freqüències amb les dades obtingudes.

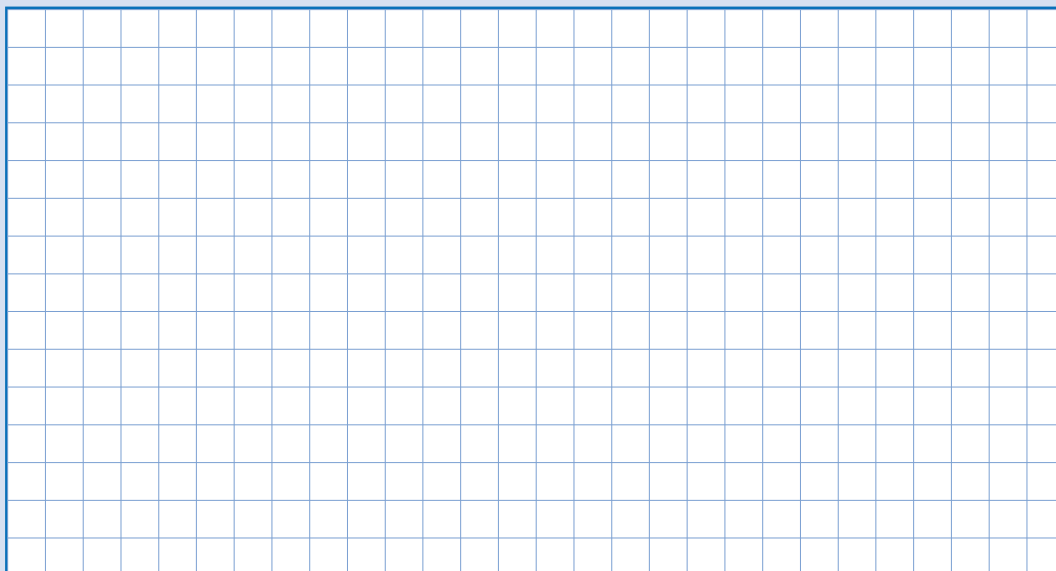


**Activitat 10**

En Joan demana a 25 persones quants llibres van llegir l'any passat. Heus aquí els resultats:

2	1	2	1	1	8	1	0	5	3
3	4	4	2	1	0	0	1	4	1
6	3	9	1	2					

Agrupa les dades obtingudes en classes de dos valors i dibuixa un pictograma i un polígon de freqüències acumulades.

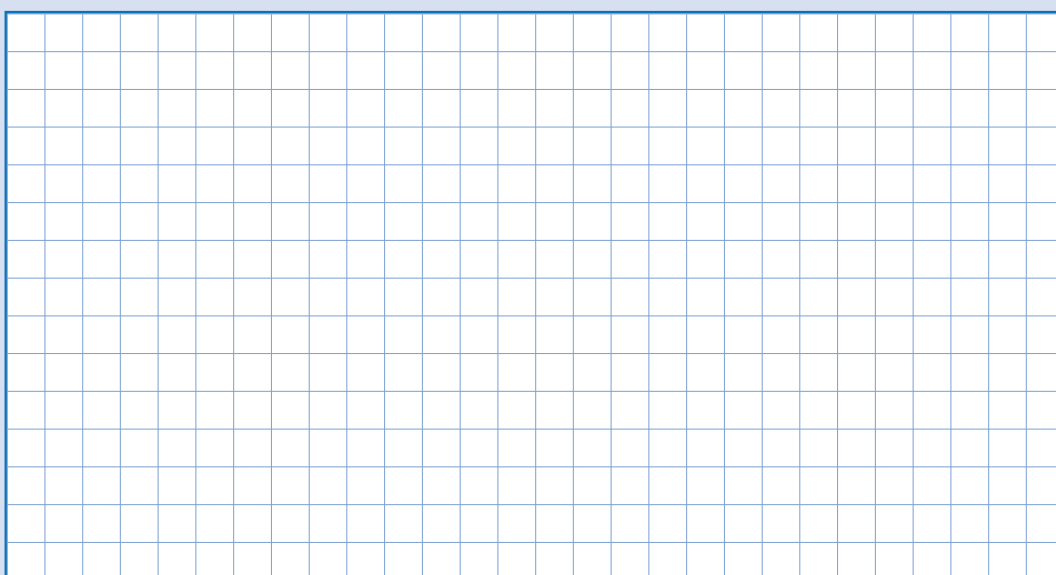
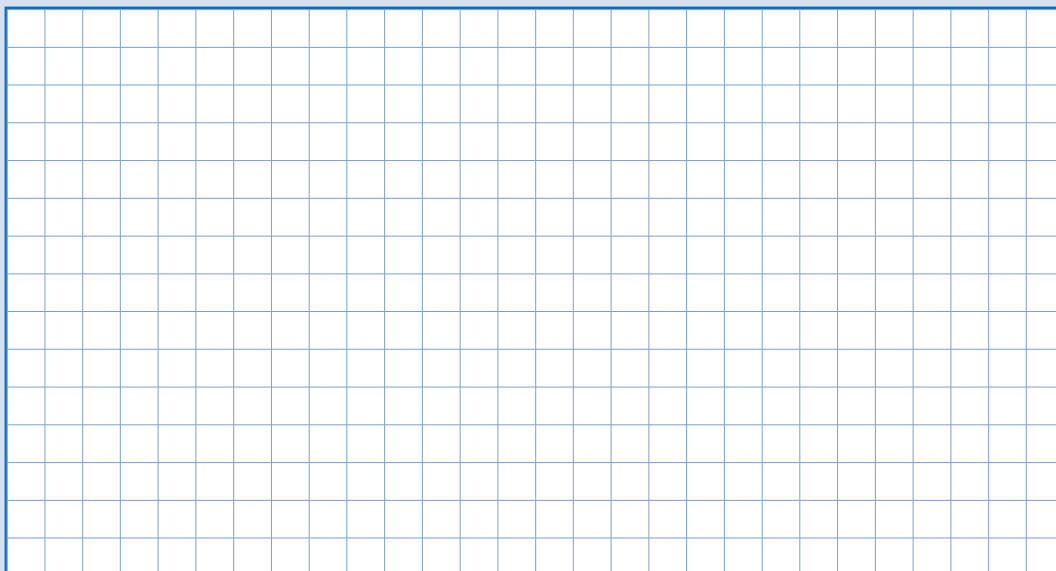


**Activitat 11**

En Francesc demana als seus companys de classe quina edat tenen. Heus aquí els resultats en anys:

18 24 28 18 19 21 24 23 28 45  
29 33 46 20 23 27 21 35 37 40

Agrupa les dades obtingudes en classes de sis valors i dibuixa un diagrama de sectors circular (on constin les freqüències relatives en forma de percentatge dins de cada sector) i un polígon de freqüències.

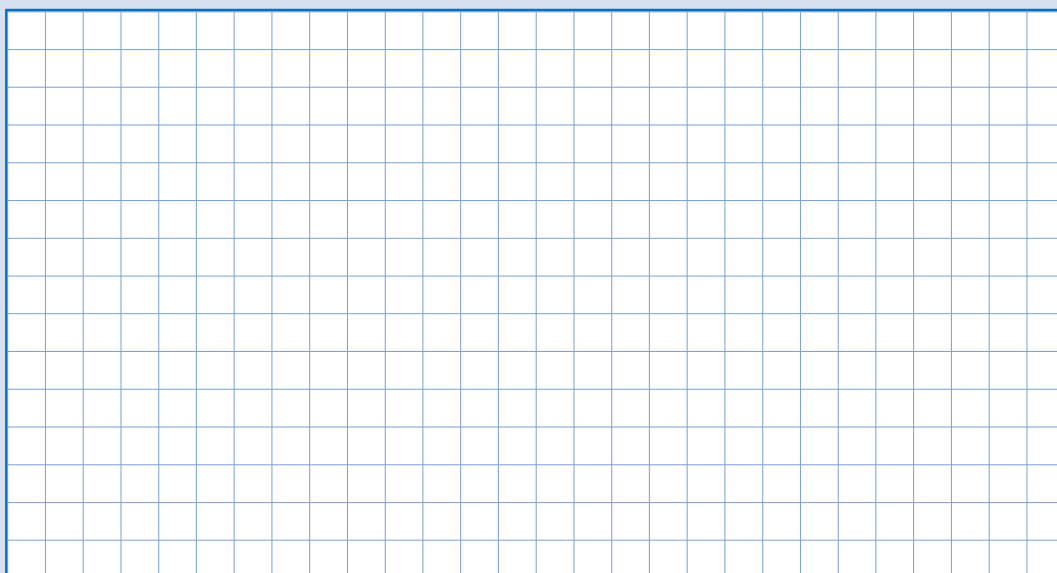


**Activitat 12**

A les darreres eleccions locals de Sinera, el nombre de regidors obtinguts per les diverses forces polítiques fou:

Partits polítics	Regidors
Partit Republicà de Sinera (PRS)	18
Liberals per Sinera (LS)	6
Partit Monàrquic (PM)	3
Partit Nacional Sinerenc (PNS)	12
Partit dels Treballadors de Sinera (PTS)	6

Amb aquestes dades, dibuixa un diagrama de sectors semicircular (on constin les freqüències absolutes) i un diagrama de barres.

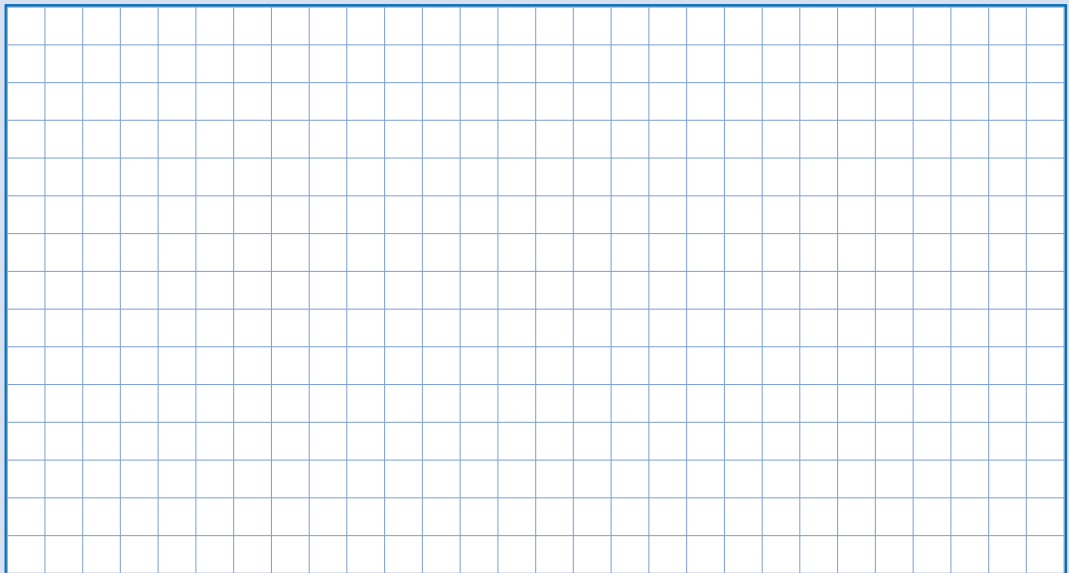


**Activitat 13**

El nombre de gols aconseguït pel F.C. Sinera durant els partits jugats a la darrera lliga nacional foren:

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

Amb aquestes dades, dibuixa un diagrama de sectors semicircular (on constin les freqüències relatives en forma de percentatges) i un pictograma.



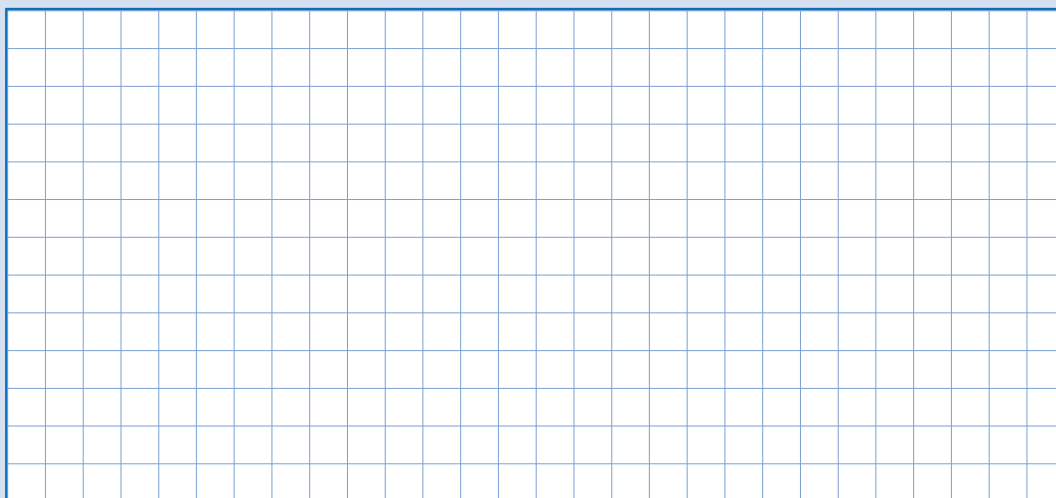
**Activitat 14**

L'Annabel té una petita estació meteorològica on ha recollit dades sobre quantitat de pluja durant els vint-i-cinc primers dies d'octubre. Heus aquí les dades recollides, en litres/m<sup>2</sup>:

1	0	0	1	6	3	8	5	3	9
3	11	3	4	6	2	1	1	1	0
7	6	2	4	3					

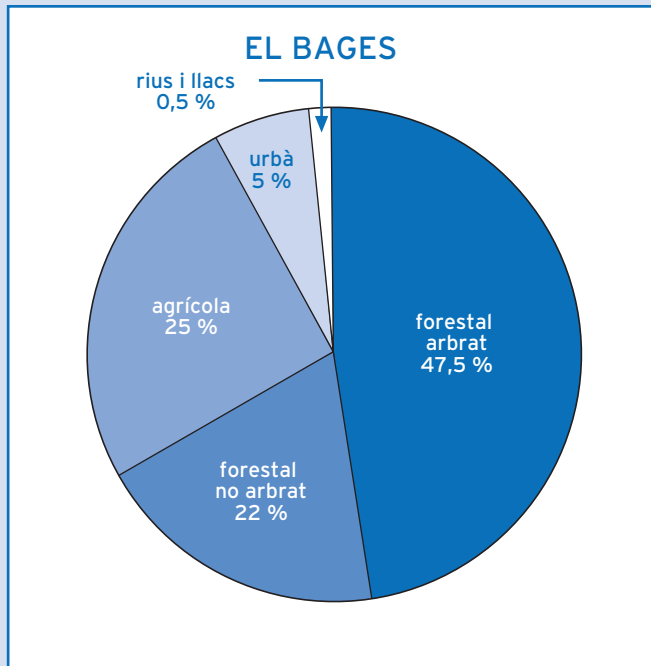
Agrupa les dades en classes de dos valors i dibuixa:

- Un diagrama de sectors circular on consti la freqüència relativa, en forma de percentatge, dins de cada sector.
- Un polígon de freqüències acumulades.



**Activitat 15**

El diagrama de sectors adjunt representa, en percentatges, els usos del territori a la comarca del Bages.



A partir de la informació que dóna aquest diagrama de sectors, i tenint en compte que la superfície de la comarca és de 129.800 hectàrees, construeix una taula de freqüències amb totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

## 66 ACTIVITATS COMPLEMENTÀRIES

### Activitat 1

Aquesta és una activitat de recerca. Tria quin estudi estadístic vols fer. Pot ésser el que vulguis, una variable numèrica o nominal, per exemple, el nombre de pel·lícules vistes al cinema el mes passat, el color dels ulls, etc. Has de definir quina població serà l'objecte del teu estudi (companys, familiars, etc.).

Un cop tinguis definida la variable i la població, caldrà que recullis dades i les ordenis, construint una taula on hi hagin totes les freqüències estudiades. Agrupa les dades en classes, si treballes amb una variable quantitativa que pren molts valors.

Després, per fer més entenedores les dades, dibuixa un diagrama de sectors circular i un diagrama de barres amb les dades de l'estudi. És possible que a l'hora de determinar el valor dels angles, per poder dibuixar el diagrama de sectors circular, et trobis amb nombres decimals. Si et passa això, mira d'aproximar els valors dels angles a graus sencers, tenint en compte que, en qualsevol cas, han de sumar  $360^\circ$ .

Finalment, fes un comentari i interpretació dels resultats obtinguts.



## Activitat 2

Busca una representació gràfica de les estudiades en aquesta unitat, que il·lustri un estudi estadístic. Pots cercar-la en diaris, revistes o a través d'internet. Si optes per diaris o revistes, és molt útil anar a una biblioteca i fer una ullada a les revistes que tenen. Segur que en trobes, especialment en les de temàtica científica i econòmica. A internet en pots trobar utilitzant un cercador i introduint paraules com gràfica estadística, estudi estadístic, etc.

Fes una còpia de la gràfica, enganxa-la i fes-ne una valoració seguint aquest guió:

- Tipus de representació gràfica
- Claredat i disseny
- Avenença de la gràfica al tipus d'estudi estadístic
- Errades

Per jutjar l'avenença de la gràfica al tipus d'estudi estadístic, tingues en compte si copes fàcilment la informació.

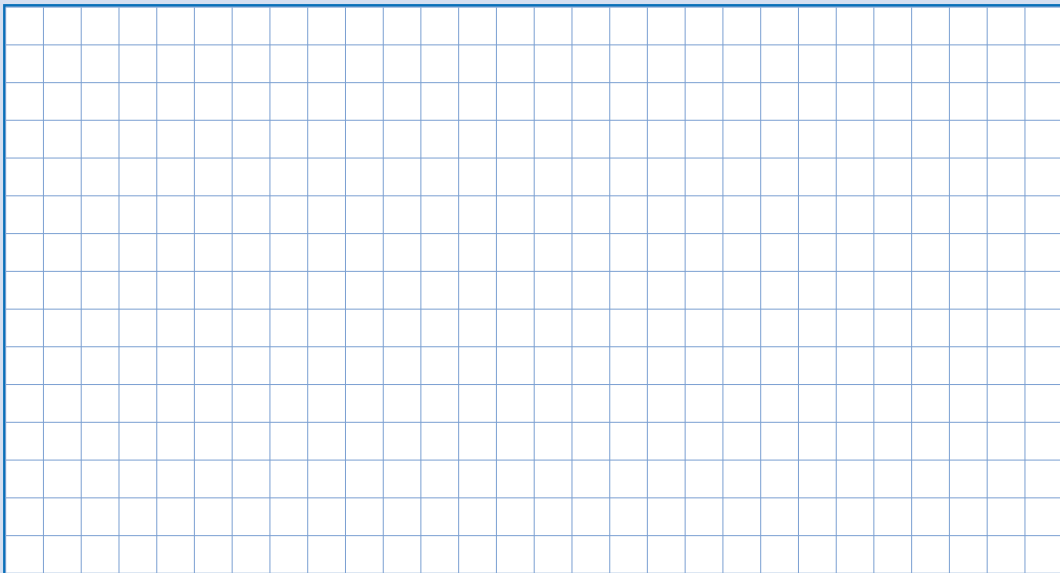
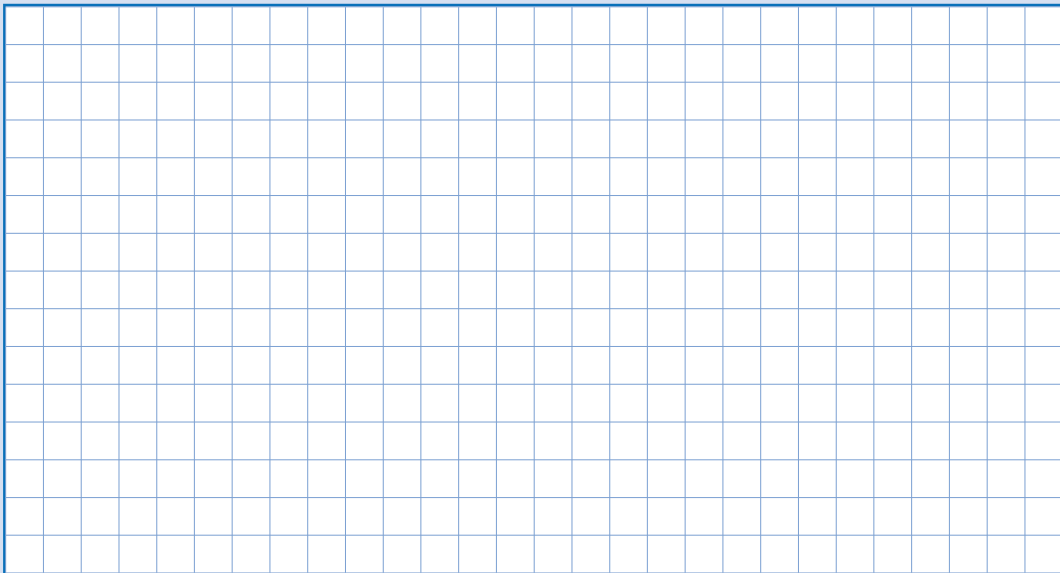
Fixa't bé si hi ha errades. Per fer-ho cal contrastar les freqüències dels valors de la variable amb les representacions corresponents a la gràfica.

**Activitat 1**

Les qualificacions de trigonometria obtingudes per un grup de 18 estudiants són:

9 10 8 7 9 6 9 7 5  
9 3 5 7 8 3 10 3 9

Dibuixa un polígon de freqüències acumulades i un diagrama de sectors circular on constin les freqüències relatives en percentatges.

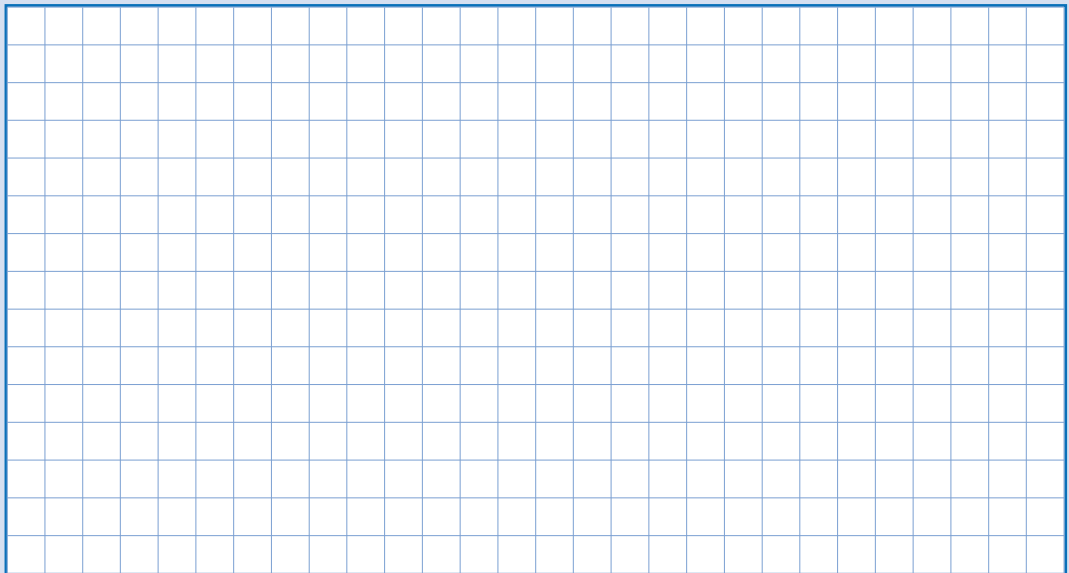
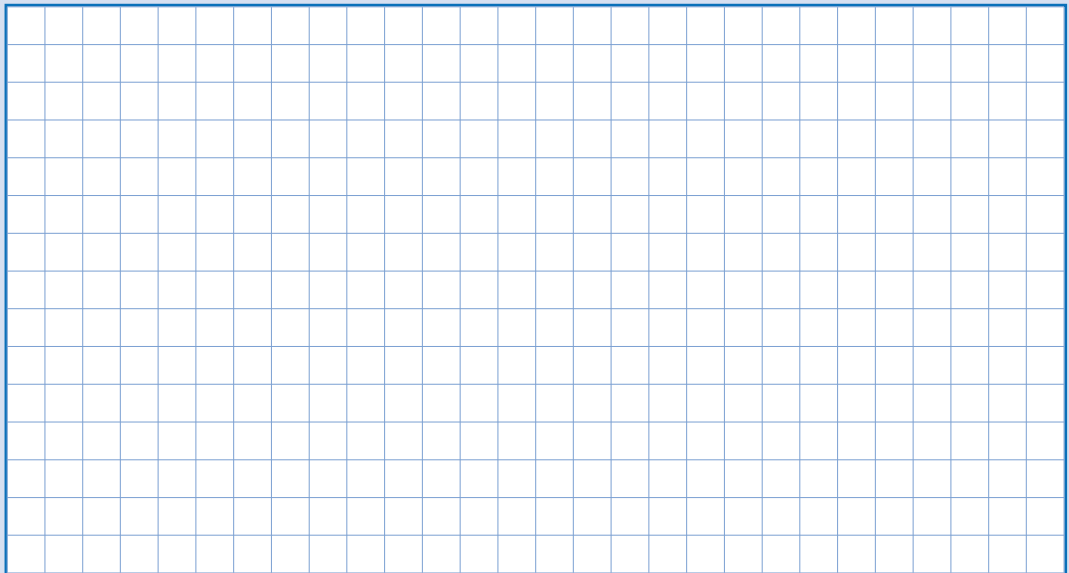


**Activitat 2**

La Montse demana a 40 persones quants gots d'aigua van beure ahir i obté aquests resultats:

7	2	3	3	2	5	4	8	0	0
7	2	4	6	3	3	2	1	3	5
4	6	6	7	3	8	4	5	3	6
1	5	5	2	0	3	0	2	1	4

- A. Construeix una taula de freqüències, agrupant les dades en classes de tres valors de la variable.
- B. A partir de la informació de la taula, dibuixa un pictograma i un histograma amb les marques de classe.



**Activitat 3**

La Sara ha observat el color dels ulls de 45 estudiants de secundària i ha fet aquesta taula de freqüències:

Color d'ulls	Estudiants ( $n_i$ )
Marró	20
Verd	6
Blau	10
Negre	9

Amb aquestes dades, dibuixa un diagrama de sectors semicircular on constin les freqüències absolutes.



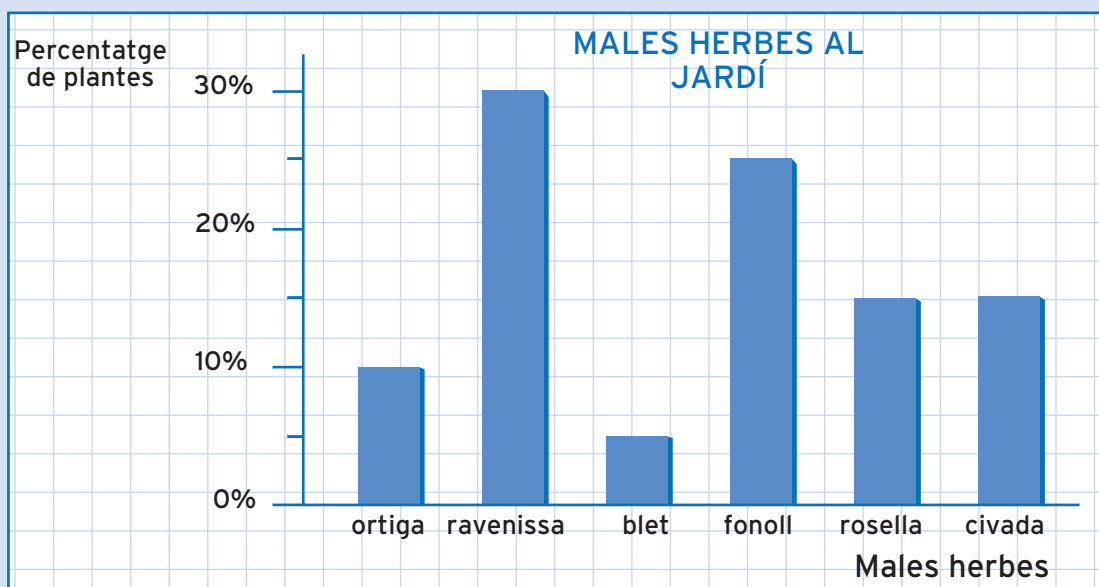
## Activitat 1

En Miquel arrenca totes les males herbes que han envaït la gespa del seu jardí, i, mentre ho fa, va apuntant el nom de cada una de les plantes arrencades:

Ortiga	Ravenissa	Ravenissa	Blet	Fonoll	Ortiga
Ravenissa	Ravenissa	Rosella	Civada	Civada	Fonoll
Civada	Fonoll	Rosella	Ravenissa	Fonoll	Ravenissa
Rosella	Fonoll				

Dibuixa un diagrama de barres amb aquestes dades, utilitzant les freqüències relatives en forma de percentatges.

Males herbes	$n_i$	$f_i$
Ortiga	2	0,1
Ravenissa	6	0,3
Blet	1	0,05
Fonoll	5	0,25
Rosella	3	0,15
Civada	3	0,15
	$N = 20$	Total $f_i = 1$



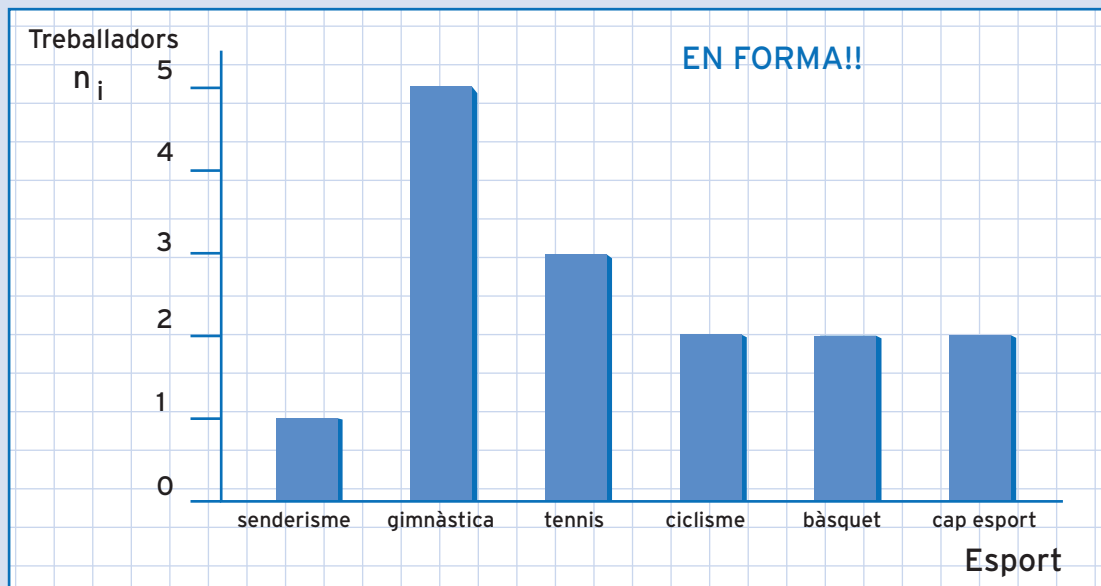
### Activitat 2

Un estudi sobre l'esport que practiquen els treballadors d'una petita empresa dóna aquests resultats:

Senderisme	Gimnàstica	Ciclisme	Ciclisme	Tennis
Bàsquet	Gimnàstica	Gimnàstica	Tennis	Cap esport
Gimnàstica	Ciclisme	Cap esport	Bàsquet	Gimnàstica

Dibuixa un diagrama de barres amb aquestes dades.

Esport	$n_i$
Senderisme	1
Gimnàstica	5
Ciclisme	3
Tennis	2
Bàsquet	2
Cap esport	2
N = 15	



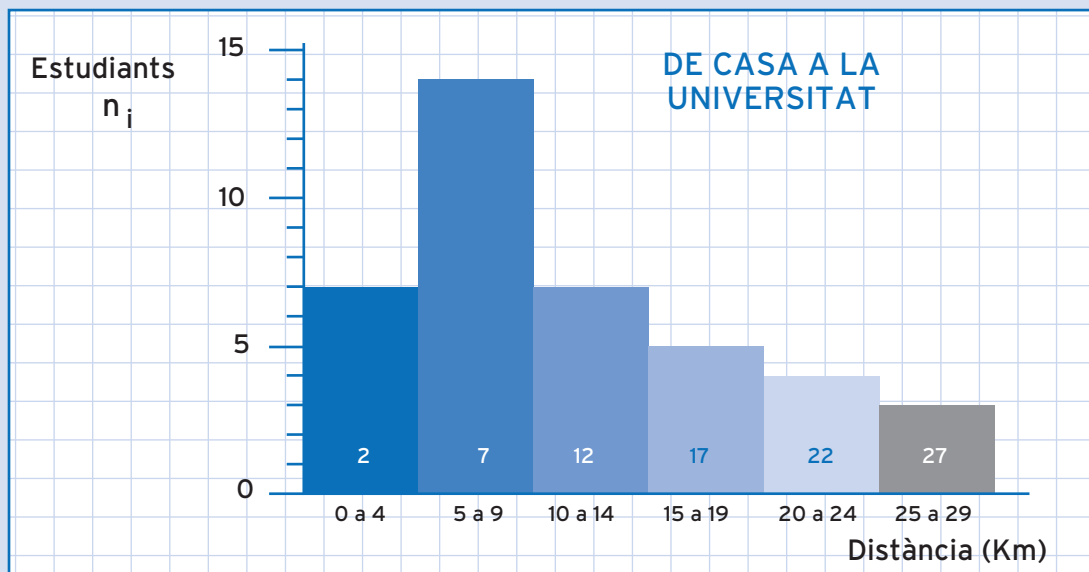
**Activitat 4**

Es va fer un estudi sobre la distància que els estudiants d'un grup de segon curs de filosofia havien de recórrer per arribar a la Universitat. Heus aquí les dades que mostren la distància en quilòmetres que recorria cada un dels estudiants:

2 1 8 23 19 0 21 2 27 5  
 6 20 18 14 14 19 5 3 10 9  
 9 8 7 3 6 27 13 10 19 12  
 3 21 25 15 5 8 14 6 6 9

Construeix una taula de freqüències, agrupant les distàncies en classes de 5 valors, i dibuixa un histograma, assenyalant les marques de classe.

Distància (km)	Marca de classe	$n_i$
0 a 4	2	7
5 a 9	7	14
10 a 14	12	7
15 a 19	17	5
20 a 24	22	4
25 a 29	27	3
		N = 40



**Activitat 4**

El diagrama de barres adjunt mostra les causes de mortalitat a l'Àfrica central. L'estudi es va fer amb una mostra de 10.000 casos.

A partir de la informació que dóna aquest diagrama de barres, construeix una taula de freqüències amb totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

A partir dels percentatges podem obtenir les freqüències relatives, dividint cada un dels valors entre cent. Per obtenir les freqüències absolutes només cal multiplicar les freqüències relatives per 10.000, que és el nombre total de casos estudiats.

<b>Causes de mortalitat</b>	$n_i$	$N_i$	$f_i$	$F_i$	<b>Percentatge</b>
Infeccions	4.000	4.000	0,4	0,4	40 %
Càncer	500	4.500	0,05	0,45	5 %
Malalties cardiovasculars	1.500	6.000	0,15	0,6	15 %
Accidents	500	6.500	0,05	0,65	5 %
Part	1.000	7.500	0,1	0,75	10 %
Altres	2.500	10.000	0,25	1	25 %
	$n = 10.000$		Total $f_i = 1$		



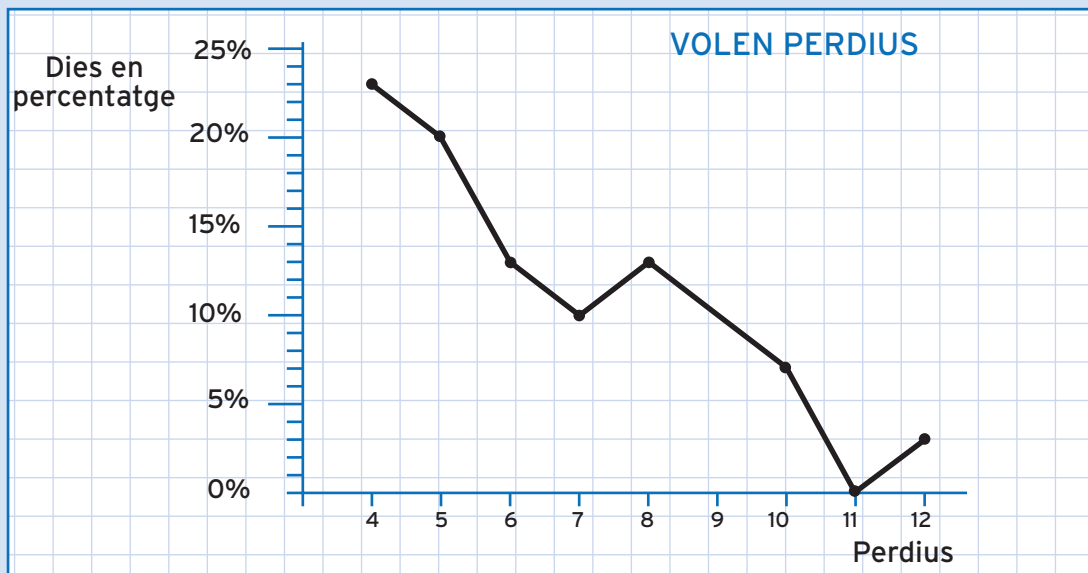
**Activitat 5**

Durant un mes de primavera, l'Adela, una noia que viu a pagès, compta el nombre de perdius que veu caminant des de casa fins al poble, i recull aquestes dades:

5 4 6 7 8 6 5 4 7 9  
 8 4 8 5 12 10 9 6 5 4  
 4 4 4 7 10 5 8 9 5 6

Amb aquestes dades, dibuixa un polígon de freqüències relatives en percentatges.

Perdius	$n_i$	$f_i$
4	7	0,23
5	6	0,2
6	4	0,13
7	3	0,1
8	4	0,13
9	3	0,1
10	2	0,07
11	0	0
12	1	0,03
N = 30		Total $f_i = 0,99$



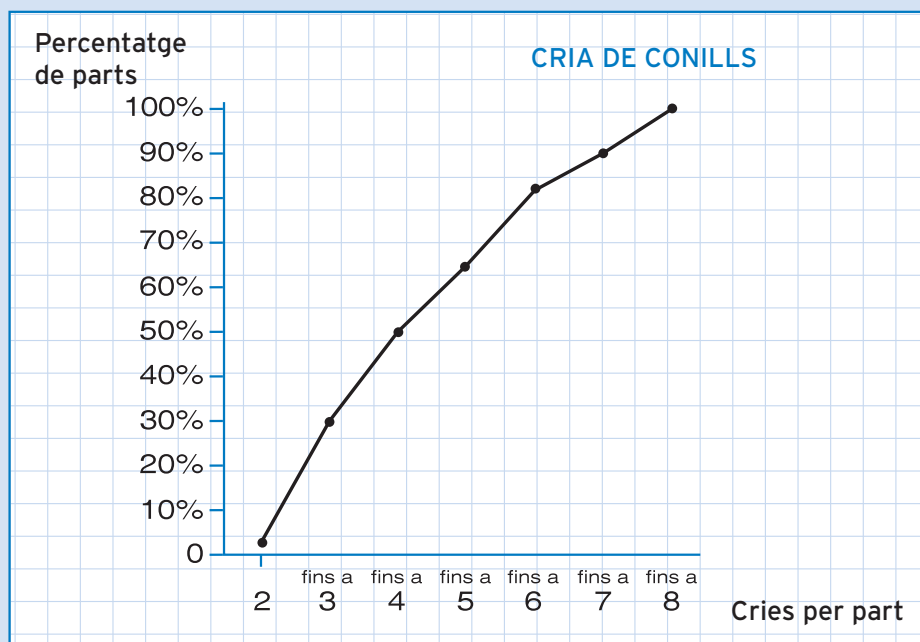
## Activitat 6

En Rossend estudia el nombre de cries de conill per cada part, en una granja de conills, i recull aquestes dades:

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

A partir d'aquestes dades, dibuixa un polígon de freqüències relatives acumulades en percentatges.

Cries	$n_i$	$N_i$	$f_i$	$F_i$
2	1	1	0,025	0,025
3	11	12	0,275	0,3
4	8	20	0,2	0,5
5	6	26	0,15	0,65
6	7	33	0,175	0,825
7	3	36	0,075	0,9
8	4	40	0,1	1
		$N = 40$	Total $f_i = 1$	



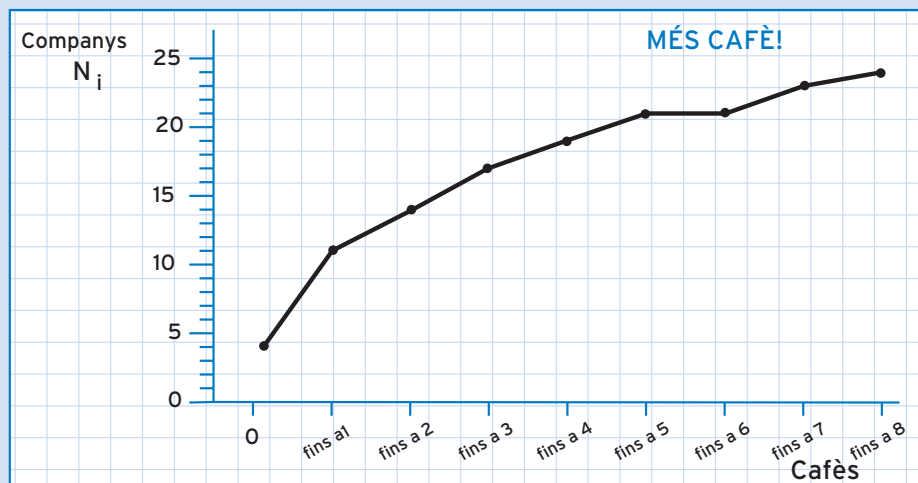
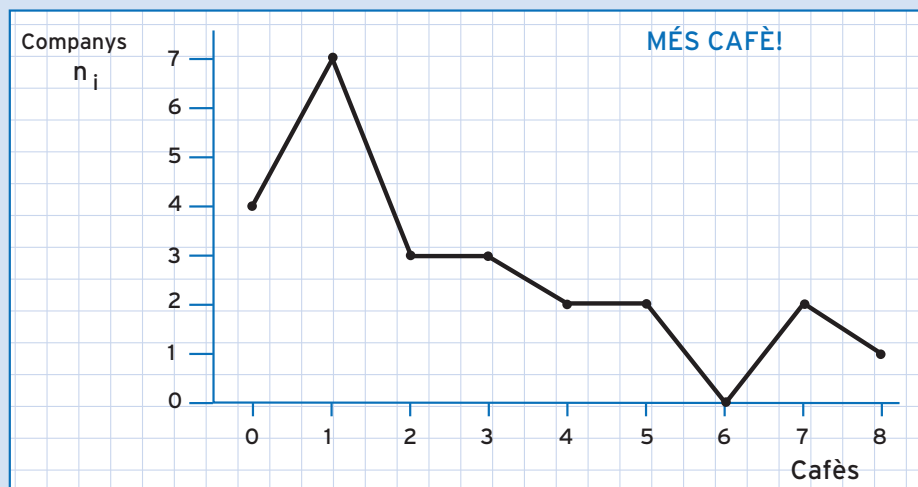
**Activitat 7**

La Laura demana als companys de classe quants cafès van prendre ahir. Heus aquí els resultats:

3	0	5	7	3	2	1	1	1	0
2	1	1	2	4	0	5	8	7	4
3	0	1	1						

Amb aquestes dades, dibuixa un polígon de freqüències i un polígon de freqüències acumulades.

Nombre de cafès	$n_i$	$N_i$
0	4	4
1	7	11
2	3	14
3	3	17
4	2	19
5	2	21
6	0	21
7	2	23
8	1	24
$N = 24$		



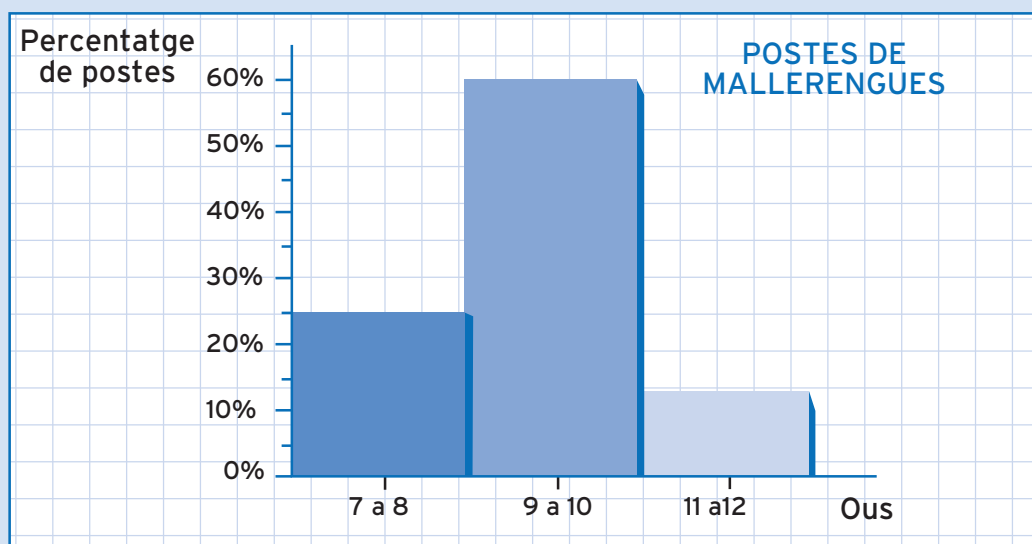
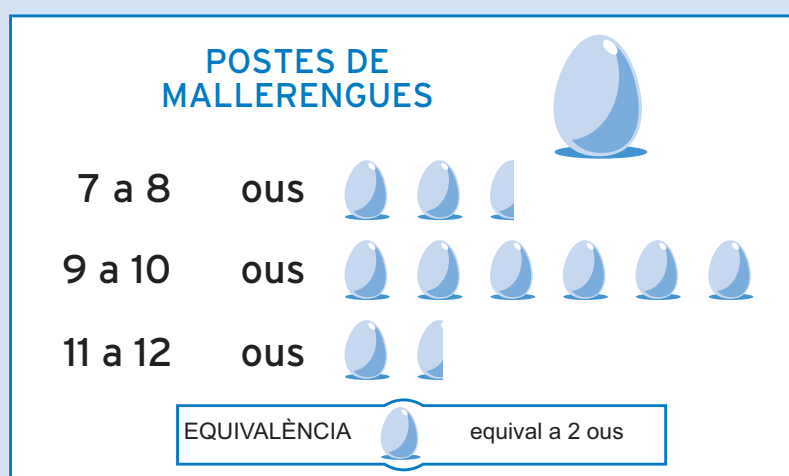
## Activitat 8

Un investigador recull dades sobre el nombre d'ous de les postes de vint mallarengues. Heus aquí els resultats:

8 10 10 9 8 9 10 11 8 10  
9 10 12 9 8 10 10 9 11 7

Amb aquestes dades, agrupades en classes de dos valors de la variable, dibuixa un pictograma i un histograma amb les freqüències relatives en percentatges.

Ous	$n_i$	$N_i$	$f_i$	$F_i$
7 - 8	5	5	0,25	0,25
9 - 10	12	17	0,6	0,85
11 - 12	3	20	0,15	1
	N = 20		Total $f_i = 1$	



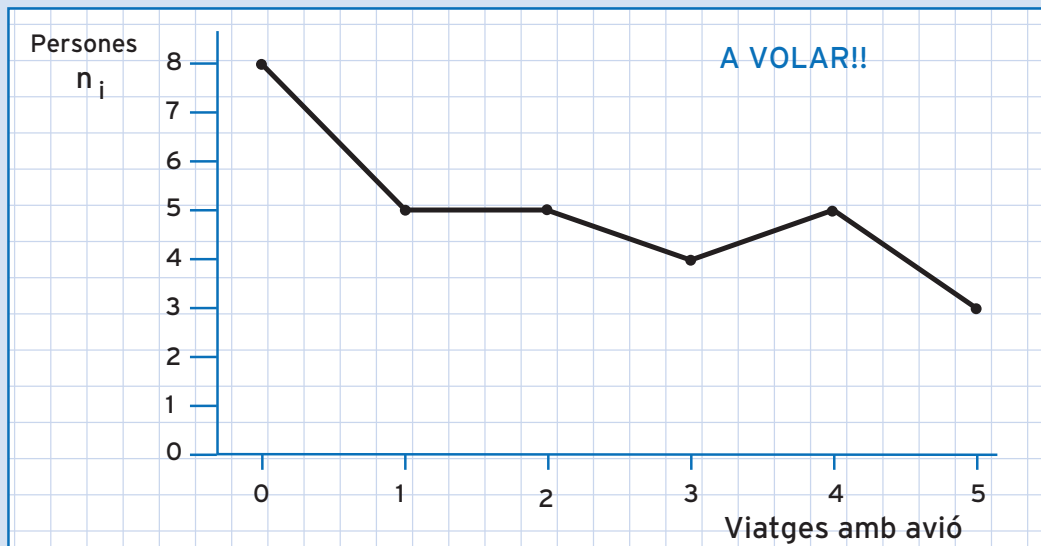
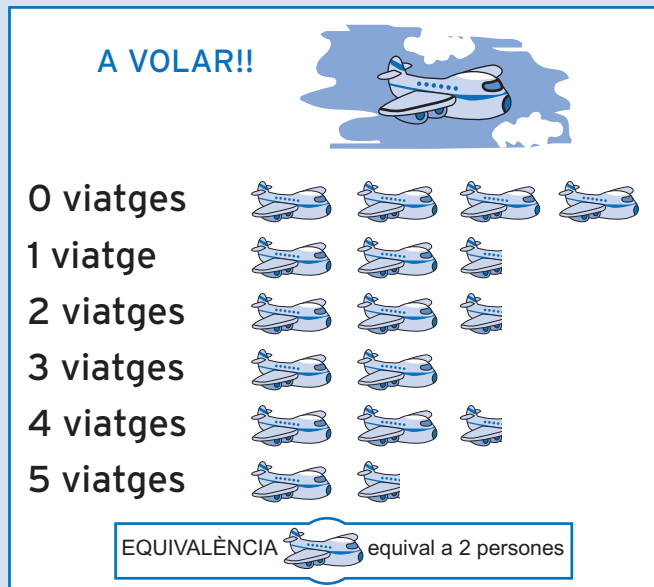
**Activitat 9**

La Maria demana a 30 persones quants viatges amb avió van fer l'any passat. Aquests són els resultats:

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

Dibuixa un pictograma i un polígon de freqüències amb les dades obtingudes.

Viatges amb avió	$n_i$
0	8
1	5
2	5
3	4
4	5
5	3
N = 30	



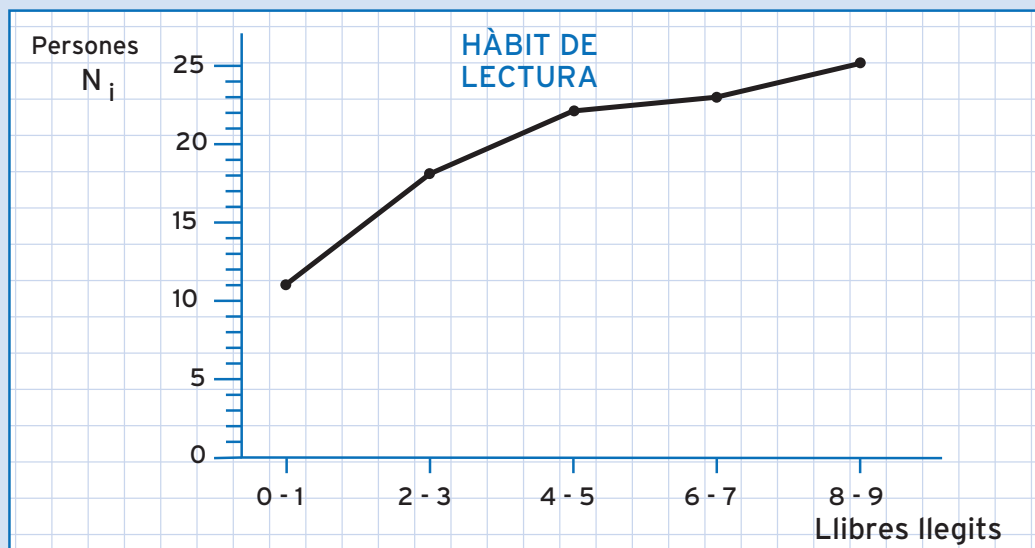
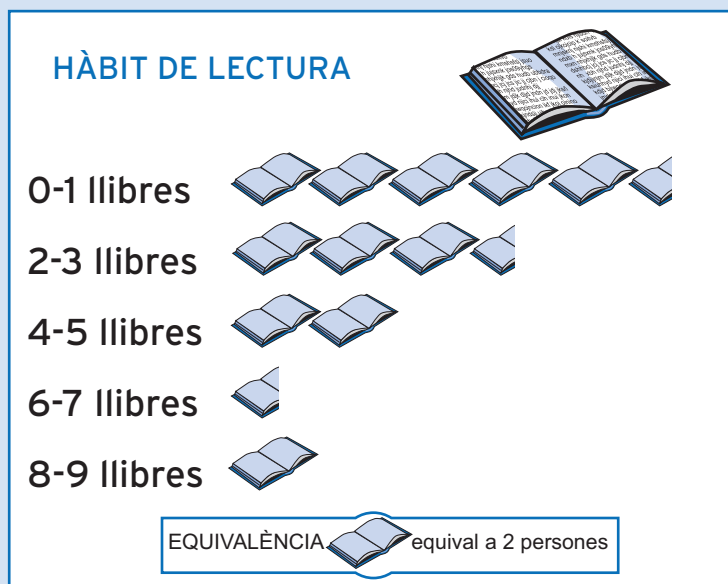
## Activitat 10

En Joan demana a 25 persones quants llibres van llegir l'any passat. Heus aquí els resultats:

2	1	2	1	1	8	1	0	5	3
3	4	4	2	1	0	0	1	4	1
6	3	9	1	2					

Agrupa les dades obtingudes en classes de dos valors i dibuixa un pictograma i un polígon de freqüències acumulades.

Llibres llegits	$n_i$	$N_i$
0-1	11	11
2-3	7	18
4-5	4	22
6-7	1	23
8-9	2	25
N = 25		



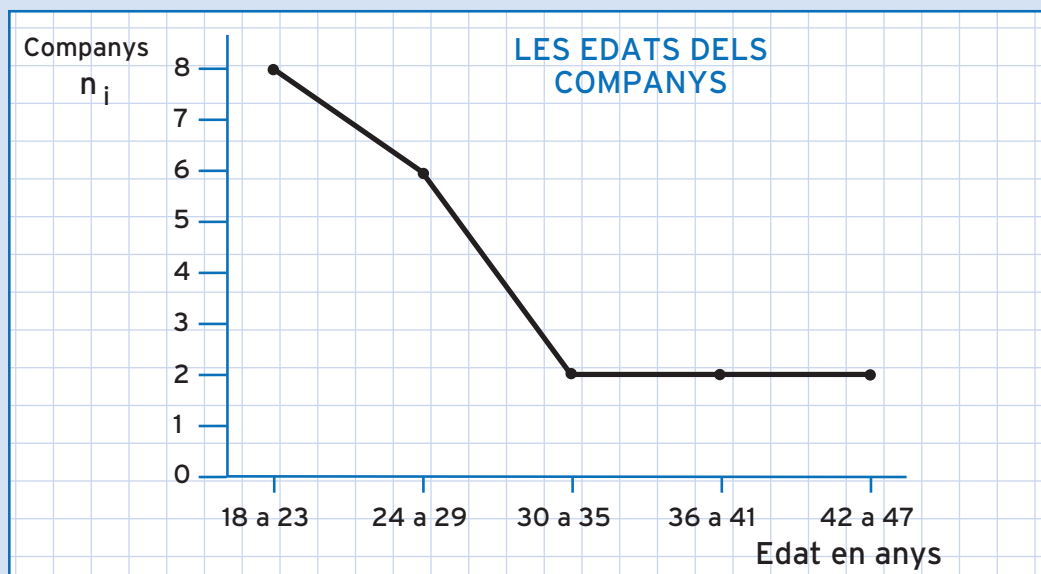
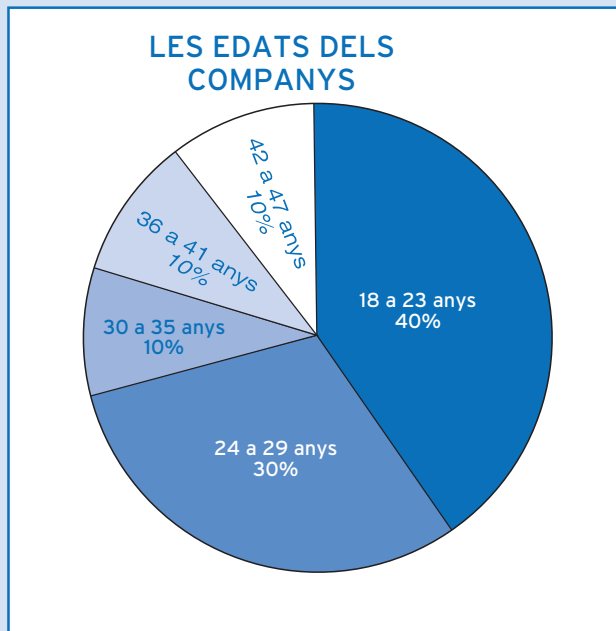
**Activitat 11**

En Francesc demana als seus companys de classe quina edat tenen. Heus aquí els resultats en anys:

18 24 28 18 19 21 24 23 28 45  
29 33 46 20 23 27 21 35 37 40

Agrupa les dades obtingudes en classes de sis valors i dibuixa un diagrama de sectors circular (on constin les freqüències relatives en forma de percentatge dins de cada sector) i un polígon de freqüències.

Edat en anys	$n_i$	$f_i$	Angle
18 a 23	8	0,4	$8 \times 360^\circ : 20 = 144^\circ$
24 a 29	6	0,3	$6 \times 360^\circ : 20 = 108^\circ$
30 a 35	2	0,1	$2 \times 360^\circ : 20 = 36^\circ$
36 a 41	2	0,1	$2 \times 360^\circ : 20 = 36^\circ$
42 a 47	2	0,1	$2 \times 360^\circ : 20 = 36^\circ$
N = 20		Total $f_i = 1$	



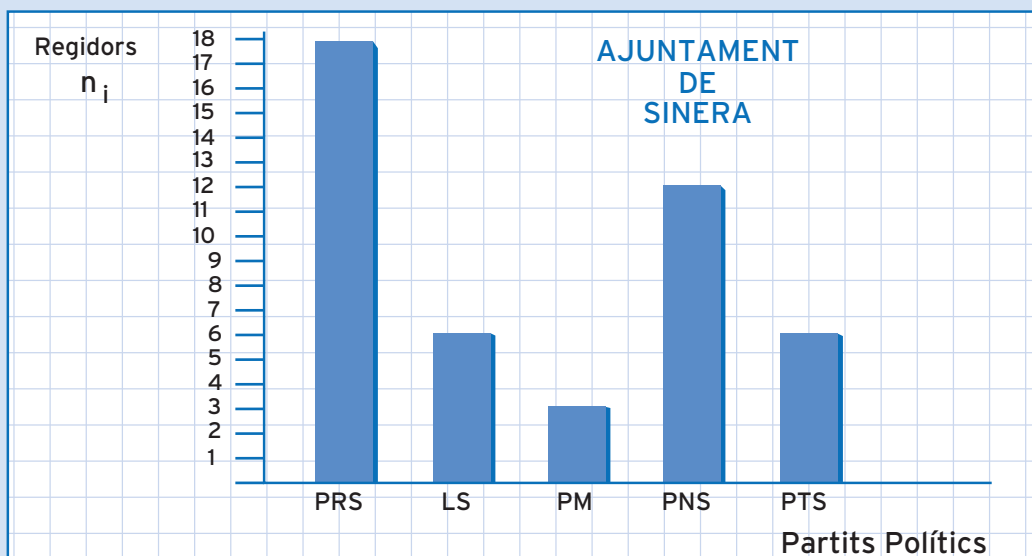
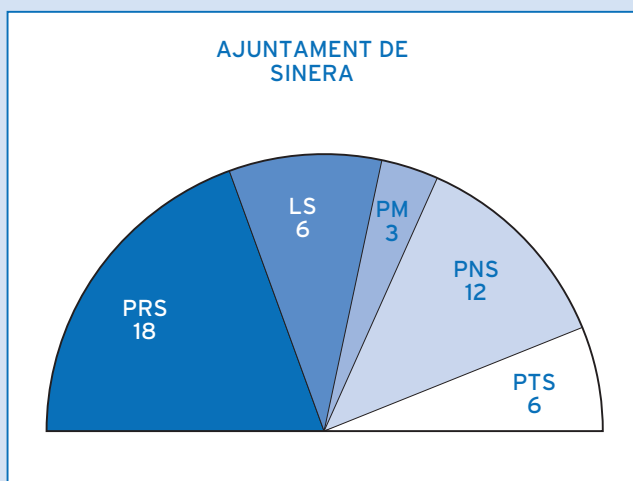
## Activitat 12

A les darreres eleccions locals de Sinera, el nombre de regidors obtinguts per les diverses forces polítiques fou:

Partits polítics	Regidors
Partit Republicà de Sinera (PRS)	18
Liberals per Sinera (LS)	6
Partit Monàrquic (PM)	3
Partit Nacional Sinerenc (PNS)	12
Partit dels Treballadors de Sinera (PTS)	6

Amb aquestes dades, dibuixa un diagrama de sectors semicircular (on constin les freqüències absolutes) i un diagrama de barres.

Partits polítics	Regidors	Angle
Partit Republicà de Sinera (PRS)	18	$18 \times 180^\circ : 45 = 72^\circ$
Liberals per Sinera (LS)	6	$6 \times 180^\circ : 45 = 24^\circ$
Partit Monàrquic (PM)	3	$3 \times 180^\circ : 45 = 12^\circ$
Partit Nacional Sinerenc (PNS)	12	$12 \times 180^\circ : 45 = 48^\circ$
Partit dels Treballadors de Sinera (PTS)	6	$6 \times 180^\circ : 45 = 24^\circ$





**Activitat 13**

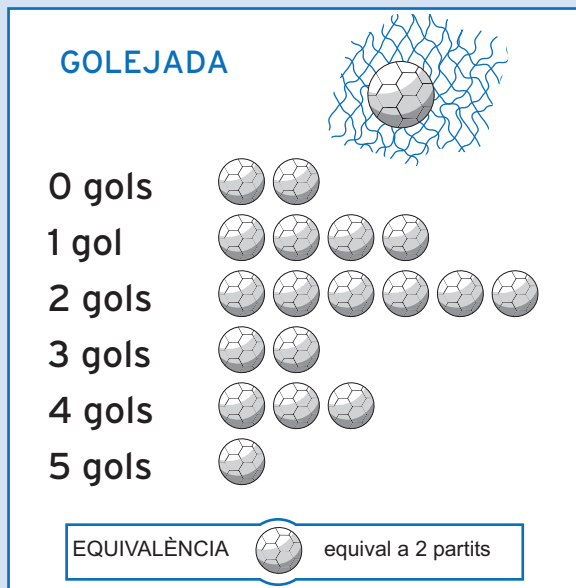
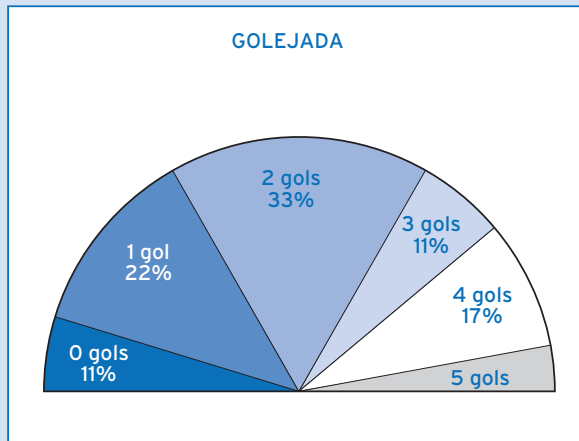
El nombre de gols aconseguït pel F.C. Sinera durant els partits jugats a la darrera lliga nacional foren:

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

Amb aquestes dades, dibuixa un diagrama de sectors semicircular (on constin les freqüències relatives en forma de percentatges) i un pictograma.

Gols marcats	$n_i$	$f_i$	Angle
0	4	0,11	$4 \times 180^\circ : 36 = 20^\circ$
1	8	0,22	$8 \times 180^\circ : 36 = 40^\circ$
2	12	0,33	$12 \times 180^\circ : 36 = 60^\circ$
3	4	0,11	$4 \times 180^\circ : 36 = 20^\circ$
4	6	0,17	$6 \times 180^\circ : 36 = 30^\circ$
5	2	0,056	$2 \times 180^\circ : 36 = 10^\circ$
	N = 36	Total $f_i = 0,996$	

Per tal que la suma dels percentatges de tots els sectors sigui 100 %, ajustarem la  $f_i$  del valor de la variable 5 gols al 6 %.



## Activitat 14

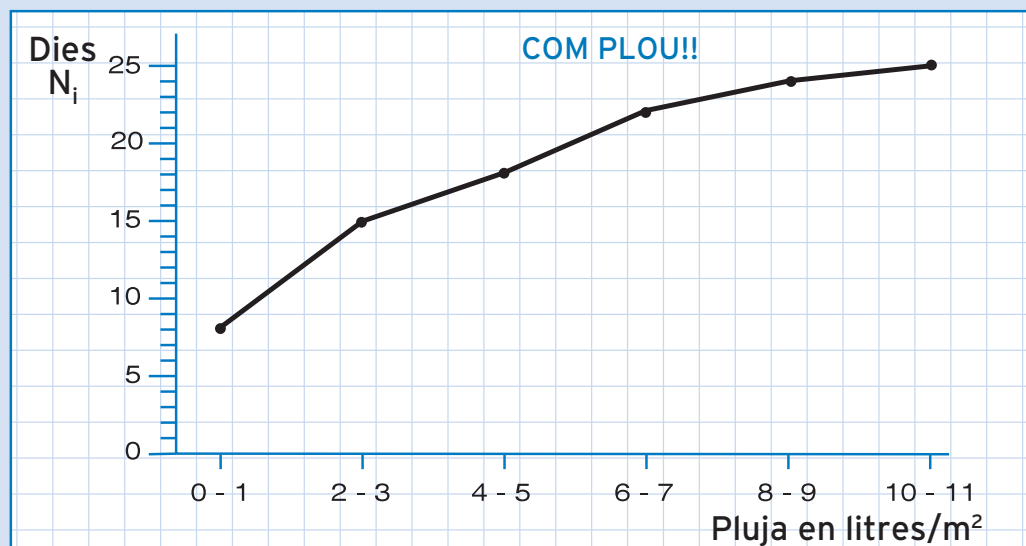
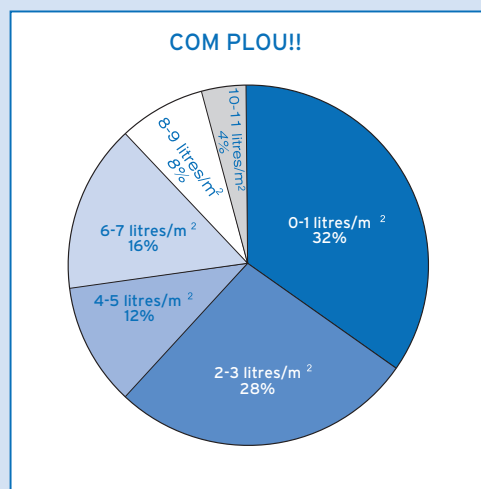
L'Annabel té una petita estació meteorològica on ha recollit dades sobre quantitat de pluja durant els vint-i-cinc primers dies d'octubre. Heus aquí les dades recollides, en litres/m<sup>2</sup>:

1	0	0	1	6	3	8	5	3	9
3	11	3	4	6	2	1	1	1	0
7	6	2	4	3					

Agrupar les dades en classes de dos valors i dibuixa:

- A. Un diagrama de sectors circular on consti la freqüència relativa, en forma de percentatge, dins de cada sector.
- B. Un polígon de freqüències acumulades.

Pluja en litres/m <sup>2</sup>	n <sub>i</sub>	N <sub>i</sub>	f <sub>i</sub>	Angle
0-1	8	8	0,32	8 x 360° : 25 = 115°
2-3	7	15	0,28	7 x 360° : 25 = 101°
4-5	3	18	0,12	3 x 360° : 25 = 43°
6-7	4	22	0,16	4 x 360° : 25 = 58°
8-9	2	24	0,08	2 x 360° : 25 = 29°
10-11	1	25	0,04	1 x 360° : 25 = 14°
		N = 25	Total f <sub>i</sub> = 1	



**Activitat 15**

El diagrama de sectors adjunt representa, en percentatges, els usos del territori a la comarca del Bages.

A partir de la informació que dóna aquest diagrama de sectors, i tenint en compte que la superfície de la comarca és de 129.800 hectàrees, construeix una taula de freqüències amb totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .

A partir dels percentatges podem obtenir les freqüències relatives, dividint cada un dels valors entre cent. Per obtenir les freqüències absolutes (el nombre d'hectàrees destinades a cada ús) només cal multiplicar les freqüències relatives per 129.800, que és el nombre total d'hectàrees.

Ús del territori	$n_i$	$N_i$	$f_i$	$F_i$	Percentatge
Forestal arbrat	61.655	61.655	0,475	0,475	47,5 %
Forestal no arbrat	28.556	90.211	0,22	0,695	22 %
Agrícola	32.450	122.661	0,25	0,945	25 %
Urbà	6.490	129.151	0,05	0,995	5 %
Rius i llacs	649	129.800	0,005	1	0,5 %
	N = 129.800		Total $f_i = 1$		

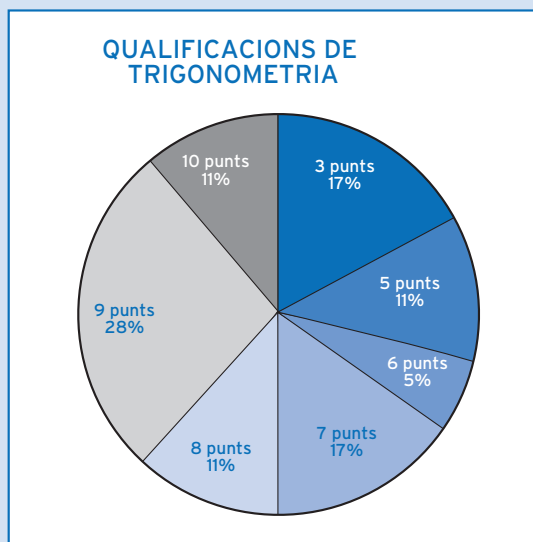
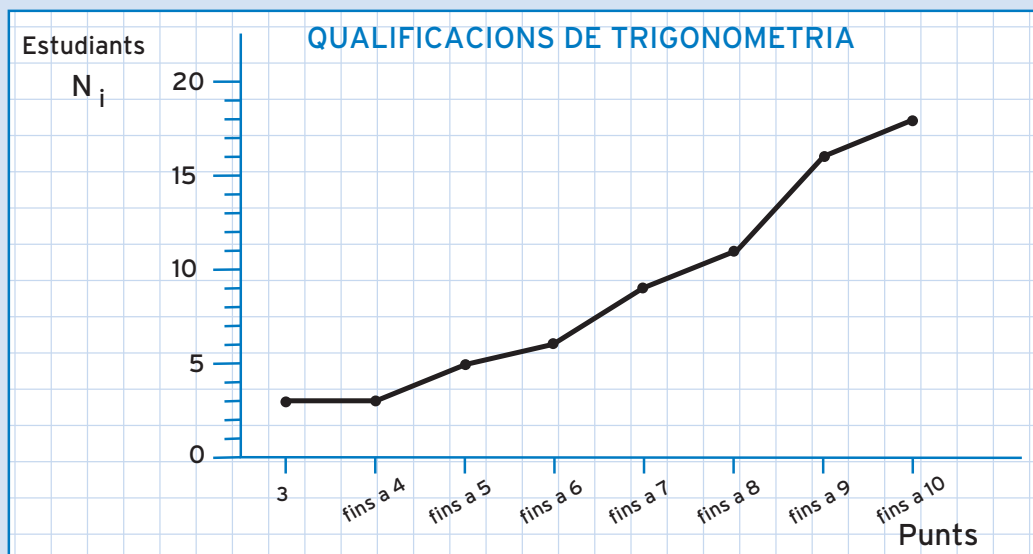
Activitat 1

Les qualificacions de trigonometria obtingudes per un grup de 18 estudiants són:

9 10 8 7 9 6 9 7 5  
 9 3 5 7 8 3 10 3 9

Dibuixa un polígon de freqüències acumulades i un diagrama de sectors circular on constin les freqüències relatives en percentatges.

Qualificacions	Estudiants ( $n_i$ )	$N_i$	$f_i$	Angle
3	3	3	0,17	$3 \times 360^\circ : 18 = 60^\circ$
4	0	3	0	$0 \times 360^\circ : 18 = 0^\circ$
5	2	5	0,11	$2 \times 360^\circ : 18 = 40^\circ$
6	1	6	0,056	$1 \times 360^\circ : 18 = 20^\circ$
7	3	9	0,17	$3 \times 360^\circ : 18 = 60^\circ$
8	2	11	0,11	$2 \times 360^\circ : 18 = 40^\circ$
9	5	16	0,28	$5 \times 360^\circ : 18 = 100^\circ$
10	2	18	0,11	$2 \times 360^\circ : 18 = 40^\circ$
		N = 18	Total $f_i = 1,006$	



Per tal que la suma dels percentatges de tots els sectors sigui 100 %, ajustarem la  $f_i$  del valor de la variable 6 punts al 5 %. Una altra opció hagués estat agafar totes les  $f_i$  amb tres decimals i, per tant, els percentatges amb un decimal. D'aquesta manera haguéssim minimitzat l'error.

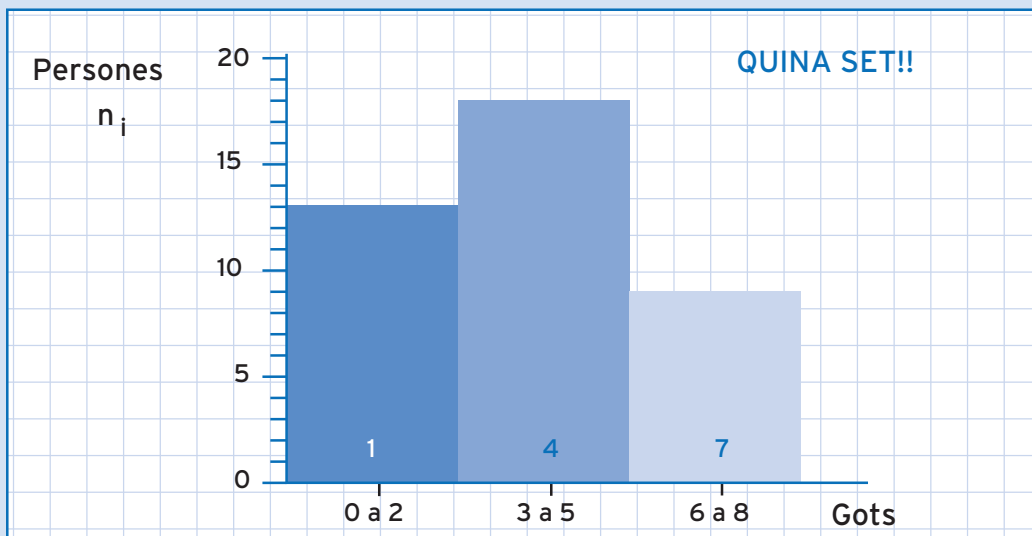
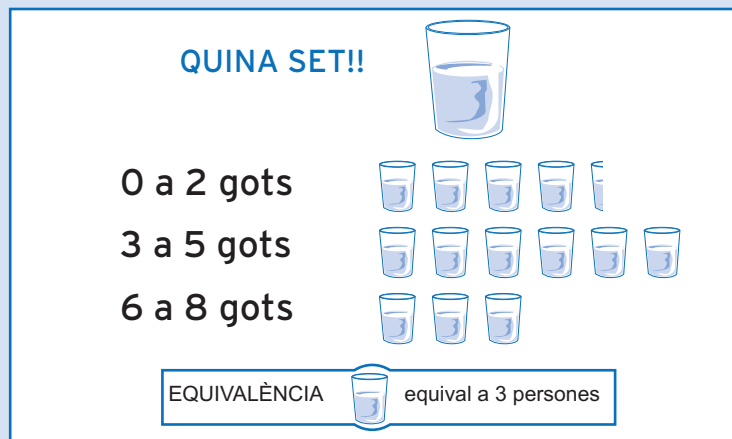
### Activitat 2

La Montse demana a 40 persones quants gots d'aigua van beure ahir i obté aquests resultats:

7 2 3 3 2 5 4 8 0 0  
 7 2 4 6 3 3 2 1 3 5  
 4 6 6 7 3 8 4 5 3 6  
 1 5 5 2 0 3 0 2 1 4

- A. Construeix una taula de freqüències, agrupant les dades en classes de tres valors de la variable.
- B. A partir de la informació de la taula, dibuixa un pictograma i un histograma amb les marques de classe.

Nombre de gots	Marca de classe	$n_i$
0 a 2	1	13
3 a 5	4	18
6 a 8	7	9
		N = 40



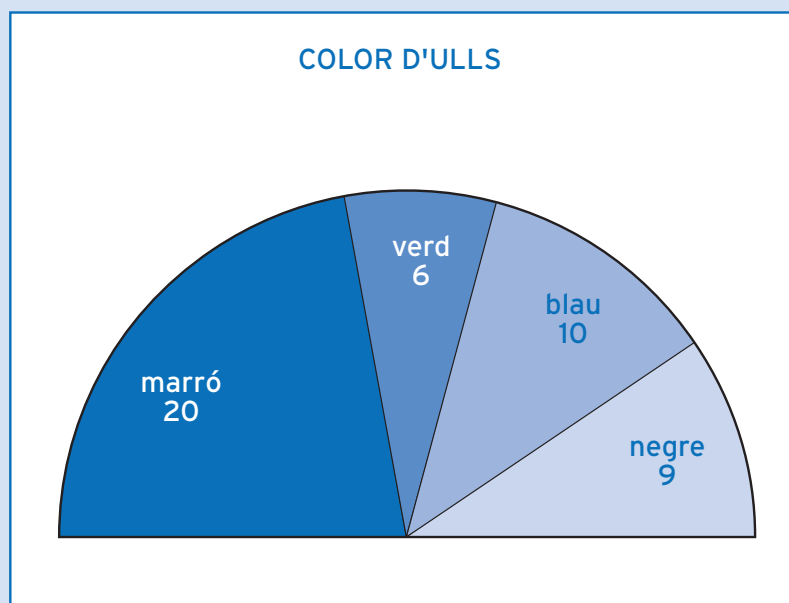
## Activitat 3

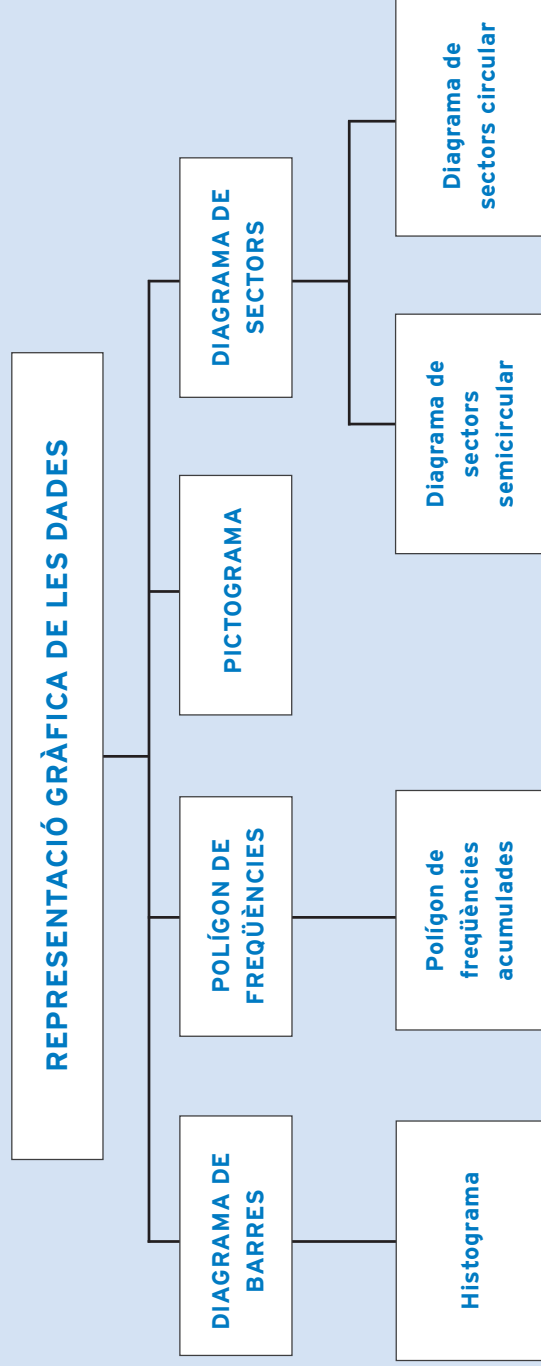
La Sara ha observat el color dels ulls de 45 estudiants de secundària i ha fet aquesta taula de freqüències:

Color d'ulls	Estudiants ( $n_i$ )
Marró	20
Verd	6
Blau	10
Negre	9

Amb aquestes dades, dibuixa un diagrama de sectors semicircular on constin les freqüències absolutes.

Color d'ulls	Estudiants ( $n_i$ )	Angle
Marró	20	$20 \times 180^\circ : 45 = 80^\circ$
Verd	6	$6 \times 180^\circ : 45 = 24^\circ$
Blau	10	$10 \times 180^\circ : 45 = 40^\circ$
Negre	9	$9 \times 180^\circ : 45 = 36^\circ$
	N = 45	





# com ho porto?

Omple la graella següent posant una creu on correspongui:

En acabar la unitat sóc capaç de:	Bé	Regular	Malament
Elaborar un diagrama de barres a partir de la informació continguda en un recull de dades.			
Elaborar un polígon de freqüències a partir de la informació continguda en un recull de dades.			
Elaborar un pictograma a partir de la informació continguda en un recull de dades.			
Elaborar un diagrama de sectors a partir de la informació continguda en un recull de dades.			
Descriure la informació reflectida en les representacions gràfiques.			



# Unitat 3

## TERMES MITJANS

91

UNITAT 3 TERMES MITJANS

UNITAT 3

19. ESTADÍSTICA

Matemàtiques, Ciència i Tecnologia

# què treballaràs?

En acabar aquesta unitat has de ser capaç de:

- Explicar el significat dels diferents tipus de valors numèrics que permeten resumir el comportament d'un conjunt de dades quantitatives.
- Calcular els valors numèrics que permeten resumir el comportament d'un conjunt de dades quantitatives.
- Comparar les dades estadístiques corresponents a diferents estudis.
- Interpretar la informació que ens aporten els termes mitjans.

Els termes mitjans ens permeten resumir en uns quants valors numèrics el conjunt de dades amb les quals fem l'estudi estadístic.

Hi ha dos tipus de termes mitjans: les **mesures de centralització** i les **mesures de dispersió**.

Les **mesures de centralització** ens indiquen quins són els punts centrals de la distribució de dades, és a dir, cap a quins valors de la variable s'agrupen majoritàriament les dades.

Les **mesures de dispersió** ens indiquen com d'escampades estan les dades.

Per entendre millor el significat i el càlcul dels diferents termes mitjans, treballarem amb exemples. Exposarem dos exemples per a cada terme mitjà. El primer exemple constarà d'un recull petit de dades. Aquest exemple anirà canviant segons el terme mitjà que s'està estudiant. En el segon exemple hi haurà un recull de dades molt més gran, la qual cosa comportarà treballar amb taules de freqüències per tal d'evitar errades. Aquest segon exemple serà sempre el mateix, per tal de poder comparar els valors dels diferents termes mitjans.

## 1. Mesures de centralització

Hi ha tres tipus de mesures de centralització: la **mitjana**, la **mediana** i el **mode**.

### La mitjana

La **mitjana** s'obté dividint la suma de les dades entre el nombre total de dades.

El símbol de la mitjana és  $\bar{X}$ .

#### Exemple 1

L'estiu passat cinc, amics van recaptar un total de 60 € per fer un regal.

Ens podem fer una idea aproximada del que va aportar cada una de les persones si dividim el total de 60 € entre 5. Així s'obté **la mitjana**.

<b>Mitjana de diners recollits</b>
$60 \text{ €} : 5 = 12 \text{ €}$

En altres paraules, cada persona va recollir, de mitjana, 12 €.

En realitat, les quantitats aportades per cada un dels amics van ésser:

Jordi 10 €    Eva 6 €    Joana 16 €    Mònica 13 €    Pere 15 €

Observa que, en realitat, cap d'ells no va aportar 12 €. De fet, no cal que la mitjana coincideixi amb un dels valors utilitzats per calcular-la.

#### Exemple 2

En aquest exemple, aprendràs a trobar la mitjana a partir d'una taula de freqüències. Utilitzarem una taula perquè treballarem amb moltes dades, la qual cosa fa que convingui ordenar-les en una taula per tal d'evitar errades.

Demaneu als vint-i-cinc estudiants d'una classe quants parells de sabates tenen. Heus aquí els resultats:

2 3 6 10 3 6 3 5 4 4  
 2 1 6 3 5 11 5 7 10 4  
 3 5 3 4 10

Fem una taula de freqüències per tal d'ordenar les dades recollides:

Parells de sabates	$n_i$
1	1
2	2
3	6
4	4
5	4
6	3
7	1
8	0
9	0
10	3
11	1
N = 25	

Ara, per tal de trobar la mitjana, cal que segueixis els passos següents:

- Multiplica cada un dels valors de la variable per la seva freqüència absoluta. Per estalviar-te feina, no cal que multipliquis els valors de la variable que no recullen cap dada, és a dir, amb  $n_i = 0$ , ja que el resultat de la multiplicació serà zero, la qual cosa farà que no afectin al resultat final.
- Suma tots els resultats obtinguts.
- Divideix la suma total entre el nombre de dades (N).

Fem-ho:

$$\bar{X} = \frac{(1 \times 1) + (2 \times 2) + (3 \times 6) + (4 \times 4) + (5 \times 4) + (6 \times 3) + (7 \times 1) + (10 \times 3) + (11 \times 1)}{25}$$

$$\bar{X} = \frac{1 + 4 + 18 + 16 + 20 + 18 + 7 + 30 + 11}{25} = \frac{125}{25} = 5 \text{ parells}$$

- **Fes les activitats d'aprenentatge 1 i 2**

## La mediana

La **mediana** és la dada que ocupa la posició central quan les dades estan disposades en ordre.

El símbol de la mediana és **Me**.

### Exemple 1

La Sara va fer cinc exàmens de matemàtiques durant el trimestre, i va treure aquestes notes sobre 10 punts:

9      8      1      8      9

La Sara no va fer bé el tercer examen perquè estava malalta.

El seu professor va calcular la nota mitjana dels exàmens fets per la Sara:

$$\bar{X} = \frac{9+8+1+8+9}{5} = \frac{35}{5} = 7 \text{ punts}$$

Aquesta desició no li va semblar justa a la Sara, perquè havia tret més de 7 punts en quatre dels exàmens.

El professor va resoldre el problema utilitzant la mediana de les notes com a terme mitjà.

Per fer-ho, va ordenar les notes des de la més petita fins a la més gran:

1      8      **8**      9      9

La nota que ocupa la posició central és 8 ( $Me = 8$ ). És el que anomenem mediana. El professor va agafar la mediana per posar la nota final del trimestre i aquest fet li va semblar molt bé a la Sara.

### Exemple 2

Ara utilitzarem un altre cop l'exemple del nombre de parells de sabates que tenen els integrants d'un grup d'estudiants, per tal d'aprendre com calcular la mediana a partir d'una taula de freqüències.

Parells de sabates	$n_i$	$N_i$
1	1	1
2	2	3
3	6	9
<b>4</b>	4	13
5	4	17
6	3	20
7	1	21
8	0	21
9	0	21
10	3	24
11	1	25
	<b>N = 25</b>	

La taula de freqüències mostra les dades ordenades, des de la més petita fins a la més gran.

En aquest exemple el nombre de dades recollides és de 25 ( $N=25$ ). Per tant, la posició central està ocupada per la tretzena dada (ja que en té dotze al davant i dotze al darrere), que correspon al valor de la variable 4 parells de sabates.

Així doncs,  $Me = 4$  parells de sabates.

Quan el nombre de dades és parell, per obtenir la mediana hem de calcular la mitjana de les dues dades centrals. Agafem un altre cop l'exemple de les notes de matemàtiques de la Sara, però considerant que va fer sis exàmens durant el trimestre amb aquestes notes:

9 8 1 8 9 9

Primer ordenem les notes des de la més petita fins a la més gran:

1 8 **8** **9** 9 9

Després calculem la mitjana de les dues dades centrals:

$$Me = \frac{8+9}{2} = 8,5 \text{ punts}$$

Si treballem amb una taula i un nombre parell de dades, la situació és similar. Si tornem al segon exemple, però hi afegim una dada més ( $N = 26$ ):

Parells de sabates	$n_i$	$N_i$
1	1	1
2	2	3
3	6	9
4	4	13
5	5	18
6	3	21
7	1	22
8	0	22
9	0	22
10	3	25
11	1	26
		$N = 26$

Atès que el nombre de dades és de 26, la posició central de les dades ordenades l'ocupen les dades 13a i 14a. Fixa't que abans de la dada 13a hi ha 12 dades (de la 1a a la 12a) i que després de la dada 14a també hi ha 12 dades (de la 15a a la 26a). La dada 13a correspon al valor de la variable 4 parells i la dada 14a correspon al valor de la variable 5 parells. Això ho podem veure fàcilment a partir de les  $N_i$ . Fixa't que la  $N_i$  del valor de la variable 4 parells és 13. Això ens indica que fins a 4 parells hi ha 13 dades. L'última d'aquestes dades, la número 13, correspon, doncs, al valor de la variable 4 parells. La  $N_i$  del valor de la variable 5 parells és 17, la qual cosa indica que les dades 14a, 15a, 16a i 17a corresponen al valor de la variable 5 parells.

Ara ja sabem els valors de la variable que ocupen la posició central: 4 parells i 5 parells. Calculem ara la mitjana d'aquests dos valors per obtenir la mediana:

$$Me = \frac{4+5}{2} = 4,5 \text{ punts}$$

## El mode

El **mode** és la dada que més succeeix.

El símbol del mode és **Mo**.

### Exemple 1

«Quants exàmens us agradaria fer durant aquest trimestre?», va preguntar el professor als seus alumnes.

Els alumnes van respondre i el professor va escriure les respostes a la pissarra:

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

En Joan va calcular ràpidament el valor de la mitjana, 2,5. «Hem de fer 2,5 exàmens durant el trimestre», va dir.

La Maria intervingué aleshores: «Això no té ni solta ni volta. Hem de fer dos exàmens durant el trimestre, perquè és la tria majoritària». La Maria té raó. Hi ha més dosos a la pissarra que qualsevol altre nombre. Així doncs, el 2 és el nombre més comú i l'anomenem **mode** o **valor modal**.

Per tant, en aquest cas **Mo = 2 exàmens**.

De vegades el mode és més útil que no pas la mitjana.

### Exemple 2

Ara utilitzarem un altre cop l'exemple del nombre de parells de sabates que tenen els integrants d'un grup d'estudiants, per tal d'aprendre com calcular el mode a partir d'una taula de freqüències.

Parells de sabates	$n_i$
1	1
2	2
<b>3</b>	6
4	4
5	4
6	3
7	1
8	0
9	0
10	3
11	1
N = 25	

En aquest cas, podem veure fàcilment que el nombre de parells de sabates més habitual, amb sis dades enregistrades, correspon al valor de la variable 3 parells de sabates (en color blau a la taula).

Així doncs, **Mo = 3 parells de sabates**.

• **Fes les activitats d'aprenentatge 3, 4, 5 i 6**

## 98 2. Mesures de dispersió

Si només utilitzem les mesures de centralització, no donem una informació acurada. També calen les mesures de dispersió, que ens indiquen com d'escampades estan les dades. Amb la utilització conjunta de les mesures de centralització i de dispersió podem conèixer amb gran precisió quina és la distribució de les dades.

Les mesures de dispersió que veurem són: el **rang**, la **desviació mitjana**, la **variança** i la **desviació típica**.

### El rang

El rang o **recorregut** és la mesura de dispersió més senzilla, i és la diferència entre el valor més gran i el més petit que pren la variable.

Per calcular-lo, únicament hem de restar el valor més petit al valor més gran.

El símbol del rang és **R**.

Passem ara als exemples. Mantindrem els mateixos exemples per a totes les mesures de dispersió. D'aquesta manera, podrem comparar les característiques de cada mesura. El primer exemple contindrà poques dades i el segon exemple constarà d'un nombre més gran de dades, per la qual cosa caldrà treballar amb una taula de freqüències.

#### Exemple 1

En Sebastià agafa el tren de dilluns a divendres per anar a la feina. La setmana passada va anotar en una llibreta els minuts d'endarreriment que duia el tren cada dia:

2 minuts      5 minuts      3 minuts      7 minuts      8 minuts

El rang d'aquesta col·lecció de dades és:

$$R = \text{valor més gran} - \text{valor més petit}$$

$$R = 8 \text{ minuts} - 2 \text{ minuts}$$

$$R = 6 \text{ minuts}$$



**Exemple 2**

Com ja vam comentar, el segon exemple serà el mateix que ja vam veure en l'estudi de les mesures de centralització, el del nombre de parells de sabates que tenen els vint-i-cinc estudiants d'una classe. D'aquesta manera, podrem comparar, a través d'un únic cas, tots els termes mitjans, els de centralització i els de dispersió.

Recorda la taula que ja vam construir quan vam estudiar les mesures de centralització:

Parells de sabates	$n_i$
1	1
2	2
3	6
4	4
5	4
6	3
7	1
8	0
9	0
10	3
11	1
N = 25	

El rang del nombre de parells de sabates que tenen és:

$$R = \text{valor més gran} - \text{valor més petit}$$

$$R = 11 \text{ parells} - 1 \text{ parells}$$

$$R = 10 \text{ parells de sabates}$$

**La desviació mitjana**

La **desviació mitjana** ens indica de quina manera s'allunyen de la mitjana les dades d'un estudi estadístic.

El símbol de la desviació mitjana és **DM**.

Per tal de calcular la desviació mitjana, has de seguir aquests passos:

- Compta el nombre de dades per obtenir **N**.
- Troba la mitjana.
- Compara cada una de les dades amb la mitjana i fes-ne la resta. Resta el nombre petit al nombre gran. Tant se val si el gran és la mitjana o la dada en consideració. Si una dada té el mateix valor que la mitjana, el resultat d'aquesta operació serà 0.
- Suma tots els resultats de les restes.
- Divideix el nombre obtingut entre **N**.

Tornarem a utilitzar els exemples que vam fer servir per estudiar el rang.

**Exemple 1**

Recorda els minuts d'endarreriment del tren durant la setmana laboral, enregistrats per en Sebastià:

2 minuts      5 minuts      3 minuts      7 minuts      8 minuts

Primer, hem de calcular la mitjana d'aquesta col·lecció de dades:

$$\bar{X} = \frac{2+5+3+7+8}{5} = \frac{25}{5} = 5 \text{ minuts}$$

Ara, per calcular la desviació mitjana, buscarem les diferències entre cada una de les dades i la mitjana, les sumarem totes i dividirem el resultat de la suma entre N:

$$DM = \frac{(5-2)+(5-5)+(5-3)+(7-5)+(8-5)}{5} = \frac{3+0+2+2+3}{5} = \frac{10}{5} = 2 \text{ minuts}$$

**Exemple 2**

Continuem amb l'exemple dels parells de sabates per aprendre a trobar la desviació mitjana a partir d'una taula de freqüències.

Recorda la taula de freqüències:

Parells de sabates	$n_i$
1	1
2	2
3	6
4	4
5	4
6	3
7	1
8	0
9	0
10	3
11	1
N = 25	

Primer cal trobar la mitjana del nombre de parells de sabates. Malgrat que ja ho vam fer quan vam estudiar la mitjana, tornarem a calcular-la:

$$\bar{X} = \frac{(1 \times 1) + (2 \times 2) + (3 \times 6) + (4 \times 4) + (5 \times 4) + (6 \times 3) + (7 \times 1) + (10 \times 3) + (11 \times 1)}{25}$$

$$\bar{X} = \frac{1+4+18+16+20+18+7+30+11}{25} = \frac{125}{25} = 5 \text{ parells}$$

Ara, per calcular la desviació mitjana, hem de calcular les diferències entre cada un dels valors de la variable i la mitjana (excepte pels valors sense cap dada

recollida) i multiplicar aquestes diferències pel nombre de dades de cada valor de la variable. Després, s'han de sumar tots els resultats i dividir entre N:

$$DM = \frac{(5-1) \times 1 + (5-2) \times 2 + (5-3) \times 6 + (5-4) \times 4 + (5-5) \times 4 + (6-5) \times 3 + (7-5) \times 1 + (10-5) \times 3 + (11-5) \times 1}{25}$$

$$DM = \frac{4+6+12+4+0+3+2+15+6}{25} = \frac{52}{25} = 2,08 \text{ parells}$$

- Fes les activitats d'aprenentatge 7, 8, 9 i 10

## La variància i la desviació típica

La **variància** i la **desviació típica** (també anomenada **desviació estàndard**) són unes altres maneres d'indicar en quina mesura, per terme mitjà, les dades s'allunyen de la mitjana.

El símbol de la variància és  $S^2$  i el símbol de la desviació típica és  $S$ .

Per calcular la **variància** cal que segueixis aquests passos:

- Compta el nombre de dades, per obtenir  $N$ .
- Calcula la mitjana.
- Compara cada una de les dades amb la mitjana i fes-ne la resta (resta-li al nombre gran el nombre petit, sigui quin sigui el gran o el petit). Si una dada té el mateix valor que la mitjana, el resultat d'aquesta operació serà 0.
- Suma els quadrats de les diferències.
- Divideix el resultat de la suma entre  $N$ .

La **desviació típica** és l'arrel quadrada de la variància.

Tornarem a utilitzar els exemples que vam fer servir per estudiar la desviació mitjana.

### Exemple 1

Recorda els minuts d'endarreriment del tren durant la setmana laboral, enregistrats per en Sebastià:

2 minuts      5 minuts      3 minuts      7 minuts      8 minuts

Ja vam calcular la mitjana d'aquestes dades,  $\bar{X} = 5$  minuts.

Ara, per calcular la variància, hem de calcular els quadrats de les diferències entre cada una de les dades i la mitjana, sumar-les i dividir el resultat de la suma entre  $N$ :

$$S^2 = \frac{(5-2)^2 + (5-5)^2 + (5-3)^2 + (7-5)^2 + (8-5)^2}{5} = \frac{3^2+0^2+2^2+2^2+3^2}{5}$$

$$S^2 = \frac{9+0+4+4+9}{5} = \frac{26}{5} = 5,2 \text{ minuts}$$

Finalment, per trobar la desviació típica, cal calcular l'arrel quadrada de la variància:

$$S = \sqrt{5,2} = 2,28 \text{ minuts}$$

### Exemple 2

Reprenem ara l'exemple dels parells de sabates per aprendre a trobar la variància i la desviació típica a partir d'una taula de freqüències.

Tornem a la taula que ja vam fer servir pels altres termes mitjans:

Parells de sabates	$n_i$
1	1
2	2
3	6
4	4
5	4
6	3
7	1
8	0
9	0
10	3
11	1
N = 25	

Primer, cal trobar la mitjana del nombre de parells de sabates. Recorda que ja ho vam fer a l'apartat anterior, quan vam estudiar el procediment per obtenir la desviació mitjana, i que vam obtenir una mitjana de 5 parells ( $\bar{X} = 5$  parells).

Ara, per calcular la variància, hem de trobar el quadrats de les diferències entre cada un dels valors de la variable i la mitjana (excepte pels valors sense cap dada recollida) i multiplicar-los pel nombre de dades de cada un dels valors de la variable. Després, hem de sumar tots els resultats obtinguts i dividir el nombre obtingut entre N.

$$S^2 = \frac{(5-1)^2 \times 1 + (5-2)^2 \times 2 + (5-3)^2 \times 6 + (5-4)^2 \times 4 + (5-5)^2 \times 4 + (6-5)^2 \times 3 + (7-5)^2 \times 1 + (10-5)^2 \times 3 + (11-5)^2 \times 1}{25}$$

$$S^2 = \frac{16+18+24+4+0+3+4+75+36}{25} = \frac{180}{25} = 7,2 \text{ parells}$$

Finalment, per trobar la desviació típica, cal calcular l'arrel quadrada de la variància:

$$S = \sqrt{7,2} = 2,68 \text{ parells}$$

- Fes les activitats d'aprenentatge 11, 12, 13, 14, 15 i 16

### 3. Les distribucions estadístiques

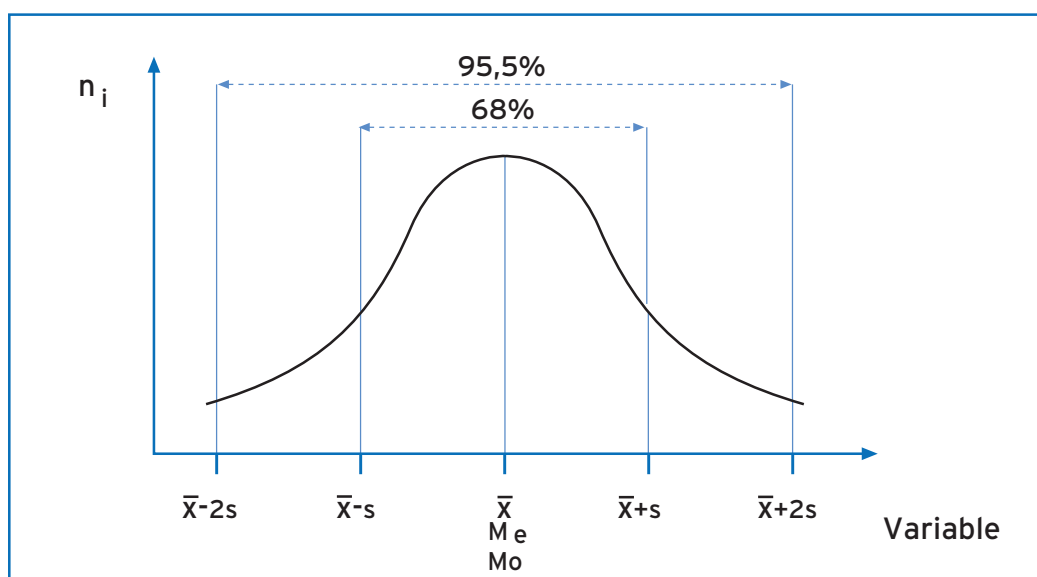
Fins ara, quan hem fet estudis estadístics, hem parlat de grups o col·leccions de dades. La paraula **distribució** fa referència a les dades que segueixen un patró determinat. Parlem de variables que, quan s'estudien a partir d'un nombre elevat de dades, segueixen un patró que també es repeteix en l'estudi d'altres variables. Com més gran sigui el nombre de dades recollides, més clarament ens n'adonarem de si les dades segueixen una determinada distribució estadística.

Per representar les distribucions, ens basarem en un polígon de freqüències amb els angles suavitzats fins a formar una línia sinuosa. El funcionament és, per tant, idèntic al d'un polígon de freqüències: per determinar la freqüència d'un valor de la variable només cal determinar el punt de la corba que relaciona el valor amb la seva freqüència, de manera semblant a com ho fem amb els polígons de freqüències.

Només farem referència a tres tipus de distribucions estadístiques: la **normal**, les **asimètriques** i la **bimodal**. Només veurem les característiques principals de cada una, però no hi entrarem a fons, ja que no és l'objectiu d'aquest mòdul.

#### La distribució normal

Aquesta distribució també s'anomena **campana de Gauss**, pels estudis que va fer Carl Friedrich Gauss (1777-1855) sobre aquesta distribució. És una distribució simètrica en forma de campana, que està associada a molts fenòmens naturals i quotidians. Moltes característiques morfològiques, com per exemple l'alçada o el pes, segueixen aquesta distribució.



Les característiques principals d'aquesta distribució són:

- La mitjana, la mediana i el mode coincideixen i marquen el punt de la variable on es divideix el nombre de dades en dues parts iguals (50 % i 50 %).
- La forma de la campana depèn de  $\bar{X}$  (la mitjana) i  $S$  (la desviació típica).  $\bar{X}$  indica la posició de la campana sobre l'eix d'abscisses. Com més gran sigui el seu valor, més «cap a la dreta» se situarà la campana.  $S$  determina si la

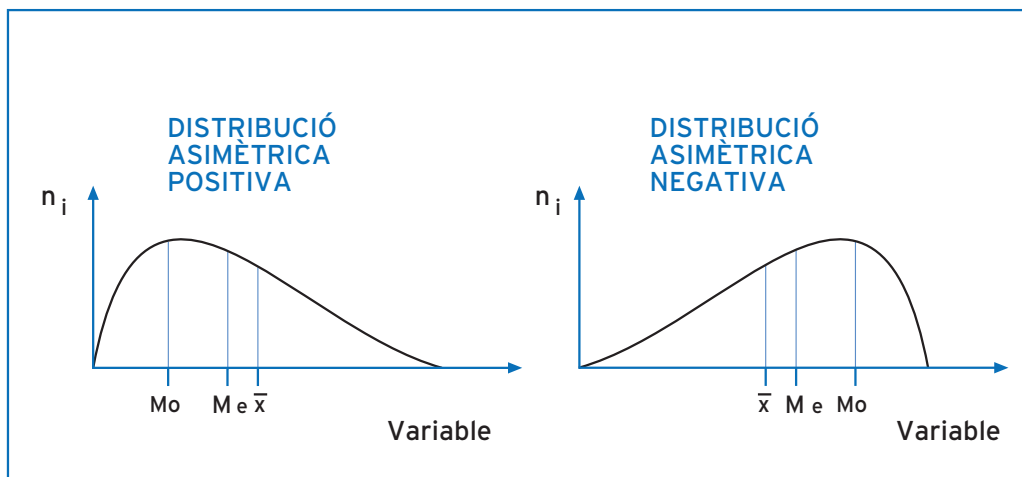
campana és més o menys aplanada. Com més gran sigui el seu valor, més disperses estaran les dades i, per tant, la campana serà més aplanada.

- Un 68 % de les dades se situen entre la mitjana i una desviació típica ( $S$ ). Un 95,5 % de les dades se situen entre la mitjana i dues desviacions típiques ( $2S$ ). Així doncs, si en l'estudi del pes d'una població de persones adultes que segueix una distribució normal,  $\bar{X}=72$  kg i  $S = 8$  kg, el 68 % de la població pesa entre 64 kg ( $\bar{X}-S$ ) i 80 kg ( $\bar{X}+S$ ) i el 95,5 % de la població pesa entre 56 kg ( $\bar{X}-2S$ ) i 88 kg ( $\bar{X}+2S$ ).

### Les distribucions asimètriques

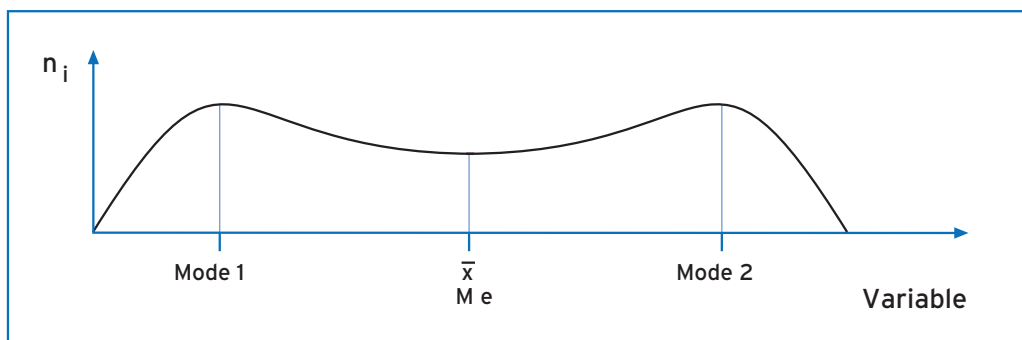
De vegades les dades adopten campanes asimètriques. Aleshores parlem de **distribucions asimètriques**. En aquests tipus de distribucions, la distància entre la mediana i el mode és aproximadament el doble de la distància entre la mediana i la mitjana.

Quan la campana «s'allarga» cap als valors més alts de la variable, parlem de **distribució asimètrica positiva**, mentre que, quan la campana «s'allarga» cap als valors més baixos de la variable, parlem de **distribució asimètrica negativa**.



### La distribució bimodal

Hi ha estudis que recullen dades que segueixen dues tendències força diferenciades. En aquests casos, parlem de distribucions bimodals, atès que hi ha dos modes, entre els quals es troben les altres mesures de centralització: la mitjana i la mediana.



**Activitat 1**

L'Albert va cronometrar la durada d'un mateix trajecte d'autobús de dilluns a divendres. Aquests són els resultats en minuts:

20 15 16 20 27

Calcula la duració mitjana del trajecte durant aquests dies.

**Activitat 2**

El Sergi va mesurar les llargades de 20 cues de sargantana de paret i va recollir aquestes dades en centímetres:

10 12 15 10 17 10 15 12 15 11  
14 14 17 13 14 15 10 17 11 12

Fes una taula de freqüències absolutes i troba la llargada mitjana de les cues.

**Activitat 3**

L'Annabel demana els seus amics quantes pel·lícules han vist durant la setmana i recull aquestes dades:

1 0 2 1 0 1 1

Calcula la mitjana, la mediana i el mode del nombre de pel·lícules vistes pels amics de l'Annabel.

**Activitat 4**

La Maria recull dades sobre el nombre de peces defectuoses detectades diàriament en un petit taller de fabricació de components elèctrics. Heus aquí el recull de dades sobre peces defectuoses trobades durant els dies de control:

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

Calcula la mitjana, la mediana i el mode del nombre de peces defectuoses detectades diàriament.



**Activitat 5**

En Joan, el davanter centre del F.C. Sinera, va jugar dotze partits el mes passat. Aquí hi ha un recull dels gols marcats a cada partit:

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

Calcula la mitjana, la mediana i el mode del nombre de gols marcats.

**Activitat 6**

En Lluís compta el nombre de maduixes en quaranta-una safates de 250 grams i recull aquestes dades:

8	9	10	9	11	8	15	12	9	11
15	15	9	14	12	10	8	12	12	10
11	13	15	13	8	10	16	13	10	10
8	9	12	14	10	15	11	10	15	8
14									

Calcula la mitjana, la mediana i el mode del nombre de maduixes.

**Activitat 7**

En Pere ha mesurat la longitud total de set salamandres. Heus aquí les dades recollides, en centímetres:

22 31 25 26 20 22 29

Calcula el rang i la desviació mitjana de la longitud total de les salamandres mesurades.

**Activitat 8**

La Marina estudia el nombre de cries de ratolí espiguer nascudes per cada part. Heus aquí les dades del nombre de cries per part, després d'haver-ne controlat trenta:

3	2	2	5	1	3	2	6	3	2
4	2	3	1	2	3	2	2	3	5
8	2	6	3	2	4	1	1	4	3

Calcula el rang i la desviació mitjana del nombre de cries per part.

**Activitat 9**

La Sara va fer deu exàmens de ciències durant el curs passat i va obtenir aquestes notes:

8 4 5 9 10 7 4 6 4 8

Calcula el rang i la desviació mitjana de les notes.

**Activitat 10**

En Jordi compta el nombre de llumins que hi ha en vint-i-set caixes de llumins i recull aquestes dades:

32 29 30 33 31 34 27 26 32 32  
26 26 32 33 26 30 29 28 29 33  
34 29 28 25 31 33 32

Calcula el rang i la desviació mitjana del nombre de llumins per caixa.

**Activitat 11**

Es pregunta a deu persones de Terrassa i a deu persones de Manresa quants llibres van llegir l'any passat, i es recullen aquestes dades:

Grup de Terrassa    2    6    6    3    2    1    4    2    3    1

Grup de Manresa    4    1    1    0    3    1    7    1    6    6

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les dades més escampades?

**Activitat 12**

Les notes obtingudes en l'assignatura d'anatomia per dos grups de 15 estudiants universitaris de primer curs de medicina són:

Universitat de Michigan

0	3	8	3	6	2	7	4	7	3
1	4	3	5	4					

Universitat de Chicago

10	4	6	5	8	5	9	7	5	5
6	5	7	3	5					

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les puntuacions més escampades?

**Activitat 13**

Un estudi sobre el nombre de diaris llegits diàriament per dos grups de vint homes i dones de negocis, un de Mumbai i l'altre de Delhi, ha donat aquests resultats:

Grup de Mumbai

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

Grup de Delhi

1	0	4	0	2	1	0	9	8	6
8	4	8	9	5	0	2	1	1	1

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les dades més escampades?

**Activitat 14**

Durant la primera quinzena d'agost de l'any passat, les temperatures recollides a mitjanit, en graus centígrads, a Karachi i a Mumbai van ésser:

Karachi

35	32	32	35	31	31	31	36	32	31
33	33	36	36	31					

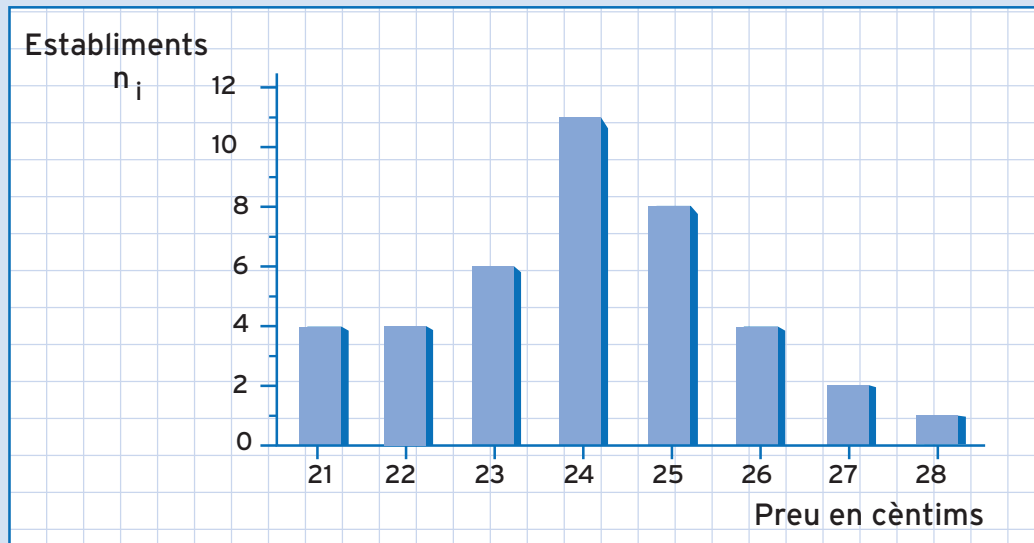
Mumbai

28	29	30	30	30	32	31	29	27	25
26	30	34	35	34					

Calcula totes les mesures de centralització i de dispersió per a cada recull de temperatures. A quina ciutat s'han recollit dades més disperses?

**Activitat 15**

S'ha fet un estudi sobre el preu d'una ampolla d'aigua mineral d'una determinada marca a diferents establiments comercials. Heus aquí la gràfica que il·lustra els resultats obtinguts:

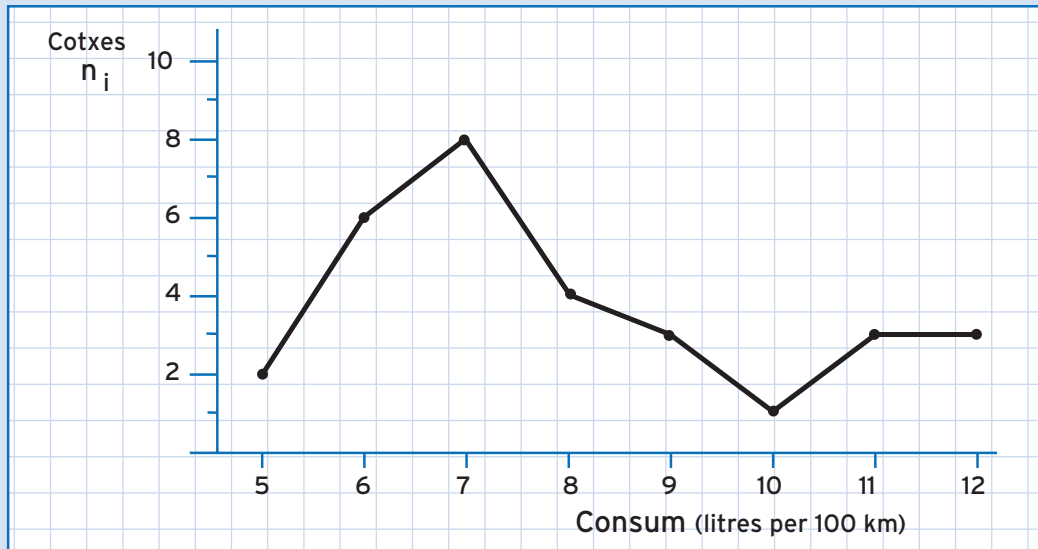


A partir d'aquesta gràfica, calcula totes les mesures de centralització i de dispersió que es desprenen d'aquest estudi.



**Activitat 16**

S'ha fet un estudi sobre el consum de benzina dels vehicles dels trenta treballadors d'una empresa. El consum s'ha mesurat en litres per cada cent quilòmetres. Heus aquí la gràfica que il·lustra els resultats obtinguts:



A partir d'aquesta gràfica, calcula totes les mesures de centralització i de dispersió que es desprenen d'aquest estudi.

**Activitat 1**

Recull dades sobre la cotització del dòlar dels EUA durant deu dies laborables consecutius. Pots obtenir la informació a través d'un diari, d'una pàgina d'internet o del teletext. Agafa les cotitzacions amb dues xifres decimals. Un cop recollides les deu dades, calcula totes les mesures de centralització i de dispersió que co-neixes.

Cal tenir en compte només els dies laborables, ja que en els dies festius no hi ha canvis en les cotitzacions. Perquè l'estudi sigui més acurat, cal que agafis les dades en les mateixes condicions. Si utilitzes un diari, cal que sempre sigui el mateix. Si utilitzes internet o teletext, cal utilitzar sempre la mateixa font i a la mateixa hora.

Atès que només tindrem en compte dues xifres decimals, tingues en compte el valor de la tercera xifra decimal per tal de pujar un dígit el valor de la segona xifra decimal, si s'escau (si la tercera és 5 o més de 5).

**Activitat 2**

Busca algunes variables que segueixin una distribució normal. Pots cercar en llibres de ciència, enciclopèdies o a través d'internet. Hi ha molts fenòmens morfològics i estudis mèdics que segueixen aquesta distribució.

**Activitat 1**

En Francesc té dues gallines, una blanca i una rossa, i compta els ous que pon cada una durant cada un dels dies d'una setmana. Heus aquí els resultats:

Gallina blanca:      4    7    6    6    6    7    6

Gallina rossa:        3    3    3    6    3    4    6

Calcula totes les mesures de centralització i de dispersió del nombre d'ous posats diàriament per cada gallina. Quina gallina ha post els ous de manera més dispersa?

**Activitat 2**

Un estudi sobre el nombre de nens que han tingut dos grups de dones de més de cinquanta anys ha donat aquests resultats:

Grup de Bangkok

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

Grup de Bali

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

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les dades més escampades?

**Activitat 1**

L'Albert va cronometrar la durada d'un mateix trajecte d'autobús de dilluns a divendres. Aquests són els resultats en minuts:

20 15 16 20 27

Calcula la duració mitjana del trajecte durant aquests dies.

$$\bar{X} = \frac{20+15+16+20+27}{5} = \frac{98}{5} = 19,6 \text{ minuts}$$

Ho podem expressar en minuts i segons: 19 minuts i 36 segons.

Recorda que per passar la part decimal de minuts a segons, cal multiplicar-la per 60:  $0,6 \times 60 = 36$

**Activitat 2**

El Sergi va mesurar les llargades de 20 cues de sargantana de paret i va recollir aquestes dades en centímetres:

10 12 15 10 17 10 15 12 15 11  
14 14 17 13 14 15 10 17 11 12

Fes una taula de freqüències absolutes i troba la llargada mitjana de les cues.

Llargada cua	$n_i$
10	4
11	2
12	3
13	1
14	3
15	4
16	0
17	3
N = 20	

$$\bar{X} = \frac{(10 \times 4) + (11 \times 2) + (12 \times 3) + (13 \times 1) + (14 \times 3) + (15 \times 4) + (17 \times 3)}{20}$$

$$\bar{X} = \frac{40 + 22 + 36 + 13 + 42 + 60 + 51}{20} = \frac{264}{20} = 13,2 \text{ cm}$$

**Activitat 3**

L'Annabel demana els seus amics quantes pel·lícules han vist durant la setmana i recull aquestes dades:

1 0 2 1 0 1 1

Calcula la mitjana, la mediana i el mode del nombre de pel·lícules vistes pels amics de l'Annabel.

Primer ordenem les dades: 0 0 1 1 1 1 2

$$\bar{X} = \frac{0+0+1+1+1+1+2}{7} = \frac{6}{7} = 0,86 \text{ pel·lícules}$$

Me = 1 pel·lícula Ja que la dada 1 pel·lícula és la que ocupa la posició central de les dades ordenades.

Mo = 1 pel·lícula Ja que la majoria d'amics van veure 1 pel·lícula.

**Activitat 4**

La Maria recull dades sobre el nombre de peces defectuoses detectades diàriament en un petit taller de fabricació de components elèctrics. Heus aquí el recull de dades sobre peces defectuoses trobades durant els dies de control:

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

Calcula la mitjana, la mediana i el mode del nombre de peces defectuoses detectades diàriament.

Peces defectuoses	$n_i$	$N_i$
0	8	8
1	4	12
2	6	18
3	3	21
4	3	24
5	2	26
N = 26		

$$\bar{X} = \frac{(0 \times 8) + (1 \times 4) + (2 \times 6) + (3 \times 3) + (4 \times 3) + (5 \times 2)}{26}$$

$$\bar{X} = \frac{0+4+12+9+12+10}{26} = \frac{47}{26} = 1,81 \text{ peces}$$

Me = 2 peces Ja que les dues posicions centrals (13a i 14a) corresponen al valor de la variable 2 peces. Això es desprèn de les  $N_i$ , atès que ens indiquen que les dades compreses entre les posicions 13a i 18a corresponen al valor de la variable 2 peces.

Mo = 0 peces Ja que el nombre de dies en què no es va detectar cap peça defectuosa és el majoritari.

**Activitat 5**

En Joan, el davanter centre del F.C. Sinera, va jugar dotze partits el mes passat. Aquí hi ha un recull dels gols marcats a cada partit:

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

Calcula la mitjana, la mediana i el mode del nombre de gols marcats.

Primer ordenem les dades:

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

$$\bar{X} = \frac{0+0+0+0+1+1+1+2+2+2+3+3}{12} = \frac{15}{12} = 1,25 \text{ gols}$$

Me = 1 gol Ja que les dues posicions centrals (6a i 7a) corresponen al valor de la variable 1 gol.

Mo = 0 gols Ja que **0 gols** és el valor de la variable que més s'ha produït.

**Activitat 6**

En Lluís compta el nombre de maduixes en quaranta-una safates de 250 grams i recull aquestes dades:

8 9 10 9 11 8 15 12 9 11  
 15 15 9 14 12 10 8 12 12 10  
 11 13 15 13 8 10 16 13 10 10  
 8 9 12 14 10 15 11 10 15 8  
 14

Calcula la mitjana, la mediana i el mode del nombre de maduixes.

Atès l'elevat nombre de dades, farem una taula de freqüències:

Maduixes	$n_i$	$N_i$
8	6	6
9	5	11
10	8	19
11	4	23
12	5	28
13	3	31
14	3	34
15	6	40
16	1	41
		N = 41

$$\bar{X} = \frac{(8 \times 6) + (9 \times 5) + (10 \times 8) + (11 \times 4) + (12 \times 5) + (13 \times 3) + (14 \times 3) + (15 \times 6) + (16 \times 1)}{41}$$

$$\bar{X} = \frac{48 + 45 + 80 + 44 + 60 + 39 + 42 + 90 + 16}{41} = \frac{464}{41} = 11,32 \text{ maduixes}$$

Me = 11 maduixes Ja que la dada central, la 21a, correspon al valor de la variable 11 maduixes.

Mo = 10 maduixes Ja que hi ha més safates amb deu maduixes que no pas amb cap altra quantitat.

**Activitat 7**

En Pere ha mesurat la longitud total de set salamandres. Heus aquí les dades recollides, en centímetres:

22 31 25 26 20 22 29

Calcula el rang i la desviació mitjana de la longitud total de les salamandres mesurades.

Primer ordenem les dades: 20 22 22 25 26 29 31

$$R = 31 - 20 = 11 \text{ cm}$$

Per poder calcular la desviació mitjana, primer cal calcular la mitjana:

$$\bar{X} = \frac{20+22+22+25+26+29+31}{7} = \frac{175}{7} = 25 \text{ cm}$$

Ara ja podem calcular la desviació mitjana:

$$DM = \frac{(25-20)+(25-22)+(25-22)+(25-25)+(26-25)+(29-25)+(31-25)}{7}$$

$$DM = \frac{5+3+3+0+1+4+6}{7} = \frac{22}{7} = 3,14 \text{ cm}$$

**Activitat 8**

La Marina estudia el nombre de cries de ratolí espiguer nascudes per cada part. Heus aquí les dades del nombre de cries per part, després d'haver-ne controlat trenta:

3 2 2 5 1 3 2 6 3 2  
4 2 3 1 2 3 2 2 3 5  
8 2 6 3 2 4 1 1 4 3

Calcula el rang i la desviació mitjana del nombre de cries per part.

Atès l'elevat nombre de dades, farem una taula de freqüències:

Cries	$n_i$
1	4
2	10
3	8
4	3
5	2
6	2
7	0
8	1
N = 30	

$$R = 8 - 1 = 7 \text{ cries}$$

Per poder calcular la desviació mitjana, primer cal calcular la mitjana:

$$\bar{X} = \frac{(1 \times 4) + (2 \times 10) + (3 \times 8) + (4 \times 3) + (5 \times 2) + (6 \times 2) + (8 \times 1)}{30}$$

$$\bar{X} = \frac{4+20+24+12+10+12+8}{30} = \frac{90}{30} = 3 \text{ cries}$$

Ara ja podem calcular la desviació mitjana:

$$DM = \frac{(3-1) \times 4 + (3-2) \times 10 + (3-3) \times 8 + (4-3) \times 3 + (5-3) \times 2 + (6-3) \times 2 + (8-3) \times 1}{30}$$

$$DM = \frac{8+10+0+3+4+6+5}{30} = \frac{36}{30} = 1,2 \text{ cries}$$



**Activitat 9**

La Sara va fer deu exàmens de ciències durant el curs passat i va obtenir aquestes notes:

8 4 5 9 10 7 4 6 4 8

Calcula el rang i la desviació mitjana de les notes.

Primer ordenem les dades:

4 4 4 5 6 7 8 8 9 10

$R = 10 - 4 = 6$  punts

Per poder calcular la desviació mitjana, primer cal calcular la mitjana:

$$\bar{X} = \frac{4+4+4+5+6+7+8+8+9+10}{10} = \frac{65}{10} = 6,5 \text{ punts}$$

Ara ja podem calcular la desviació mitjana:

$$DM = \frac{(6,5-4)+(6,5-4)+(6,5-4)+(6,5-5)+(6,5-6)+(7-6,5)+(8-6,5)+(8-6,5)+(9-6,5)+(10-6,5)}{10}$$

$$DM = \frac{2,5+2,5+2,5+1,5+0,5+0,5+1,5+1,5+2,5+3,5}{10} = \frac{19}{10} = 1,9 \text{ punts}$$

**Activitat 10**

En Jordi compta el nombre de llumins que hi ha en vint-i-set caixes de llumins i recull aquestes dades:

32 29 30 33 31 34 27 26 32 32  
26 26 32 33 26 30 29 28 29 33  
34 29 28 25 31 33 32

Calcula el rang i la desviació mitjana del nombre de llumins per caixa.

Atès l'elevat nombre de dades, farem una taula de freqüències:

Llumins	$n_i$
25	1
26	4
27	1
28	2
29	4
30	2
31	2
32	5
33	4
34	2
N = 27	

$R = 34 - 25 = 9$  llumins

Per poder calcular la desviació mitjana, primer cal calcular la mitjana:

$$\bar{X} = \frac{(25 \times 1) + (26 \times 4) + (27 \times 1) + (28 \times 2) + (29 \times 4) + (30 \times 2) + (31 \times 2) + (32 \times 5) + (33 \times 4) + (34 \times 2)}{27}$$

$$\bar{X} = \frac{25+104+27+56+116+60+62+160+132+68}{27} = \frac{810}{27} = 30 \text{ llumins}$$

Ara ja podem calcular la desviació mitjana:

$$DM = \frac{(30-25) \times 1 + (30-26) \times 4 + (30-27) \times 1 + (30-28) \times 2 + (30-29) \times 4 + (30-30) \times 2 + (31-30) \times 2 + (32-30) \times 5 + (33-30) \times 4 + (34-30) \times 2}{27}$$

$$DM = \frac{5+16+3+4+4+0+2+10+12+8}{27} = \frac{64}{27} = 2,37 \text{ llumins}$$

**Activitat 11**

Es pregunta a deu persones de Terrassa i a deu persones de Manresa quants llibres van llegir l'any passat i es recullen aquestes dades:

Grup de Terrassa	2	6	6	3	2	1	4	2	3	1
Grup de Manresa	4	1	1	0	3	1	7	1	6	6

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les dades més escampades?

**Grup de Terrassa.** Primer ordenem les dades:

1 1 2 2 2 3 3 4 6 6

$$\bar{X} = \frac{1+1+2+2+2+3+3+4+6+6}{10} = \frac{30}{10} = 3 \text{ llibres}$$

$$Me = \frac{2+3}{2} = \frac{5}{2} = 2,5 \text{ llibres}$$

$$Mo = 2 \text{ llibres}$$

$$R = 6 - 1 = 5 \text{ llibres}$$

$$DM = \frac{(3-1)+(3-1)+(3-2)+(3-2)+(3-2)+(3-3)+(3-3)+(4-3)+(6-3)+(6-3)}{10}$$

$$DM = \frac{2+2+1+1+1+0+0+1+3+3}{10} = \frac{14}{10} = 1,4 \text{ llibres}$$

$$S^2 = \frac{(3-1)^2+(3-1)^2+(3-2)^2+(3-2)^2+(3-2)^2+(3-3)^2+(3-3)^2+(4-3)^2+(6-3)^2+(6-3)^2}{10}$$

$$S^2 = \frac{4+4+1+1+1+0+0+1+9+9}{10} = \frac{30}{10} = 3 \text{ llibres}$$

$$S = \sqrt{3} = 1,73 \text{ llibres}$$

**Grup de Manresa.** Primer ordenem les dades:

0 1 1 1 1 3 4 6 6 7

$$\bar{X} = \frac{0+1+1+1+1+3+4+6+6+7}{10} = \frac{30}{10} = 3 \text{ llibres}$$

$$\bar{X} = \frac{1+3}{2} = \frac{4}{2} = 2 \text{ llibres}$$

$$Mo = 1 \text{ llibre}$$

$$R = 7 - 0 = 7 \text{ llibres}$$

$$DM = \frac{(3-0)+(3-1)+(3-1)+(3-1)+(3-1)+(3-3)+(4-3)+(6-3)+(6-3)+(7-3)}{10}$$

$$DM = \frac{3+2+2+2+2+0+1+3+3+4}{10} = \frac{22}{10} = 2,2 \text{ llibres}$$

$$S^2 = \frac{(3-0)^2+(3-1)^2+(3-1)^2+(3-1)^2+(3-1)^2+(3-3)^2+(4-3)^2+(6-3)^2+(6-3)^2+(7-3)^2}{10}$$

$$S^2 = \frac{9+4+4+4+4+0+1+9+9+16}{10} = \frac{60}{10} = 6 \text{ llibres}$$

$$S = \sqrt{6} = 2,45 \text{ llibres}$$

Les dades estan més escampades al grup de Manresa, perquè les mesures de dispersió són més grans.

**Activitat 12**

Les notes obtingudes en l'assignatura d'anatomia per dos grups de 15 estudiants universitaris de primer curs de medicina són:

Universitat de Michigan	0	3	8	3	6	2	7	4	7	3
	1	4	3	5	4					
Universitat de Chicago	10	4	6	5	8	5	9	7	5	5
	6	5	7	3	5					

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les puntuacions més escampades?

Atès l'elevat nombre de dades, treballarem amb taules de freqüències.

**Universitat de Michigan**

Punts	$n_i$	$N_i$
0	1	1
1	1	2
2	1	3
3	4	7
4	3	10
5	1	11
6	1	12
7	2	14
8	1	15
N = 15		

$$\bar{X} = \frac{(0 \times 1) + (1 \times 1) + (2 \times 1) + (3 \times 4) + (4 \times 3) + (5 \times 1) + (6 \times 1) + (7 \times 2) + (8 \times 1)}{15}$$

$$\bar{X} = \frac{0 + 1 + 2 + 12 + 12 + 5 + 6 + 14 + 8}{15} = \frac{60}{15} = 4 \text{ punts}$$

$$Me = 4 \text{ punts}$$

$$Mo = 3 \text{ punts}$$

$$R = 8 - 0 = 8 \text{ punts}$$

$$DM = \frac{(4-0)x1 + (4-1)x1 + (4-2)x1 + (4-3)x4 + (4-4)x3 + (5-4)x1 + (6-4)x1 + (7-4)x2 + (8-4)x1}{15}$$

$$DM = \frac{4 + 3 + 2 + 4 + 0 + 1 + 2 + 6 + 4}{15} = \frac{26}{15} = 1,73 \text{ punts}$$

$$S^2 = \frac{(4-0)^2x1 + (4-1)^2x1 + (4-2)^2x1 + (4-3)^2x4 + (4-4)^2x3 + (5-4)^2x1 + (6-4)^2x1 + (7-4)^2x2 + (8-4)^2x1}{15}$$

$$S^2 = \frac{16 + 9 + 4 + 4 + 0 + 1 + 4 + 18 + 16}{15} = \frac{72}{15} = 4,8 \text{ punts}$$

$$S = \sqrt{4,8} = 2,19 \text{ punts}$$

Punts	$n_i$	$N_i$
3	1	1
4	1	2
5	6	8
6	2	10
7	2	12
8	1	13
9	1	14
10	1	15
N = 15		

$$\bar{X} = \frac{(3 \times 1) + (4 \times 1) + (5 \times 6) + (6 \times 2) + (7 \times 2) + (8 \times 1) + (9 \times 1) + (10 \times 1)}{15}$$

$$\bar{X} = \frac{3 + 4 + 30 + 12 + 14 + 8 + 9 + 10}{15} = \frac{90}{15} = 6 \text{ punts}$$

Me = 5 punts

Mo = 5 punts

R = 10 - 3 = 7 punts

$$DM = \frac{(6-3)x1 + (6-4)x1 + (6-5)x6 + (6-6)x2 + (7-6)x2 + (8-6)x1 + (9-6)x1 + (10-6)x1}{15}$$

$$DM = \frac{3 + 2 + 6 + 0 + 2 + 2 + 3 + 4}{15} = \frac{22}{15} = 1,47 \text{ punts}$$

$$S^2 = \frac{(6-3)^2x1 + (6-4)^2x1 + (6-5)^2x6 + (6-6)^2x2 + (7-6)^2x2 + (8-6)^2x1 + (9-6)^2x1 + (10-6)^2x1}{15}$$

$$S^2 = \frac{9 + 4 + 6 + 0 + 2 + 4 + 9 + 16}{15} = \frac{50}{15} = 3,33 \text{ punts}$$

$$S = \sqrt{3,33} = 1,82 \text{ punts}$$

Les dades estan més escampades a la Universitat de Michigan, perquè les mesures de dispersió són més grans.

**Activitat 13**

Un estudi sobre el nombre de diaris llegits diàriament per dos grups de vint homes i dones de negocis, un de Mumbai i l'altre de Delhi, ha donat aquests resultats:

Grup de Mumbai	3	2	6	6	5	3	5	4	2	2
	4	5	5	5	3	3	5	5	2	5
Grup de Delhi	1	0	4	0	2	1	0	9	8	6
	8	4	8	9	5	0	2	1	1	1

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les dades més escampades?

Atès l'elevat nombre de dades, treballarem amb taules de freqüències.

**Grup de Mumbai**

Diaris	$n_i$	$N_i$
2	4	4
3	4	8
4	2	10
5	8	18
6	2	20
N = 20		

$$\bar{X} = \frac{(2 \times 4) + (3 \times 4) + (4 \times 2) + (5 \times 8) + (6 \times 2)}{20} = \frac{8 + 12 + 8 + 40 + 12}{20} = \frac{80}{20} = 4 \text{ diaris}$$

$$Me = \frac{4 + 5}{2} = 4,5 \text{ diaris}$$

$$Mo = 5 \text{ diaris}$$

$$R = 6 - 2 = 4 \text{ diaris}$$

$$DM = \frac{(4-2) \times 4 + (4-3) \times 4 + (4-4) \times 2 + (5-4) \times 8 + (6-4) \times 2}{20} = \frac{8 + 4 + 0 + 8 + 4}{20} = \frac{24}{20} = 1,2 \text{ diaris}$$

$$S^2 = \frac{(4-2)^2 \times 4 + (4-3)^2 \times 4 + (4-4)^2 \times 2 + (5-4)^2 \times 8 + (6-4)^2 \times 2}{20} = \frac{16 + 4 + 0 + 8 + 8}{20} = \frac{36}{20} = 1,8 \text{ diaris}$$

$$S = \sqrt{1,8} = 1,34 \text{ diaris}$$

## Grup de Delhi

Diaris	$n_i$	$N_i$
0	4	4
1	5	9
2	2	11
3	0	11
4	2	13
5	1	14
6	1	15
7	0	15
8	3	18
9	2	20
N = 20		

$$\bar{X} = \frac{(0 \times 4) + (1 \times 5) + (2 \times 2) + (4 \times 2) + (5 \times 1) + (6 \times 1) + (8 \times 3) + (9 \times 2)}{20}$$

$$\bar{X} = \frac{0 + 5 + 4 + 8 + 5 + 6 + 24 + 18}{20} = \frac{70}{20} = 3,5 \text{ diaris}$$

Me = 2 diaris

Mo = 1 diari

R = 9 - 0 = 9 diaris

$$DM = \frac{(3,5-0) \times 4 + (3,5-1) \times 5 + (3,5-2) \times 2 + (4-3,5) \times 2 + (5-3,5) \times 1 + (6-3,5) \times 1 + (8-3,5) \times 3 + (9-3,5) \times 2}{20}$$

$$DM = \frac{14 + 12,5 + 3 + 1 + 1,5 + 2,5 + 13,5 + 11}{20} = \frac{59}{15} = 2,95 \text{ diaris}$$

$$S^2 = \frac{(3,5-0)^2 \times 4 + (3,5-1)^2 \times 5 + (3,5-2)^2 \times 2 + (4-3,5)^2 \times 2 + (5-3,5)^2 \times 1 + (6-3,5)^2 \times 1 + (8-3,5)^2 \times 3 + (9-3,5)^2 \times 2}{20}$$

$$S^2 = \frac{49 + 31,25 + 4,5 + 0,5 + 2,25 + 6,25 + 60,75 + 60,5}{20} = \frac{215}{20} = 10,75 \text{ diaris}$$

$$S = \sqrt{10,75} = 3,28 \text{ diaris}$$

Les dades estan més escampades al grup de Delhi perquè les mesures de dispersió són més grans.

**Activitat 14**

Durant la primera quinzena d'agost de l'any passat, les temperatures recollides a mitjanit, en graus centígrads, a Karachi i a Mumbai van ésser:

Karachi	35	32	32	35	31	31	31	36	32	31
	33	33	36	36	31					
Mumbai	28	29	30	30	30	32	31	29	27	25
	26	30	34	35	34					

Calcula totes les mesures de centralització i de dispersió per a cada recull de temperatures. A quina ciutat s'han recollit dades més disperses?

Atesa l'elevada quantitat de dades, utilitzarem taules de freqüències.

**Karachi**

Temperatura	$n_i$	$N_i$
31°	5	5
32°	3	8
33°	2	10
34°	0	10
35°	2	12
36°	3	15
	<b>N = 15</b>	

$$\bar{X} = \frac{(31 \times 5) + (32 \times 3) + (33 \times 2) + (35 \times 2) + (36 \times 3)}{15} = \frac{155 + 96 + 66 + 70 + 108}{15} = \frac{495}{15} = 33^\circ$$

$$Me = 32^\circ$$

$$Mo = 31^\circ$$

$$R = 36 - 31 = 5^\circ$$

$$DM = \frac{(33-31) \times 5 + (33-32) \times 3 + (33-33) \times 2 + (35-33) \times 2 + (36-33) \times 3}{15} = \frac{10 + 3 + 0 + 4 + 9}{15} = \frac{26}{15} = 1,73^\circ$$

$$S^2 = \frac{(33-31)^2 \times 5 + (33-32)^2 \times 3 + (33-33)^2 \times 2 + (35-33)^2 \times 2 + (36-33)^2 \times 3}{15} = \frac{20 + 3 + 0 + 8 + 27}{15} = \frac{58}{15} = 3,87^\circ$$

$$S = \sqrt{3,87} = 1,97^\circ$$

## Mumbai

Temperatura	$n_i$	$N_i$
25°	1	1
26°	1	2
27°	1	3
28°	1	4
29°	2	6
30°	4	10
31°	1	11
32°	1	12
33°	0	12
34°	2	14
35°	1	15
$N = 15$		

$$\bar{X} = \frac{(25 \times 1) + (26 \times 1) + (27 \times 1) + (28 \times 1) + (29 \times 2) + (30 \times 4) + (31 \times 1) + (32 \times 1) + (34 \times 2) + (35 \times 1)}{15}$$

$$\bar{X} = \frac{25 + 26 + 27 + 28 + 58 + 120 + 31 + 32 + 68 + 35}{15} = \frac{450}{15} = 30^\circ$$

$$Me = 30^\circ$$

$$Mo = 30^\circ$$

$$R = 35 - 25 = 10^\circ$$

$$DM = \frac{(30-25) \times 1 + (30-26) \times 1 + (30-27) \times 1 + (30-28) \times 1 + (30-29) \times 2 + (30-30) \times 4 + (31-30) \times 1 + (32-30) \times 1 + (34-30) \times 2 + (35-30) \times 1}{15}$$

$$DM = \frac{5 + 4 + 3 + 2 + 2 + 0 + 1 + 2 + 8 + 5}{15} = \frac{32}{15} = 2,13^\circ$$

$$S^2 = \frac{(30-25)^2 \times 1 + (30-26)^2 \times 1 + (30-27)^2 \times 1 + (30-28)^2 \times 1 + (30-29)^2 \times 2 + (30-30)^2 \times 4 + (31-30)^2 \times 1 + (32-30)^2 \times 1 + (34-30)^2 \times 2 + (35-30)^2 \times 1}{15}$$

$$S^2 = \frac{25 + 16 + 9 + 4 + 2 + 0 + 1 + 4 + 32 + 25}{15} = \frac{118}{15} = 7,87^\circ$$

$$S = \sqrt{7,87} = 2,81^\circ$$

Les dades estan més escampades a Mumbai, perquè les mesures de dispersió són més grans.



**Activitat 15**

S'ha fet un estudi sobre el preu d'una ampolla d'aigua mineral d'una determinada marca a diferents establiments comercials. Heus aquí la gràfica que il·lustra els resultats obtinguts:

A partir d'aquesta gràfica, calcula totes les mesures de centralització i de dispersió que es desprenen d'aquest estudi.

Primer cal construir una taula de freqüències amb les freqüències absolutes:

Preu en cèntims	$n_i$	$N_i$
21	4	4
22	4	8
23	6	14
24	11	25
25	8	33
26	4	37
27	2	39
28	1	40
	$N = 40$	

$$\bar{X} = \frac{(21 \times 4) + (22 \times 4) + (23 \times 6) + (24 \times 11) + (25 \times 8) + (26 \times 4) + (27 \times 2) + (28 \times 1)}{40}$$

$$\bar{X} = \frac{84 + 88 + 138 + 264 + 200 + 104 + 54 + 28}{40} = \frac{960}{40} = 24 \text{ cèntims}$$

$$Me = 24 \text{ cèntims}$$

$$Mo = 24 \text{ cèntims}$$

$$R = 28 - 21 = 7 \text{ cèntims}$$

$$DM = \frac{(24-21) \times 4 + (24-22) \times 4 + (24-23) \times 6 + (24-24) \times 11 + (25-24) \times 8 + (26-24) \times 4 + (27-24) \times 2 + (28-24) \times 1}{40}$$

$$DM = \frac{12 + 8 + 6 + 0 + 8 + 8 + 6 + 4}{40} = \frac{52}{40} = 1,3 \text{ cèntims}$$

$$S^2 = \frac{(24-21)^2 \times 4 + (24-22)^2 \times 4 + (24-23)^2 \times 6 + (24-24)^2 \times 11 + (25-24)^2 \times 8 + (26-24)^2 \times 4 + (27-24)^2 \times 2 + (28-24)^2 \times 1}{40}$$

$$S^2 = \frac{36 + 16 + 6 + 0 + 8 + 16 + 18 + 16}{40} = \frac{116}{40} = 2,9 \text{ cèntims}$$

$$S = \sqrt{2,9} = 1,7 \text{ cèntims}$$

**Activitat 16**

S'ha fet un estudi sobre el consum de benzina dels vehicles dels trenta treballadors d'una empresa. El consum s'ha mesurat en litres per cada cent quilòmetres. Heus aquí la gràfica que il·lustra els resultats obtinguts:

A partir d'aquesta gràfica, calcula totes les mesures de centralització i de dispersió que es desprenen d'aquest estudi.

Primer cal construir una taula de freqüències amb les freqüències absolutes:

Consum (litres/100 km)	$n_i$	$N_i$
5	2	2
6	6	8
7	8	16
8	4	20
9	3	23
10	1	24
11	3	27
12	3	30
<b>N = 30</b>		

$$\bar{X} = \frac{(5 \times 2) + (6 \times 6) + (7 \times 8) + (8 \times 4) + (9 \times 3) + (10 \times 1) + (11 \times 3) + (12 \times 3)}{30}$$

$$\bar{X} = \frac{10 + 36 + 56 + 32 + 27 + 10 + 33 + 36}{30} = \frac{240}{30} = 8 \text{ litres}$$

$$Me = 7 \text{ litres}$$

$$Mo = 7 \text{ litres}$$

$$R = 12 - 5 = 7 \text{ litres}$$

$$DM = \frac{(8-5) \times 2 + (8-6) \times 6 + (8-7) \times 8 + (8-8) \times 4 + (9-8) \times 3 + (10-8) \times 1 + (11-8) \times 3 + (12-8) \times 3}{30}$$

$$DM = \frac{6 + 12 + 8 + 0 + 3 + 2 + 9 + 12}{30} = \frac{52}{30} = 1,73 \text{ litres}$$

$$S^2 = \frac{(8-5)^2 \times 2 + (8-6)^2 \times 6 + (8-7)^2 \times 8 + (8-8)^2 \times 4 + (9-8)^2 \times 3 + (10-8)^2 \times 1 + (11-8)^2 \times 3 + (12-8)^2 \times 3}{30}$$

$$S^2 = \frac{18 + 24 + 8 + 0 + 3 + 4 + 27 + 48}{30} = \frac{132}{30} = 4,4 \text{ litres}$$

$$S = \sqrt{4,4} = 2,1 \text{ litres}$$

**Activitat 1**

En Francesc té dues gallines, una blanca i una rossa, i compta els ous que pon cada una durant cada un dels dies d'una setmana. Heus aquí els resultats:

Gallina blanca: 4 7 6 6 6 7 6

Gallina rossa: 3 3 3 6 3 4 6

Calcula totes les mesures de centralització i de dispersió del nombre d'ous posats diàriament per cada gallina. Quina gallina ha post els ous de manera més dispersa?

**Gallina blanca**

Primer ordenem les dades: 4 6 6 6 6 7 7

$$\bar{X} = \frac{4+6+6+6+6+7+7}{7} = \frac{42}{7} = 6 \text{ ous}$$

$$Me = 6 \text{ ous}$$

$$Mo = 6 \text{ ous}$$

$$R = 7 - 4 = 3 \text{ ous}$$

$$DM = \frac{(6-4)+(6-6)+(6-6)+(6-6)+(6-6)+(7-6)+(7-6)}{7}$$

$$DM = \frac{2+0+0+0+0+1+1}{7} = \frac{4}{7} = 0,57 \text{ ous}$$

$$S^2 = \frac{(6-4)^2+(6-6)^2+(6-6)^2+(6-6)^2+(6-6)^2+(7-6)^2+(7-6)^2}{7}$$

$$S^2 = \frac{4+0+0+0+0+1+1}{7} = \frac{6}{7} = 0,86 \text{ ous}$$

$$S = \sqrt{0,86} = 0,93 \text{ ous}$$

**Gallina rossa**

Primer ordenem les dades: 3 3 3 3 4 6 6

$$\bar{X} = \frac{3+3+3+3+4+6+6}{7} = \frac{28}{7} = 4 \text{ ous}$$

$$Me = 3 \text{ ous}$$

$$Mo = 3 \text{ ous}$$

$$R = 6 - 3 = 3 \text{ ous}$$

$$DM = \frac{(4-3)+(4-3)+(4-3)+(4-3)+(4-4)+(6-4)+(6-4)}{7}$$

$$DM = \frac{1+1+1+1+0+2+2}{7} = \frac{8}{7} = 1,14 \text{ ous}$$

$$S^2 = \frac{(4-3)^2+(4-3)^2+(4-3)^2+(4-3)^2+(4-4)^2+(6-4)^2+(6-4)^2}{7}$$

$$S^2 = \frac{1+1+1+1+0+4+4}{7} = \frac{12}{7} = 1,71 \text{ ous}$$

$$S = \sqrt{1,71} = 1,31 \text{ ous}$$

La gallina rossa és la que ha post els ous de manera més dispersa, ja que així ho indiquen les mesures de dispersió.

**Activitat 2**

Un estudi sobre el nombre de nens que han tingut dos grups de dones de més de cinquanta anys ha donat aquests resultats:

Grup de Bangkok	2	1	3	3	2	2	3	2	3	0
	4	3	5	2	3	3	1	3		
Grup de Bali	6	7	4	4	1	9	3	7	8	3
	5	7	8	7	4	3	2	2		

Calcula totes les mesures de centralització i de dispersió per a cada grup. A quin grup hi ha les dades més escampades?

Atès l'elevat nombre de dades, treballarem amb taules de freqüències.

**Grup de Bangkok**

Fills	$n_i$	$N_i$
0	1	1
1	2	3
2	5	8
3	8	16
4	1	17
5	1	18
$N = 18$		

$$\bar{X} = \frac{(0 \times 1) + (1 \times 2) + (2 \times 5) + (3 \times 8) + (4 \times 1) + (5 \times 1)}{18} = \frac{0 + 2 + 10 + 24 + 4 + 5}{18} = \frac{45}{18} = 2,5 \text{ fills}$$

$$Me = 3 \text{ fills}$$

$$Mo = 3 \text{ fills}$$

$$R = 5 - 0 = 5 \text{ fills}$$

$$DM = \frac{(2,5 - 0) \times 1 + (2,5 - 1) \times 2 + (2,5 - 2) \times 5 + (3 - 2,5) \times 8 + (4 - 2,5) \times 1 + (5 - 2,5) \times 1}{18}$$

$$DM = \frac{2,5 + 3 + 2,5 + 4 + 1,5 + 2,5}{18} = \frac{16}{18} = 0,89 \text{ fills}$$

$$S^2 = \frac{(2,5 - 0)^2 \times 1 + (2,5 - 1)^2 \times 2 + (2,5 - 2)^2 \times 5 + (3 - 2,5)^2 \times 8 + (4 - 2,5)^2 \times 1 + (5 - 2,5)^2 \times 1}{18}$$

$$S^2 = \frac{6,25 + 4,5 + 1,25 + 2 + 2,25 + 6,25}{18} = \frac{22,5}{18} = 1,25 \text{ fills}$$

$$S = \sqrt{1,25} = 1,19 \text{ fills}$$

## Grup de Bali

Fills	$n_i$	$N_i$
1	1	1
2	2	3
3	3	6
4	3	9
5	1	10
6	1	11
7	4	15
8	2	17
9	1	18
N = 18		

$$\bar{X} = \frac{(1 \times 1) + (2 \times 2) + (3 \times 3) + (4 \times 3) + (5 \times 1) + (6 \times 1) + (7 \times 4) + (8 \times 2) + (9 \times 1)}{18}$$

$$\bar{X} = \frac{1+4+9+12+5+6+28+16+9}{18} = \frac{90}{18} = 5 \text{ fills}$$

$$Me = \frac{4+5}{2} = 4,5 \text{ fills}$$

$$Mo = 7 \text{ fills}$$

$$R = 9 - 1 = 8 \text{ fills}$$

$$DM = \frac{(5-1) \times 1 + (5-2) \times 2 + (5-3) \times 3 + (5-4) \times 3 + (5-5) \times 1 + (6-5) \times 1 + (7-5) \times 4 + (8-5) \times 2 + (9-5) \times 1}{18}$$

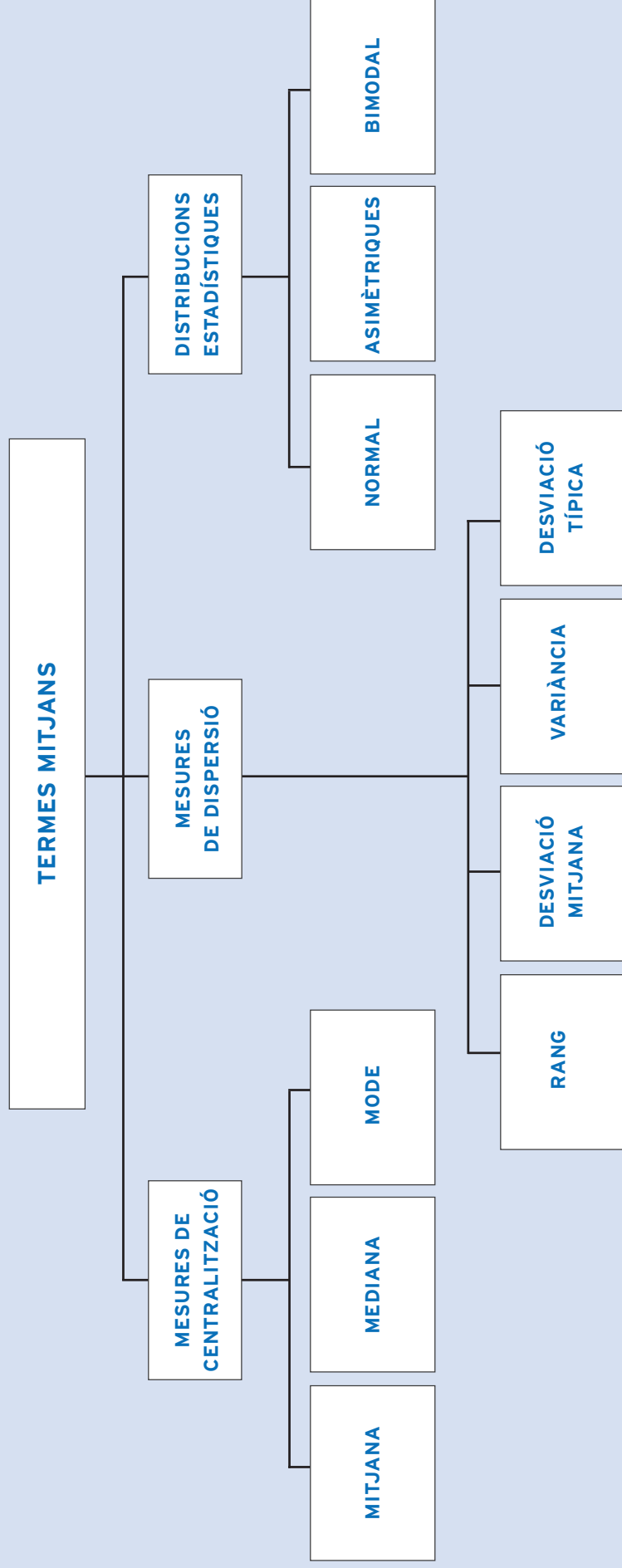
$$DM = \frac{4+6+6+3+0+1+8+6+4}{18} = \frac{38}{18} = 2,11 \text{ fills}$$

$$S^2 = \frac{(5-1)^2 \times 1 + (5-2)^2 \times 2 + (5-3)^2 \times 3 + (5-4)^2 \times 3 + (5-5)^2 \times 1 + (6-5)^2 \times 1 + (7-5)^2 \times 4 + (8-5)^2 \times 2 + (9-5)^2 \times 1}{18}$$

$$S^2 = \frac{16+18+12+3+0+1+16+18+16}{18} = \frac{100}{18} = 5,56 \text{ fills}$$

$$S = \sqrt{5,56} = 2,36 \text{ fills}$$

Les dades del grup de Bali són més disperses, ja que així ho indiquen les mesures de dispersió.



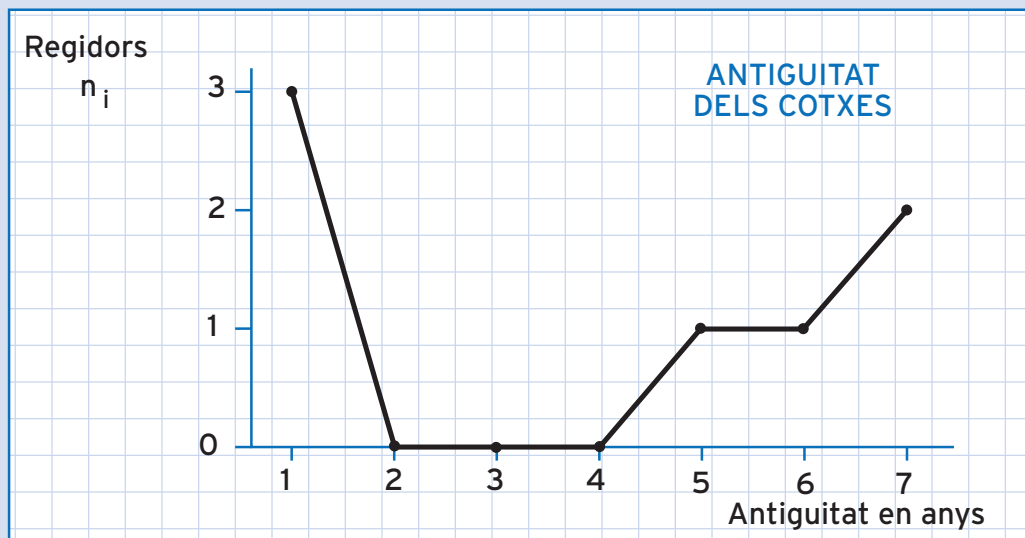
# com ho porto?

Omple la graella següent posant una creu on correspongui:

En acabar la unitat sóc capaç de:	Bé	Regular	Malament
Explicar el significat de la mitjana i calcular-ne el valor a partir d'una col·lecció de dades.			
Explicar el significat de la mediana i calcular-ne el valor a partir d'una col·lecció de dades.			
Explicar el significat del mode i calcular-ne el valor a partir d'una col·lecció de dades.			
Explicar el significat del rang i calcular-ne el valor a partir d'una col·lecció de dades.			
Explicar el significat de la desviació mitjana i calcular-ne el valor a partir d'una col·lecció de dades.			
Explicar el significat de la variància i calcular-ne el valor a partir d'una col·lecció de dades.			
Explicar el significat de la desviació típica i calcular-ne el valor a partir d'una col·lecció de dades.			
Descriure les característiques de les distribucions estadístiques normal, bimodal i asimètrica.			
Interpretar la informació que ens aporten els termes mitjans.			

**Activitat 1**

Es demana als set regidors d'una població del Maresme quina és l'antiguitat, en anys, del cotxe que condueixen habitualment. Amb els resultats obtinguts s'elabora un polígon de freqüències:



- A. Calcula totes les mesures de centralització i de dispersió de les dades recollides.



B. Amb les dades recollides, si entrevistem un dels regidors a l'atzar, quina probabilitat hi ha que l'antiguitat del seu cotxe sigui d'un any? Aquesta probabilitat és teòrica o experimental? Per què?

C. Si volem calcular la probabilitat que un conductor de Catalunya, entrevistat a l'atzar, tingui un vehicle de dos anys d'antiguitat, quina hauria d'ésser la mida de la mostra si s'accepta un error inferior al 4 % i un nivell de confiança del 95 %?

**Activitat 2**

El darrer novembre, les temperatures mínimes recollides en graus centígrads, a Kathmandu i a Chennai, van ésser:

Temperatures mínimes a Kathmandu

6	6	3	0	1	2	7	9	9	10
9	9	11	7	6	6	6	5	5	4
2	1	0	1	2	3	4	6	6	4

Temperatures mínimes a Chennai

12	12	15	15	10	12	14	15	15	13
10	8	8	9	10	13	15	15	16	15
14	15	16	15	16	13	10	13	12	14

- Construeix una taula per a cada conjunt de dades recollides on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .
- Dibuixa un diagrama de barres amb les dades enregistrades a Kathmandu i un diagrama de sectors circular amb les dades enregistrades a Chennai.
- Calcula totes les mesures de centralització i de dispersió per als dos conjunts de dades. A quina població hi ha dades més escampades?



**Activitat 1**

Es demana als set regidors d'una població del Maresme quina és l'antiguitat, en anys, del cotxe que condueixen habitualment. Amb els resultats obtinguts s'elabora un polígon de freqüències.

- Calcula totes les mesures de centralització i de dispersió de les dades recollides.
- Amb les dades recollides, si entrevistem un dels regidors a l'atzar, quina probabilitat hi ha que l'antiguitat del seu cotxe sigui d'un any? Aquesta probabilitat és teòrica o experimental? Per què?
- Si volem calcular la probabilitat que un conductor de Catalunya, entrevistat a l'atzar, tingui un vehicle de dos anys d'antiguitat, quina hauria d'ésser la mida de la mostra si s'accepta un error inferior al 4 % i un nivell de confiança del 95 %?

A. A partir de la gràfica, podem ordenar les dades recollides:

1 1 1 5 6 7 7

$$\bar{X} = \frac{1+1+1+5+6+7+7}{7} = \frac{28}{7} = 4 \text{ anys}$$

$$Me = 5 \text{ anys}$$

$$Mo = 1 \text{ any}$$

$$R = 7 - 1 = 6 \text{ anys}$$

$$DM = \frac{(4-1)+(4-1)+(4-1)+(5-4)+(6-4)+(7-4)+(7-4)}{7}$$

$$DM = \frac{3+3+3+1+2+3+3}{7} = \frac{18}{7} = 2,57 \text{ anys}$$

$$S^2 = \frac{(4-1)^2+(4-1)^2+(4-1)^2+(5-4)^2+(6-4)^2+(7-4)^2+(7-4)^2}{7}$$

$$S^2 = \frac{9+9+9+1+4+9+9}{7} = \frac{50}{7} = 5,56 \text{ anys}$$

$$S = \sqrt{5,56} = 2,36 \text{ anys}$$

B. La probabilitat que l'antiguitat del cotxe d'un dels regidors entrevistats a l'atzar sigui d'un any és:

$$p_o(1any) = f_i(1any) = \frac{n_i}{N} = \frac{3}{7} = 0,43$$

Aquesta probabilitat obtinguda és experimental, ja que ens ha fet falta realitzar un estudi empíric per tal de determinar-la.

C. Per calcular la probabilitat que un conductor de Catalunya, entrevistat a l'atzar, tingui un vehicle de dos anys d'antiguitat, acceptant un error inferior al 4 % ( $e = 0,04$ ), cal aplicar la fórmula:

$$n = \frac{1}{e^2} = \frac{1}{0,04^2} = \frac{1}{0,0016} = 625$$

Caldrà doncs entrevistar 625 conductors.

### Activitat 2

El darrer novembre, les temperatures mínimes recollides en graus centígrads, a Kathmandu i a Chennai, van ésser:

Temperatures mínimes a Kathmandu

6 6 3 0 1 2 7 9 9 10  
9 9 11 7 6 6 6 5 5 4  
2 1 0 1 2 3 4 6 6 4

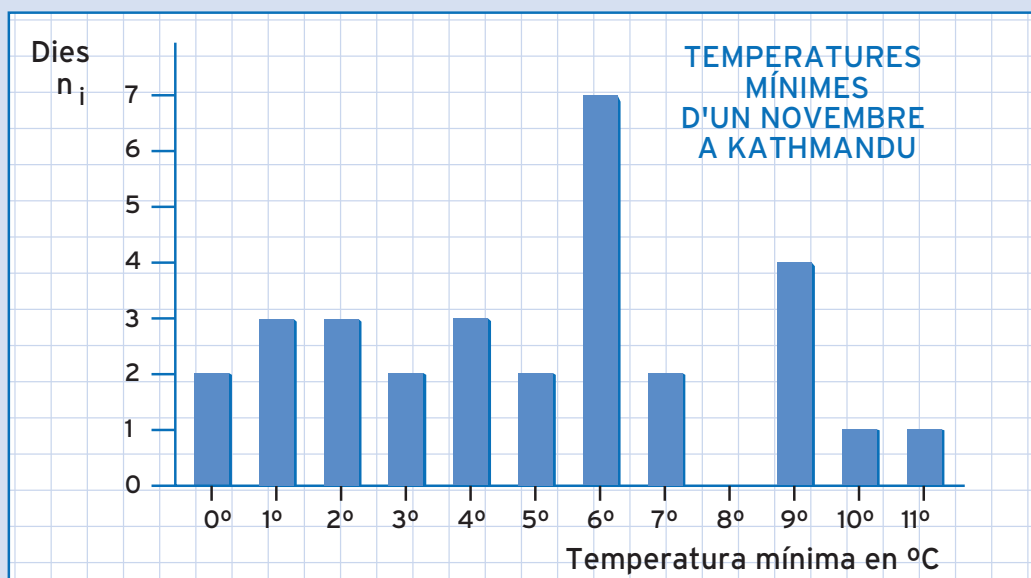
Temperatures mínimes a Chennai

12 12 15 15 10 12 14 15 15 13  
10 8 8 9 10 13 15 15 16 15  
14 15 16 15 16 13 10 13 12 14

- Construeix una taula per a cada conjunt de dades recollides on hi hagi totes les freqüències estudiades:  $n_i$ ,  $N_i$ ,  $f_i$ ,  $F_i$ .
- Dibuixa un diagrama de barres amb les dades enregistrades a Kathmandu i un diagrama de sectors circular amb les dades enregistrades a Chennai.
- Calcula totes les mesures de centralització i de dispersió per als dos conjunts de dades. A quina població hi ha dades més escampades?

#### Kathmandu

Temperatura mínima	$n_i$	$N_i$	$f_i$	$F_i$
0°	2	2	0,067	0,067
1°	3	5	0,1	0,17
2°	3	8	0,1	0,27
3°	2	10	0,067	0,33
4°	3	13	0,1	0,43
5°	2	15	0,067	0,5
6°	7	22	0,23	0,73
7°	2	24	0,067	0,8
8°	0	24	0	0,8
9°	4	28	0,13	0,93
10°	1	29	0,033	0,97
11°	1	30	0,033	1
	N = 30		Total $f_i = 0,9994$	



$$\bar{X} = \frac{(0 \times 2) + (1 \times 3) + (2 \times 3) + (3 \times 2) + (4 \times 3) + (5 \times 2) + (6 \times 7) + (7 \times 2) + (8 \times 0) + (9 \times 4) + (10 \times 1) + (11 \times 1)}{30}$$

$$\bar{X} = \frac{0 + 3 + 6 + 6 + 12 + 10 + 42 + 14 + 0 + 36 + 10 + 11}{30} = \frac{150}{30} = 5^\circ$$

$$Me = \frac{5 + 6}{2} = \frac{11}{2} = 5,5^\circ$$

$$Mo = 6^\circ$$

$$R = 11 - 0 = 11^\circ$$

$$DM = \frac{(5-0) \times 2 + (5-1) \times 3 + (5-2) \times 3 + (5-3) \times 2 + (5-4) \times 3 + (5-5) \times 2 + (6-5) \times 7 + (7-5) \times 2 + (9-5) \times 4 + (10-5) \times 1 + (11-5) \times 1}{30}$$

$$DM = \frac{10 + 12 + 9 + 4 + 3 + 0 + 7 + 4 + 16 + 5 + 6}{30} = \frac{76}{30} = 2,53^\circ$$

$$S^2 = \frac{(5-0)^2 \times 2 + (5-1)^2 \times 3 + (5-2)^2 \times 3 + (5-3)^2 \times 2 + (5-4)^2 \times 3 + (5-5)^2 \times 2 + (6-5)^2 \times 7 + (7-5)^2 \times 2 + (9-5)^2 \times 4 + (10-5)^2 \times 1 + (11-5)^2 \times 1}{30}$$

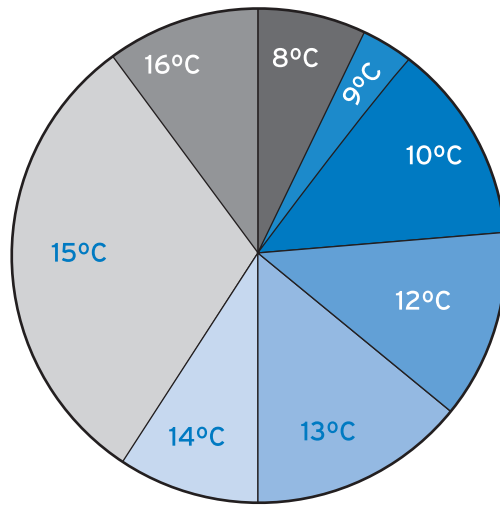
$$S^2 = \frac{50 + 48 + 27 + 8 + 3 + 0 + 7 + 8 + 64 + 25 + 36}{30} = \frac{276}{30} = 9,2^\circ$$

$$S = \sqrt{9,2} = 3,03^\circ$$

## Chennai

Temperatura mínima	n <sub>i</sub>	N <sub>i</sub>	f <sub>i</sub>	F <sub>i</sub>	angle
8°	2	2	0,067	0,067	24°
9°	1	3	0,033	0,1	12°
10°	4	7	0,13	0,23	48°
11°	0	7	0	0,23	0°
12°	4	11	0,13	0,37	48°
13°	4	15	0,13	0,5	48°
14°	3	18	0,1	0,6	36°
15°	9	27	0,3	0,9	108°
16°	3	30	0,1	1	36°
	N = 30		Total f <sub>i</sub> = 0,99		

### TEMPERATURES MÍNIMES D'UN NOVENBRE A CHENAI



$$\bar{X} = \frac{(8 \times 2) + (9 \times 1) + (10 \times 4) + (12 \times 4) + (13 \times 4) + (14 \times 3) + (15 \times 9) + (16 \times 3)}{30}$$

$$\bar{X} = \frac{16 + 9 + 40 + 48 + 52 + 42 + 135 + 48}{30} = \frac{390}{30} = 13^\circ$$

$$Me = \frac{13 + 14}{2} = \frac{27}{2} = 13,5^\circ$$

$$Mo = 15^\circ$$

$$R = 16 - 8 = 8^\circ$$

$$DM = \frac{(13-8) \times 2 + (13-9) \times 1 + (13-10) \times 4 + (13-12) \times 4 + (13-13) \times 4 + (14-13) \times 3 + (15-13) \times 9 + (16-13) \times 3}{30}$$

$$DM = \frac{10 + 4 + 12 + 4 + 0 + 3 + 18 + 9}{30} = \frac{60}{30} = 2^\circ$$

$$S^2 = \frac{(13-8)^2 \times 2 + (13-9)^2 \times 1 + (13-10)^2 \times 4 + (13-12)^2 \times 4 + (13-13)^2 \times 4 + (14-13)^2 \times 3 + (15-13)^2 \times 9 + (16-13)^2 \times 3}{30}$$

$$S^2 = \frac{50 + 16 + 36 + 4 + 0 + 3 + 36 + 27}{30} = \frac{172}{30} = 5,73^\circ$$

$$S = \sqrt{5,73} \approx 2,39^\circ$$

Les temperatures mínimes recollides estan més escampades a Kathmandu, ja que les mesures de dispersió són més grans.





# Mòdul opcional Estadística

# 19

Àmbit de les Matemàtiques, de la Ciència  
i de la Tecnologia



Generalitat de Catalunya  
Departament d'Educació i Universitats



G<sub>2</sub> R<sub>1</sub> A<sub>1</sub> D<sub>2</sub> U<sub>1</sub> ï S<sub>1</sub>

ARA EN SECUNDÀRIA