



A ESTATÍSTICA



Diagrama. Licenza de contido de Pixabay.



ÍNDICE

A ESTATÍSTICA

1. A ESTATÍSTICA.....	1
1.1 Poboación e mostra.....	1
1.2 Variables estatísticas.....	1
<i>Exercicios</i>	2
2. RECONTOS E TÁBOAS.....	4
<i>Exercicios</i>	5
3. REPRESENTACIÓNS GRÁFICAS.....	6
<i>Exercicios</i>	7
4. PARÁMETROS ESTATÍSTICOS.....	8
4.1 Parámetros de centralización.....	8
4.2 Parámetros de dispersión.....	8
4.3 Comparación de dous conxuntos de datos.....	8
<i>Exercicios</i>	9
5. FERRAMENTAS TECNOLÓXICAS.....	10
SOLUCIÓNS.....	12



1. A ESTATÍSTICA

A **estatística** é a parte das matemáticas que se encarga de estudar un fenómeno mediante a recollida, análise e interpretación de datos coa finalidade de extraer conclusións.

Un estudo estatístico pódese facer dun grupo enteiro de elementos ou persoas, ou dunha parte representativa do grupo.

Nun estudo estatístico podemos diferenciar varias fases:

- Selección das características que interesa estudar.
- Recollida de datos
- Organización e representación gráfica
- Análise de datos
- Obtención de conclusións

1.1 Poboación e mostra

A poboación é o conxunto de elementos sobre os que se quere estudar unha determinada característica.

Unha mostra é unha parte representativa da poboación. Para que unha mostra sexa representativa da poboación é necesario que sexa:

- Aleatoria: cada individuo da mostra ten a mesma probabilidade de ser elixido.
- Proporcional: os individuos da mostra conservan a mesma proporción que na poboación.

Un individuo é cada un dos elementos da poboación ou da mostra.

1.2 Variables estatísticas

Un estudo estatístico faise para analizar algunha propiedade ou característica, por exemplo, a idade, a cor de ollos, a duración dunha lámpada, a lectura preferida... Estas características ou propiedades que posúe un grupo de persoas ou elementos chámanse **variables estatísticas**. Clasifícanse en:

- **Variable cualitativa**, é a que non se pode expresar mediante números.
- **Variable cuantitativa**, é a que se expresa mediante números. No grupo das variables cuantitativas, pódense distinguir:
 - **Discreta**, é a que toma un número finito de valores numéricos.
 - **Continua**, é a que toma un número infinito de valores.



EXERCICIOS

Exercicio 1

Indica que variables estatísticas son cualitativas ou cuantitativas. Se son cuantitativas, indica se son discretas ou continuas.

- Estatura
- Cor de pelo
- Número de irmáns
- Deporte preferido
- Estado civil
- Provincia de residencia
- Número de veciños dun edificio
- Profesión
- Consumo de gasolina cada 100 km
- A lonxitude dos parafusos que produce unha máquina

Exercicio 2

Queremos realizar un estudo estatístico da talla de calzado que usan os 2.715.424 habitantes de Galicia. Preguntouse a 100 persoas galegas.

- a) Cal é a poboación?
- b) Cal é a mostra? Que tamaño ten?
- c) Que variable se vai estudar? De que tipo é?

Exercicio 3

Nunha empresa traballan 530 homes e 470 mulleres. Quérese facer unha enquisa representativa a 100 persoas traballadoras da empresa. Cantos homes e mulleres deberá ter a mostra?



Exercicio 4

A distribución por idades dos habitantes dunha localidade galega é a que se indica na seguinte táboa.

Idade (anos)	Nº de habitantes
0 a 18	700
18 a 25	1000
25 a 65	3900
Máis de 65	1100

Se se quere preguntar a 120 persoas maiores de 18 anos, a cantas persoas de cada grupo se debe preguntar para que a mostra sexa representativa?

2. RECONTOS E TÁBOAS

Para poder estudar máis comodamente os datos dunha variable estatística hai que ordenalos. Para ordenalos empregaremos **táboas estatísticas ou táboas de frecuencias**.

As frecuencias son cantidades que permiten organizar a información sobre unha característica.

x_i	Reconto	f_i	F_i	h_i	H_i	Porcentaxe	Porcentaxe acumulada
-------	---------	-------	-------	-------	-------	------------	----------------------

Nunha táboa de frecuencias recolleemos:

- **Frecuencia absoluta**, $f_{i'}$, número de veces que se repite un dato.
- **Frecuencia relativa**, $h_{i'}$, cociente entre a frecuencia absoluta, $f_{i'}$, e o número total de datos, N .
- **Frecuencia absoluta acumulada**, $F_{i'}$, suma das frecuencias absolutas ata ese dato.
- **Frecuencia relativa acumulada**, $H_{i'}$, suma das frecuencias relativas ata ese dato.
- **Porcentaxe**, expresión da frecuencia relativa coma porcentaxe.
- **Porcentaxe acumulada**, suma das porcentaxes ata ese dato.







Para construír unha táboa de frecuencias,

- se a variable é discreta (ten poucos valores distintos), colócanse nunha columna os distintos valores da variable en orde crecente e, na outra, o reconto de cada un deles.
- Se a variable é continua (ten moitos valores distintos), os datos agrúpanse en intervalos da mesma amplitude. O punto medio de cada **intervalo** é a **marca de clase**.

 EXERCICIOS

Exercicio 5

Nun ximnasio preguntouse aos usuarios polo seu deporte favorito. Os resultados foron os seguintes:

					
Fútbol:150	Baloncesto: 70	Atletismo: 130	Tenis: 60	Natación: 40	Ciclismo: 50

Completa a táboa seguinte:

Deporte	f_i	F_i	h_i	H_i	Porcentaxe	Porcentaxe acumulada
Fútbol						
Baloncesto						
Atletismo						
Tenis						
Natación						
Ciclismo						

Exercicio 6

A información que figura a continuación obtívose nunha asociación cultural que ten 25 persoas asociadas. Preguntóuselles polo número de libros que len á semana.

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

Fai unha táboa de frecuencias.

Que porcentaxe de persoas da asociación len máis de 2 libros á semana?

3. REPRESENTACIÓNS GRÁFICAS

Aínda que as táboas conteñen toda a información, expresala nun gráfico faina máis clara. Os gráficos estatísticos máis frecuentes son:

- **Diagrama de barras:** é útil para representar datos cuantitativos non agrupados e datos cualitativos. Debúxanse rectángulos sobre cada dato cunha altura igual a cada frecuencia absoluta.
- **Histograma:** é útil para representar datos agrupados. Debúxanse rectángulos sobre cada dato, sendo a base de cada rectángulo a amplitude da clase, e a altura, a frecuencia.
- **Diagrama de sectores:** é útil para calquera tipo de datos. Divídese un círculo en sectores cuxo ángulo central é proporcional á frecuencia dos datos.

O diagrama de barras e o histograma, cando representan variables cuantitativas, poden incluír o **polígono de frecuencias** que se forma ao unir o punto medio das barras.

Exemplo:

Debuxa os diagramas de barras e sectores correspondentes á seguinte táboa:

Variable	A	B	C	D	E
Frecuencia	8	3	4	5	10

Diagrama de barras: Establecemos na escala vertical as diferentes frecuencias e construímos os rectángulos sobre cada dato con altura igual á súa frecuencia.

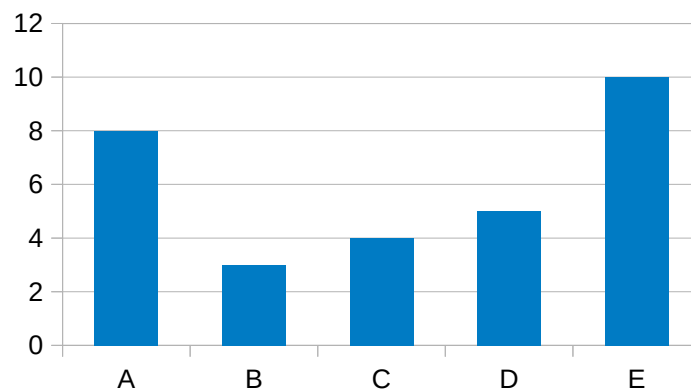
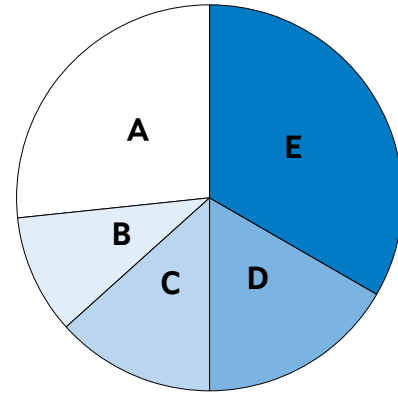


Diagrama de sectores: Calculamos os graos que lle corresponden a cada dato, como o número de datos é 30, a este valor lle corresponden 360° , logo a cada un lle corresponden $\frac{360}{30} = 12^\circ$.

Variable	Frecuencia	Graos
A	8	$8 \cdot 12^\circ = 96^\circ$
B	3	$3 \cdot 12^\circ = 36^\circ$
C	4	$4 \cdot 12^\circ = 48^\circ$
D	5	$5 \cdot 12^\circ = 60^\circ$
E	10	$10 \cdot 12^\circ = 120^\circ$



EXERCICIOS

Exercicio 7

Nunha concello con 1200 habitantes, 756 naceron no propio concello, 288 noutro concello galego, e 156 no estranxeiro.

- Que tipo de gráfico estatístico sería axeitado para representar estes datos?
- Representa os datos mediante un diagrama de barras e un diagrama de sectores.

Exercicio 8

Representa mediante un histograma, co seu correspondente polígono de frecuencias, a porcentaxe de poboación activa por idades nunha cidade.

Idades	[16, 23)	[23, 30)	[30, 37)	[37, 44)	[44, 51)	[51, 58)	[58, 65)	[65, 72)
Porcentaxe	4	6	17	25	22	14	8	4

4. PARÁMETROS ESTADÍSTICOS

Despois de obter os datos dunha distribución, necesitamos sintetizar a información para a súa posterior análise. Para iso, obteremos os parámetros estatísticos.

4.1 Parámetros de centralización

Os parámetros de centralización son cantidades que permiten facerse unha idea xeral sobre un conxunto de datos. Son a media aritmética, a mediana e a moda.

- A **media aritmética**, \bar{x} , calcúlase mediante a expresión:

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

- A **mediana**, Me , é o valor dos datos ordenados que está no medio. Se o número de datos é par, a mediana é a media aritmética dos dous valores centrais.
- A **moda**, Mo , é o dato que ten maior frecuencia absoluta. Se a variable é continua é o intervalo modal.

4.2 Parámetros de dispersión

Os parámetros de dispersión son cantidades que miden se os datos están máis ou menos separados da media aritmética. Son a varianza e a desviación típica.

- A **varianza**, σ^2 , dun conxunto de datos discretos calcúlase mediante a expresión:

$$\sigma^2 = \frac{x_1^2 \cdot f_1 + x_2^2 \cdot f_2 + \dots + x_n^2 \cdot f_n}{f_1 + f_2 + \dots + f_n} - \bar{x}^2 = \frac{\sum f_i \cdot x_i^2}{N} - \bar{x}^2$$

- A **desviación típica**, σ , é a raíz cadrada da varianza: $\sigma = \sqrt{(\sigma^2)}$. A súa vantaxe é que mide a dispersión nas mesmas unidades de medida que a variable estatística.

4.3 Comparación de dous conxuntos de datos

Para comparar se dúas distribucións de datos estatísticos son máis ou menos dispersas, emprégase o **coeficiente de variación**, CV :

$$CV = \frac{\sigma}{\bar{x}}$$

EXERCICIOS

Exercicio 9

Calcula os parámetros de centralización das idades dos asistentes a dúas festas:

a) 1, 3, 6, 7, 8, 10, 13, 14, 34, 37, 40, 40, 65, 66, 72

b)

Idade	14	15	16	17	18
Frecuencia	5	7	9	3	2

Exercicio 10

Calcula os parámetros de dispersión do número de urxencias médicas atendidas nun centro de saúde dúas semanas:

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

Exercicio 11

A media das cualificacións dos opositores do tribunal 1 de Matemáticas é $\bar{x}=7,6$, cunha desviación típica $\sigma_{T_1}=1,2$. No tribunal 2, os resultados son é $\bar{x}=6,2$ e $\sigma_{T_2}=1,2$. Analiza os resultados.

5. FERRAMENTAS TECNOLÓXICAS

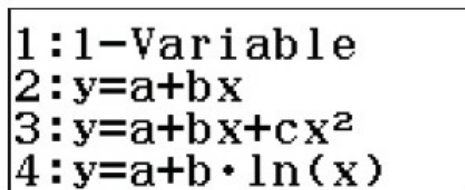
Para realizar os cálculos estatísticos podemos empregar diferentes ferramentas:

- **Calculadoras científicas.**

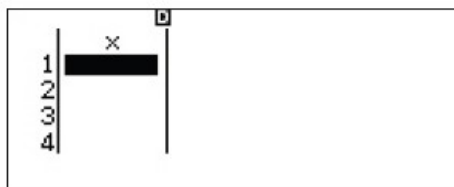
Na calculadora científica para poder traballar coa estatística hai que indicalo desde o menú principal, use as teclas de frecha para resaltar a icona Estatística e, a continuación, preme $\boxed{=}$ ou ben $\boxed{6}$.



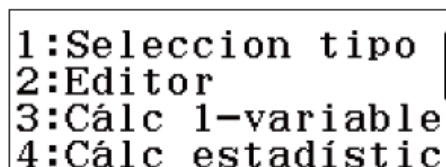
No menú que aparece, seleccione 1 para estatísticas de "1-Variable".



Aparece unha pantalla de introdución de datos.

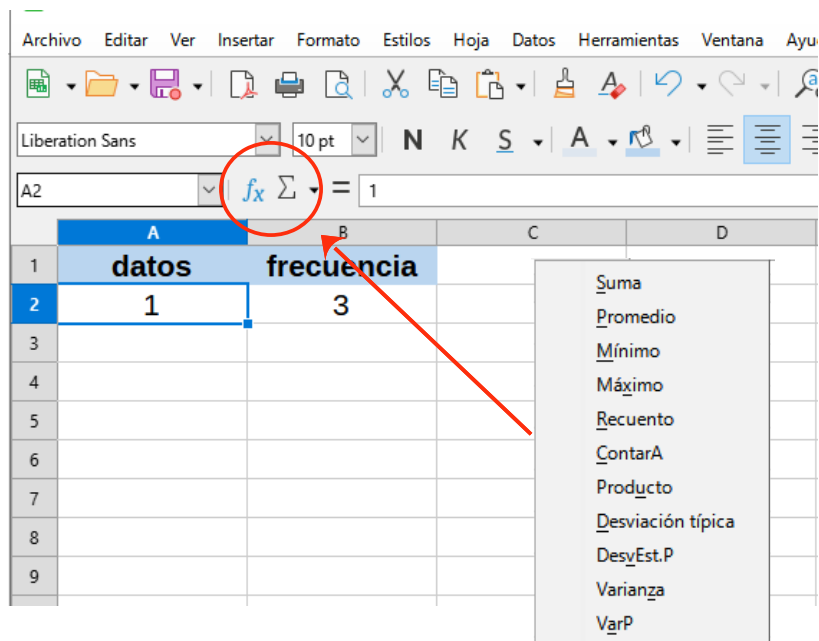


Para calcular as estatísticas dunha variable dun conxunto de datos, presionase $\boxed{\text{OPTN}}$ $\boxed{3}$ (Cálc. 1-variable).

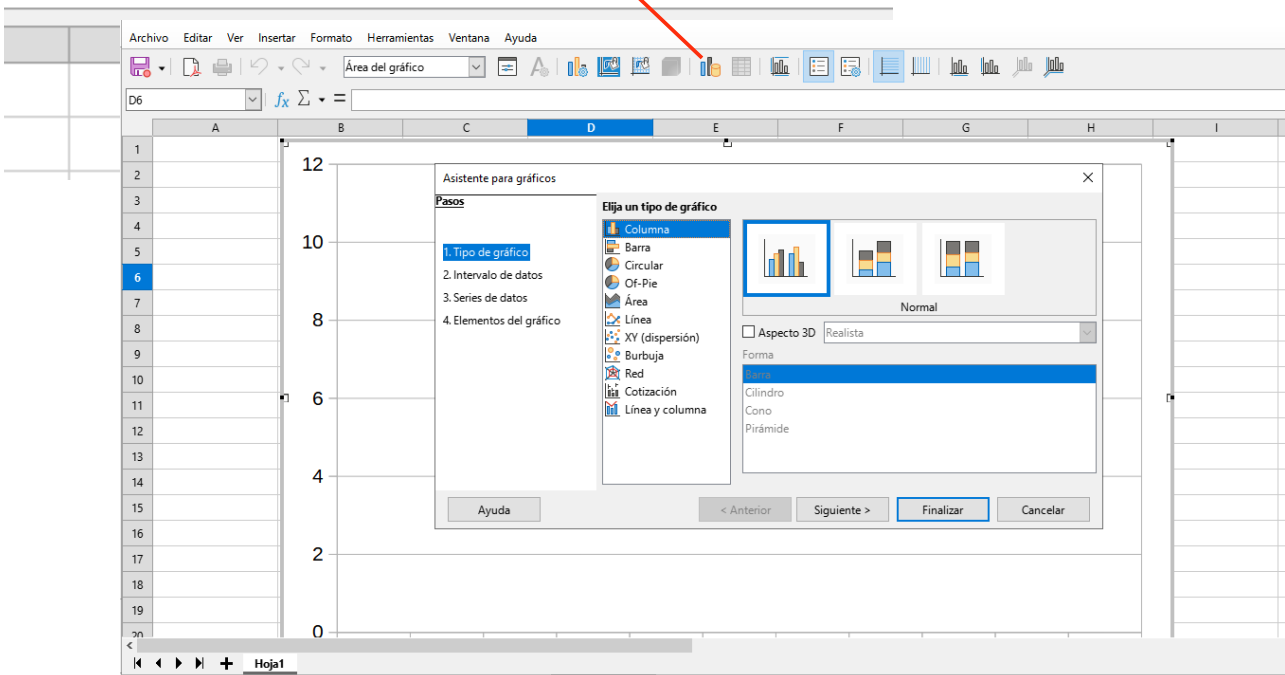
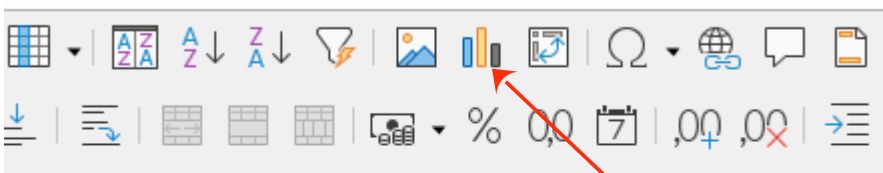


- **Follas de cálculo.**

Para introducir datos estatísticos na folla de cálculo, primeiro débense organizar os datos en columnas e filas. Logo, pódense usar as funcións da folla de cálculo para realizar cálculos estatísticos básicos como medias, desviacións....



Pódense crear varios tipos de gráficos estadísticos para visualizar datos e realizar análises. Para crear un gráfico, selecciónanse os datos, vaise á pestana "Insertar" e escóllese o tipo de gráfico desexado (columnas, liñas, circulares etc.)/ etc.)





SOLUCIÓNS

Exercicio 1

Indica que variables estatísticas son cualitativas ou cuantitativas. Se son cuantitativas, indica se son discretas ou continuas.

- Variables cualitativas: cor de pelo, deporte preferido, estado civil, provincia de residencia, profesión.
- Variables cuantitativas:
 - Discretas: número de irmáns, número de veciños dun edificio
 - Continuas: estatura, consumo de gasolina cada 100 km, lonxitude dos parafusos que produce unha máquina.

Exercicio 2

Queremos realizar un estudio estadístico do número de calzado que usan os 2.715.424 habitantes de Galicia. Preguntouse a 100 persoas galegas.

- a) Cal é a poboación? Os 2.715.424 habitantes de Galicia.
- b) Cal é a mostra? Que tamaño ten? A mostra son as 100 persoas galegas ás que se lle preguntou. O tamaño da mostra é 100.
- c) Que variable se vai estudar? De que tipo é? O número de calzado que usan. É unha variable cuantitativa discreta.

Exercicio 3

Nunha empresa traballan 530 homes e 470 mulleres. Quérese facer unha enquisa representativa a 100 persoas traballadoras da empresa. Cantos homes e mulleres deberá ter a mostra?

Hai que repartir as 100 persoas traballadoras en dúas cantidades proporcionais a 530 e 470.

$$530 + 470 = 1000$$

$$\frac{x}{100} = \frac{530}{1000} \Rightarrow x = \frac{530 \cdot 100}{1000} = 53 \text{ homes}$$

$$\frac{x}{100} = \frac{470}{1000} \Rightarrow x = \frac{470 \cdot 100}{1000} = 47 \text{ mulleres}$$

Exercicio 4

A distribución por idades dos habitantes dunha localidade galega é a que se indica na seguinte táboa.

Idade (anos)	Nº de habitantes
0 a 18	700
18 a 25	1000
25 a 65	3900
Máis de 65	1100







Se se quere preguntar a 134 persoas, a cantas persoas de cada grupo se debe preguntar para que a mostra sexa representativa?

Hai que repartir as 134 persoas traballadoras en cantidades proporcionais a 700, 1000, 3900 e 1100.

Idade (anos)	Nº de habitantes	Persoas de cada grupo ás que preguntar
0 a 18	700	$\frac{700 \cdot 134}{6700} = 14$
18 a 25	1000	$\frac{1000 \cdot 134}{6700} = 20$
25 a 65	3900	$\frac{3900 \cdot 134}{6700} = 78$
Máis de 65	1100	$\frac{1100 \cdot 134}{6700} = 22$
Total	6700	134

Exercicio 5

Nun ximnasio preguntouse aos usuarios polo seu deporte favorito. Os resultados foron os seguintes:

					
Fútbol:150	Baloncesto: 70	Atletismo: 130	Tenis: 60	Natación: 40	Ciclismo: 50

Completa a táboa seguinte:

Deporte	f_i	F_i	h_i	H_i	Porcentaxe	Porcentaxe acumulada
Fútbol	150	150	$\frac{150}{500}=0,3$	0,3	$0,3 \cdot 100=30\%$	30%
Baloncesto	70	$150+70=220$	$\frac{70}{500}=0,14$	$0,3+0,14=0,44$	$0,14 \cdot 100=14\%$	$30+14=44\%$
Atletismo	130	$220+130=350$	$\frac{130}{500}=0,26$	$0,44+0,26=0,7$	$0,26 \cdot 100=26\%$	$44+26=70\%$
Tenis	60	$350+60=410$	$\frac{60}{500}=0,12$	$0,7+0,12=0,82$	$0,12 \cdot 100=12\%$	$70+12=82\%$
Natación	40	$410+40=450$	$\frac{40}{500}=0,08$	$0,82+0,08=0,9$	$0,08 \cdot 100=8\%$	$82+0,08=90\%$
Ciclismo	50	$450+50=500$	$\frac{50}{500}=0,1$	$0,9+0,1=1$	$0,1 \cdot 100=10\%$	$90+10=100\%$
Total	500		1		100%	

Exercicio 6

A información que figura a continuación obtívose nunha asociación cultural que ten 25 persoas asociadas. Preguntóuselles polo número de libros que len á semana.

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

Fai unha táboa de frecuencias.

Que porcentaxe de persoas da asociación len máis de 2 libros á semana?

Nº libros	Reconto	f_i	F_i	h_i	H_i	Porcentaxe	Porcentaxe acumulada
0	IIIIII III	8	8	$\frac{8}{25}=0,32$	0,32	$0,32 \cdot 100=32\%$	32%
1	IIIIII IIII	9	$8+9=17$	$\frac{9}{25}=0,36$	$0,32+0,36=0,68$	$0,36 \cdot 100=36\%$	$32+36=68\%$
2	IIIIII I	6	$17+6=23$	$\frac{6}{25}=0,24$	$0,68+0,24=0,92$	$0,24 \cdot 100=24\%$	$68+24=92\%$
3		1	$23+1=24$	$\frac{1}{25}=0,04$	$0,92+0,04=0,96$	$0,04 \cdot 100=4\%$	$92+4=96\%$
4		1	$24+1=25$	$\frac{1}{25}=0,04$	$0,96+0,04=1$	$0,04 \cdot 100=4\%$	$96+4=100\%$
Total		25		1		1	

Len máis de 2 libros á semana os que len 3 ou 4 libros, é dicir, o $4\%+4\%=8\%$ de persoas da asociación len máis de 2 libros á semana.

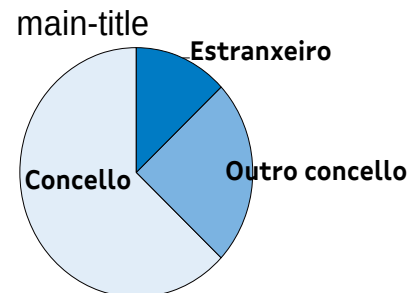
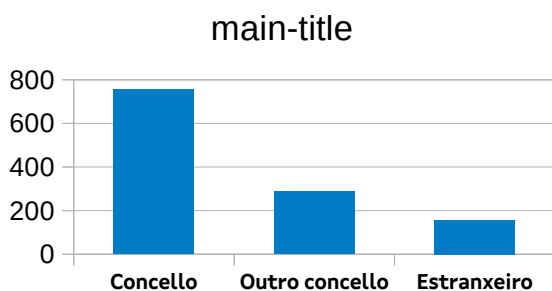
Exercicio 7

Nunha concello con 1200 habitantes, 756 naceron no propio concello, 288 noutro concello galego, e 156 no estranxeiro.

a) Que tipo de gráfico estatístico sería axeitado para representar estes datos?

Ao tratarse de datos cualitativos os gráficos axeitados serían un diagrama de barras ou un diagrama de sectores.

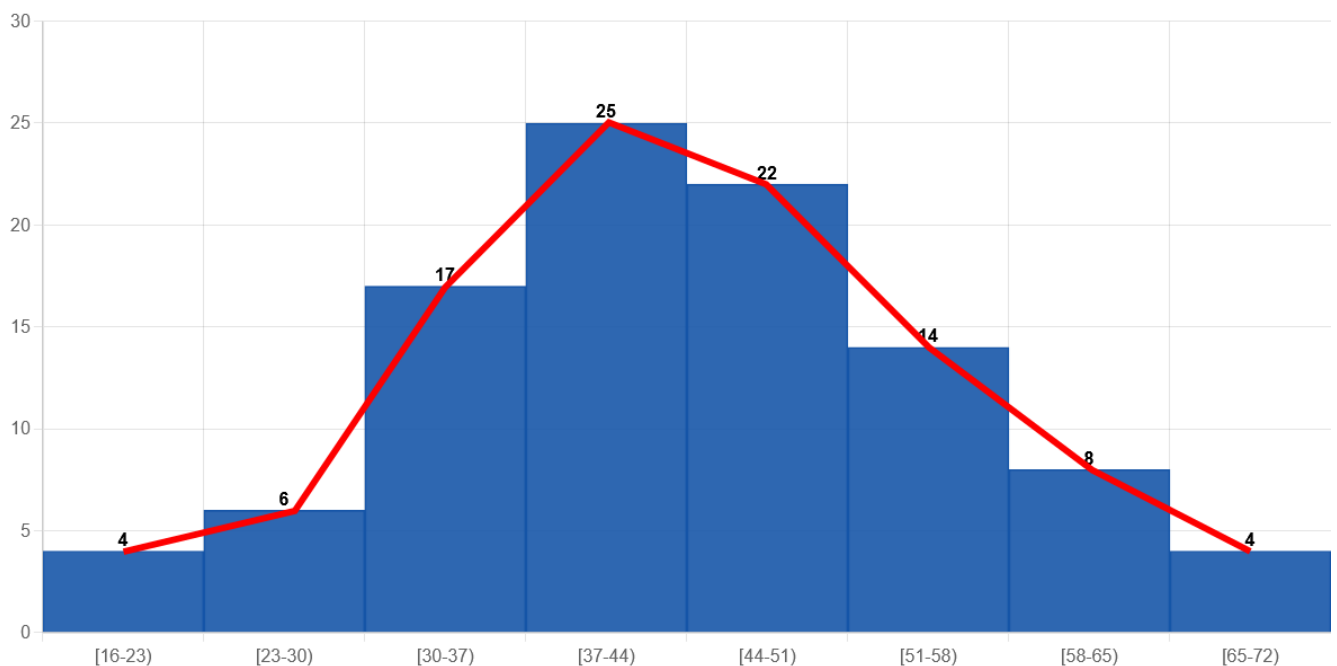
b) Representa os datos mediante un diagrama de barras e un diagrama de sectores.



Exercicio 8

Representa mediante un histograma, co seu correspondente polígono de frecuencias, a porcentaxe de poboación activa por idades nunha cidade.

Idades	[16, 23)	[23, 30)	[30, 37)	[37, 44)	[44, 51)	[51, 58)	[58, 65)	[65, 72)
Porcentaxe	4	6	17	25	22	14	8	4



Exercicio 9

Calcula os parámetros de centralización das idades dos asistentes a dúas festas:

a) 1, 3, 6, 7, 8, 10, 13, 14, 34, 37, 40, 40, 65, 66, 72

- Media: $\bar{x} = \frac{1+3+6+7+8+10+13+14+34+37+40+40+65+66+72}{15} = \frac{416}{15} = 27,73$ anos.
- Moda: $Mo = 40$ anos.
- Mediana: $Me = 14$ anos.

b)

x_i	f_i	F_i	$x_i \cdot f_i$
14	5	5	$14 \cdot 5 = 70$
15	7	12	$15 \cdot 7 = 105$
16	9	21	$16 \cdot 9 = 144$
17	3	24	$17 \cdot 3 = 51$
18	2	26	$18 \cdot 2 = 36$
Total	26		406

- Media: $\bar{x} = \frac{406}{26} = 15,62$ anos
- Moda: $Mo = 16$ anos
- Mediana. Os valores centrais son o 16 e o 16. A súa media é 16, entón $Me = \frac{16+16}{2} = 16$ anos.

Exercicio 10

Calcula os parámetros de centralización e dispersión do número de urxencias médicas atendidas nun centro de saúde dúas semanas:

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

x_i	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	2	2	$1 \cdot 2 = 2$	$1^2 \cdot 2 = 2$
2	2	$2+2=4$	$2 \cdot 2 = 4$	$2^2 \cdot 2 = 8$
3	1	$4+1=5$	$3 \cdot 1 = 3$	$3^2 \cdot 1 = 9$
4	2	$5+2=7$	$4 \cdot 2 = 8$	$4^2 \cdot 2 = 32$
5	3	$7+3=10$	$5 \cdot 3 = 15$	$5^2 \cdot 3 = 75$
7	2	$10+2=12$	$7 \cdot 2 = 14$	$7^2 \cdot 2 = 98$
8	1	$12+1=13$	$8 \cdot 1 = 8$	$8^2 \cdot 1 = 64$
10	1	$13+1=14$	$10 \cdot 3 = 10$	$10^2 \cdot 1 = 100$
Total	14		64	388



- Media: $\bar{x} = \frac{64}{14} = 4,57$
- Moda: $Mo = 5$
- Mediana: Os valores centrais son 4 e 5, entón $Me = \frac{4+5}{2} = 4,5$
- Varianza: $\sigma^2 = \frac{388}{14} - 4,57^2 = \frac{388}{14} - 20,88 = 6,83$
- Desviación típica: $\sigma = \sqrt{6,83} = 2,61$

Exercicio 11

A media das cualificacións dos opositores do tribunal 1 de Matemáticas é $\bar{x} = 7,6$, cunha desviación típica $\sigma_{T_1} = 1,2$. No tribunal 2, os resultados son é $\bar{x} = 6,2$ e $\sigma_{T_2} = 1,2$. Analiza os resultados.

$$CV_{T_1} = \frac{\sigma}{\bar{x}} = \frac{1,2}{7,6} = 0,1578$$

$$CV_{T_2} = \frac{\sigma}{\bar{x}} = \frac{1,2}{6,2} = 0,1935$$

Como $CV_{T_2} > CV_{T_1}$ as cualificacións do tribunal 2 son máis dispersas.