

# Estatística descritiva

---

# O que é a Estatística?

É um conjunto de técnicas apropriadas para recolher, classificar, apresentar e interpretar dados recolhidos numa experiência, previamente planificada, para auxílio na tomada de decisões.

---

# Conceitos básicos

**População:** conjunto de todos os objectos com uma ou mais características em comum, que se pretendem estudar.

- ❑ Cada elemento da população é uma **unidade estatística**
- ❑ O número de elementos da população designa-se por **dimensão da população** e representa-se por **N**
- ❑ A dimensão da população pode ser finita ou infinita

---

# População

Exemplos:

- Alunos de uma escola
- Potenciais eleitores para umas eleições

---

# População

**Parâmetros:** Características numéricas que descrevem a população. Estas características são, em geral, desconhecidas.

**Exemplos:**

- Alunos de uma escola
    - Altura média dos alunos
    - Classificações médias obtidas a Estatística...
  - Potenciais eleitores para umas eleições
    - Proporção de eleitores decididos a votar
    - Idade média dos eleitores...
-

---

# População

**Nem sempre é possível estudar exaustivamente todos os elementos de uma população! Porquê?**

- Pode a população ter dimensão infinita (ex. a população das temperaturas em todos os pontos da cidade)
- Pode o estudo da população levar à destruição da mesma (ex. a população dos fósforos numa caixa)
- Pode o estudo da população ser muito dispendioso em tempo ou dinheiro (ex. sondagem exaustiva de todos os eleitores)

---

# Conceitos básicos

**Amostra:** Subconjunto finito da população.

O número de elementos que fazem parte da amostra designa-se por **dimensão da amostra** e representa-se por  **$n$**

---

# Amostra

## Exemplos:

- Alunos de uma escola
  - Amostra de 100 alunos dessa escola
- Potenciais eleitores para umas eleições
  - Amostra de 1000 potenciais eleitores para essas eleições



---

# Amostra

**Estatística:** Característica numérica que descreve a amostra. Calcula-se o valor da estatística a partir dos valores observados na amostra. Utiliza-se a estatística para estimar um parâmetro desconhecido.

---

# Amostra - Estatística

## Exemplos:

- Alunos de uma escola
  - Altura média dos 100 alunos da amostra
- Potenciais eleitores para as eleições
  - Proporção de eleitores que estão decididos a votar, dos 1000 eleitores da amostra

---

# Amostra

## Será importante a fase da recolha da amostra?

Sim, pois a amostra deve ser representativa quanto possível da população de onde foi extraída, para que as conclusões possam estender-se a toda a população.

Uma amostra não representativa diz-se **enviesada**.

---

# Amostra

## Exemplo

- Utilizar uma amostra constituída por 10 benfiquistas, para prever o vencedor de um jogo do Benfica-Sporting.
- Utilizar uma amostra constituída por leitores de uma revista especializada, para tirar conclusões sobre a população em geral.

---

# Amostra

**Exercício:** Diga porque é que as seguintes situações representam más amostras:

- Para saber qual o candidato mais votado, para a Câmara de determinada cidade, auscultou-se a opinião dos clientes de determinado supermercado.
- Para conhecer a situação financeira das empresas têxteis portuguesas, verificou-se a situação das empresas que tiveram maior volume de exportações, no último ano.

---

# Censo vs Sondagem

**Censo ou Recenseamento** - estudo estatístico em que se observa toda a população.

**Exemplos:** recenseamento da população (2021); censo para o serviço militar.

**Sondagem** - estudo estatístico em que se estuda uma amostra da população

**Exemplos:** preferência musical dos alunos de uma Escola; índice de audiência de um programa da TV.

---

# **Estatística descritiva vs inferência estatística**

- **Estatística Descritiva:** conjunto de métodos estatísticos que visam sintetizar e representar de forma compreensível a informação contida nos dados.
- **Inferência Estatística:** conjunto de métodos estatísticos que visam caracterizar (ou inferir sobre) uma população a partir da informação obtida da amostra.

---

# Variável estatística

Característica da população que se pretende estudar – pode tomar vários valores possíveis.

Cada um desses valores é denominado  
categoria ou classe



---

# Variável estatística

Uma variável, enquanto representar apenas a característica e não estiver concretizada em nenhum elemento, representa-se habitualmente por uma letra maiúscula. Quando se pretende representar o valor da variável para um indivíduo em particular utiliza-se a respectiva letra minúscula.

Exemplo:

$X$  - representa a hemoglobina no sangue;

$x = 14.2$  - representa a hemoglobina de um certo indivíduo.

---

# Variável estatística - Tipo

**Qualitativas** – são aquelas que estão relacionadas com uma qualidade e apresentam-se com várias modalidades

**Exemplo:** cor dos olhos, estado civil

**Quantitativas** – são aquelas a que é possível atribuir um valor numérico e apresentam-se com diferentes intensidades ou valores

**Exemplo:** o número diário de nascimentos no hospital da Guarda

---

# Variável estatística - qualitativa

**Escala Nominal:** apresentam-se em diferentes categorias, não ordenáveis.

**Exemplos:**

- ❑ Estado civil dos empregados de uma empresa;
- ❑ Religião;
- ❑ Cor de cabelos;
- ❑ Sexo dos indivíduos de uma população;  
(**característica dicotómica ou binária**)
- ❑ Numa sondagem de opinião, a resposta à pergunta  
“É a favor da despenalização da eutanásia?”  
(**característica dicotómica ou binária**)

---

# Variável estatística - qualitativa

**Escala Ordinal:** as diversas categorias possuem uma ordem intrínseca.

Exemplos:

- ❑ O sistema de graduação militar: Soldado, Cabo, Sargento
- ❑ Num inquérito de opinião pede-se às pessoas que classifiquem um determinado produto como sendo: muito fraco, fraco, razoável, bom ou muito bom.
- ❑ Classificação dos clientes de um banco, segundo o volume de capital que movimentam mensalmente: pouco importantes, importantes ou muito importantes.
- ❑ Classificação dos alunos de uma escola segundo a sua altura: baixos (menos de 155 cm), médios (entre 155 e 170 cm) ou altos (mais de 170 cm).

---

# Variável estatística - quantitativa

**Discretas** - assumem valores num conjunto finito ou infinito numerável

**Exemplo:** o número diário de nascimentos no hospital da Guarda

**Contínuas** - assumem qualquer valor num intervalo de números reais

**Exemplo:** altura de um aluno da ESTG

---

# Variáveis discretas *vs* contínuas

## Fronteira pouco clara:

Dados discretos podem ter uma gama de valores tão dispersa que na prática funcionam como se fossem contínuos.

## Exemplos:

- ❑ O nº de células contidas em 1 ml de sangue.
- ❑ O nº de peixes que entra diariamente no rio Tejo, trazido pelas marés.

---

# Variáveis discretas *vs* contínuas

Dados contínuos são registados com precisão finita (seja grande ou pequena) e na prática são discretos.

## Exemplos:

- ❑ A idade de uma pessoa, em anos.
- ❑ O diâmetro de uma semente de papoila, em décimas de mm.

---

## Exercício:

Classifique como qualitativa (nominal ou ordinal) ou quantitativa (discreta ou contínua) as variáveis:

1. Número de filhos de um casal;
2. Grupo sanguíneo;
3. Temperatura mínima diária;
4. Classificações obtidas no 3º ciclo do ensino básico;
5. Grupo etário (crianças, jovens, adultos, idosos);
6. Comprimento das orelhas de um cão;
7. Categorias taxonómicas de plantas e animais;
8. Pulsação.



---

# Exercício:

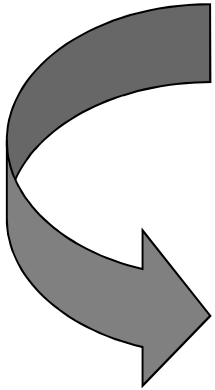
## Soluções:

1. Quantitativa discreta;
2. Qualitativa ordinal;
3. Quantitativa contínua;
4. Quantitativa discreta;
5. Qualitativa ordinal;
6. Quantitativa contínua;
7. Qualitativa ordinal;
8. Quantitativa discreta.

---

# Organização dos dados

O que fazer com os dados recolhidos?



Primeira etapa: resumo dos dados

---

# Organização dos dados

- Tabela de frequências (absolutas e relativas)
- Representações gráficas
  - Dados qualitativos
    - Diagrama de barras
    - Diagrama circular
  - Dados quantitativos
    - Discretos
      - Diagrama de barras
    - Contínuos
      - Histograma
      - Diagrama Caixa de Bigodes

# Organização dos dados

Tabela é um quadro que resume um conjunto de observações.

## **Frequência absoluta ( $f_i$ ):**

Número de elementos de cada uma das categorias ou classes  $i$

## **Frequência relativa ( $f_{r_i}$ ):**

Proporção de elementos de cada uma das categorias ou classes  $i$ . É dada pelo quociente entre a frequência absoluta e a dimensão da amostra, ou seja:

$$f_{r_i} = \frac{f_i}{n}$$

---

# Organização dos dados

Para confirmar que as frequências estão bem calculadas, basta verificar que

**A soma das frequências absolutas é igual à  
dimensão da amostra**

e

**A soma das frequências relativas é igual a 1**

# Organização dos dados

<b>Categoria ou Classe</b>	$f_i$	$f_{r_i}$
$C_1$	$n_1$	$\frac{n_1}{n}$
$C_2$	$n_2$	$\frac{n_2}{n}$
$\vdots$	$\vdots$	$\vdots$
$C_p$	$n_p$	$\frac{n_p}{n}$
<b>Total</b>	$n$	1

---

# Organização dos dados

## Exemplo:

Perguntou-se a cada um dos 25 funcionários de uma loja qual o tipo de música ambiente que preferiam ouvir durante o expediente. Os resultados foram os seguintes:

R C J P C P J J P R P O R  
J R R P R O P O C N P P

onde C- Clássica, P- Pop, R- Rock, J- Jazz,  
O- Outro tipo de música e N - Nenhum tipo

---

# Organização dos dados

Exemplo:

Preferência Musical	$f_i$	$f_{r_i}$
Clássica (C)	3	0,12
Pop (P)	8	0,32
Rock (R)	6	0,24
Jazz (J)	4	0,16
Outro (O)	3	0,12
Nenhum (N)	1	0,04
<b>Total</b>	<b>25</b>	<b>1</b>



---

# Organização dos dados

Podemos também apresentar na tabela de frequências a **frequência absoluta acumulada** e a **frequência relativa acumulada**, que se obtêm adicionando as frequências absolutas e relativas, respectivamente.

# Organização dos dados

Categoria ou classe	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
$C_1$	$n_1$	$f_{r_1} = \frac{n_1}{n}$	$n_1$	$f_{r_1}$
$C_2$	$n_2$	$f_{r_2} = \frac{n_2}{n}$	$n_1 + n_2$	$f_{r_1} + f_{r_2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$C_p$	$n_p$	$f_{r_p} = \frac{n_p}{n}$	$n_1 + n_2 + \dots + n_p = n$	$f_{r_1} + f_{r_2} + \dots + f_{r_p} = 1$
Total	$n$	1	-	-

---

# Organização dos dados

## Exemplo:

Numa turma do 10º ano, os alunos registaram o nº de irmãos, tendo-se obtido a seguinte amostra:

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

Elabore a tabela de frequências.

# Organização dos dados

Exemplo:

Nº irmãos	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
0	4	0,20	4	0,20
1	8	0,40	4+8=12	0,2+0,4=0,6
2	4	0,20	16	0,80
3	3	0,15	19	0,95
4	1	0,05	20	1
<b>Total</b>	20	1	-	-

# Organização dos dados

Exemplo:

Quantos alunos  
têm:

a) Menos de 3  
irmãos

16

Nº irmãos	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
0	4	0,20	4	0,20
1	8	0,40	12	0,60
2	4	0,20	16	0,80
3	3	0,15	19	0,95
4	1	0,05	20	1
<b>Total</b>	20	1	-	-

# Organização dos dados

Exemplo:

Quantos alunos têm:

a) Menos de 3 irmãos

b) Pelo menos 2 irmãos

20-12

ou  $4+3+1$

Nº irmãos	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
0	4	0,20	4	0,20
1	8	0,40	12	0,60
2	4	0,20	16	0,80
3	3	0,15	19	0,95
4	1	0,05	20	1
Total	20	1	-	-

# Organização dos dados

Exemplo:

Qual a  
percentagem de  
alunos com:

a) 2 irmãos

$$0,20 \times 100\% = 20\%$$

Nº irmãos	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
0	4	0,20	4	0,20
1	8	0,40	12	0,60
2	4	0,20	16	0,80
3	3	0,15	19	0,95
4	1	0,05	20	1
<b>Total</b>	20	1	-	-

# Organização dos dados

Exemplo:

Qual a  
percentagem de  
alunos com:

a) 2 irmãos

b) Menos de 2  
irmãos

$$0,60 \times 100\% = 60\%$$

Nº irmãos	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
0	4	0,20	4	0,20
1	8	0,40	12	0,60
2	4	0,20	16	0,80
3	3	0,15	19	0,95
4	1	0,05	20	1
<b>Total</b>	20	1	-	-



# Organização dos dados

Exemplo:

Qual a  
percentagem de  
alunos com:

a) 2 irmãos

b) Menos de 2  
irmãos

c) Pelo menos 3  
irmãos

$$(1 - 0,8) \times 100\% = 20\%$$

Nº irmãos	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
0	4	0,20	4	0,20
1	8	0,40	12	0,60
2	4	0,20	16	0,80
3	3	0,15	19	0,95
4	1	0,05	20	1
<b>Total</b>	20	1	-	-

---

# Organização dos dados em classes

Para o caso de os dados serem contínuos ou discretos com valores muito distintos é usual proceder-se ao agrupamento dos dados em **intervalos de classes**.

---

# Organização dos dados em classes

## Classe:

É um intervalo da forma  $[a,b[$  ou  $]a,b]$  que representa um conjunto de valores que a variável estatística pode tomar

## Amplitude da classe:

É a distância entre o limite superior e o limite inferior, ou seja,

$$\text{Amplitude} = b - a$$

---

# Organização dos dados em classes

## Valor característico:

É o ponto médio da classe e calcula-se por

$$\frac{a + b}{2}$$

---

# Organização dos dados em classes

As questões que se colocam são:

**Qual o número de classes a considerar?**

**Qual a amplitude de classe a considerar?**

---

# Organização dos dados em classes

Qual o número de classes ( $k$ ) a considerar?

- Regra de Sturges:  $k \cong 1 + 3,3 \log_{10} n$

-  $k \cong \sqrt{n}$

---

# Organização dos dados em classes

Qual a amplitude de classe a considerar?

$$\textit{amplitude de classe} = \frac{\textit{maior observação} - \textit{menor observação}}{\textit{número de classes}}$$

---

# Organização dos dados em classes

Etapas para a construção da tabela de frequências:

## 1) Definição das classes

- a) Determinar o número de classes  $k$
- b) Determinar a amplitude da amostra (máximo - mínimo)
- c) Dividir esta amplitude pelo número de classes,  $k$
- d) Tomar para amplitude de classe,  $h$ , um valor aproximado por excesso do valor obtido em c)
- e) Construir as classes de modo que tenham todas a mesma amplitude e cuja união contenha todos os elementos da amostra

## 2) Contagem do número de elementos de cada classe.

---



---

# Organização dos dados em classes

Os intervalos das classes devem ser:

- ❑ **mutuamente exclusivos** – um indivíduo não pode ser classificado em duas classes em simultâneo
- ❑ **exaustivos** – nenhum indivíduo pode ficar sem classificação

---

# Organização dos dados em classes

**Exemplo:** O tempo necessário para executar determinada cirurgia foi cronometrado (em minutos), tendo sido feitas 60 observações:

35 27 29 38 41 30 43 39 23 26 28 33 23 31 31  
29 31 35 33 35 24 35 25 38 41 41 35 36 34 35  
31 45 28 36 36 46 47 26 43 39 37 38 34 39 28  
30 22 20 47 34 47 35 34 38 33 37 30 35 26 39

Obtenha a tabela de frequências para os dados.

---

# Organização dos dados em classes

$$n^{\circ} \text{ de classes} = 1 + 3,3 \log_{10} 60 \cong 7$$

maior observação = 47 e menor observação = 20

$$\text{amplitude de classe} = \frac{47-20}{7} \cong 4$$

limite inferior da 1ª classe = 20

# Organização dos dados em classes

<b>Tempo</b>	$f_i$	$f_{r_i}$	$F_i$	$F_{r_i}$
[20,24[	4	0,07	4	0,07
[24,28[	6	0,10	10	0,17
[28,32[	12	0,20	22	0,37
[32,36[	15	0,25	37	0,62
[36,40[	13	0,22	50	0,84
[40,44[	5	0,08	55	0,92
[44,48[	5	0,08	60	1
<b>Total</b>	60	1		

---

# Representação gráfica dos dados

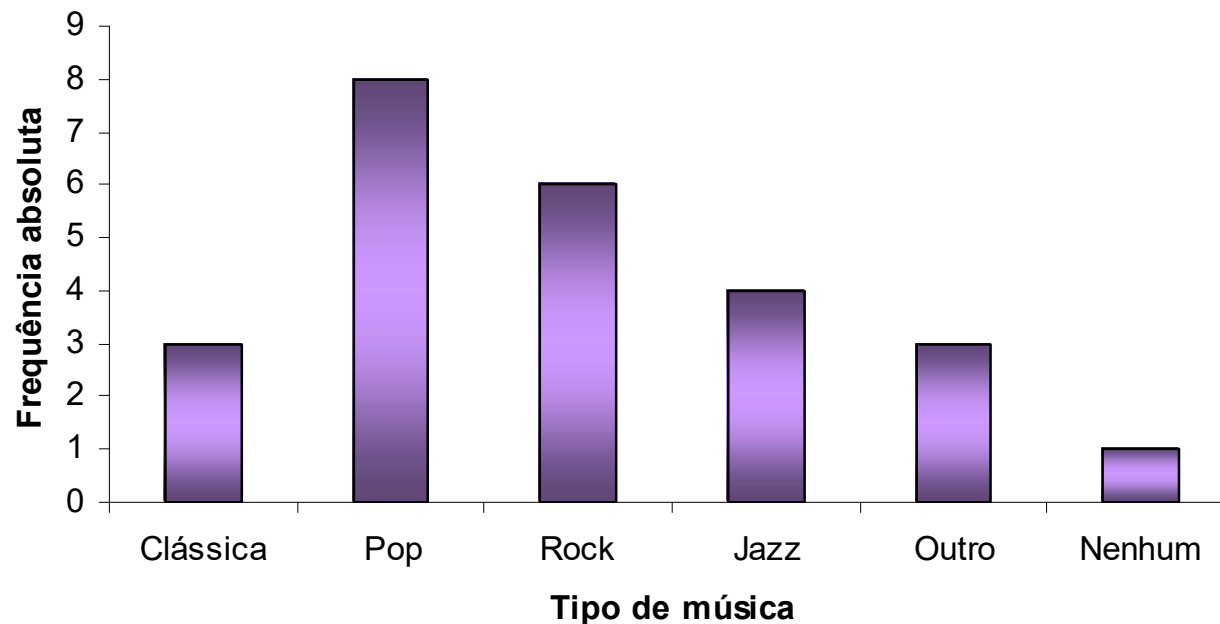
## Gráfico de barras

A metodologia a seguir na construção do gráfico de barras é:

- 1) Ordenar a amostra e considerar para classes/categorias os diferentes valores. Marcar essas classes/categorias no eixo dos XX (horizontal) do sistema de eixos coordenados.
- 2) Nos pontos onde se consideram as classes/categorias, marcar barras de altura igual à frequência absoluta ou relativa da classe/categoria.

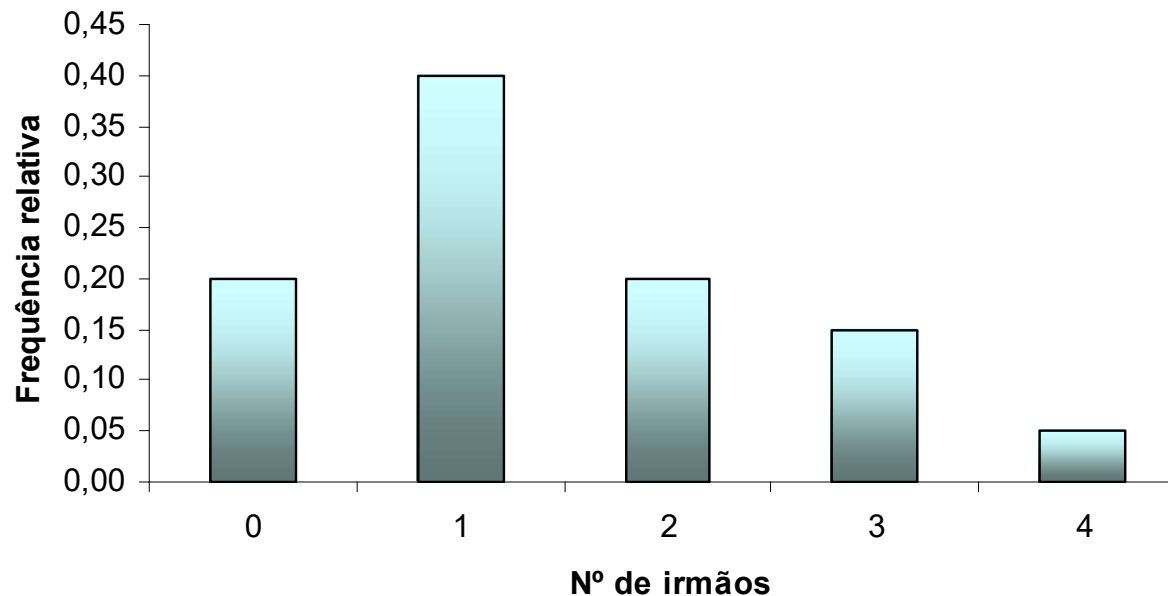
# Representação gráfica dos dados

**Exemplo:** O gráfico de barras que representa a distribuição do tipo de música ambiente preferida pelos funcionários da loja tem o seguinte aspecto:



# Representação gráfica dos dados

**Exemplo:** O gráfico de barras que representa a distribuição do número de irmãos dos alunos da turma considerada tem o seguinte aspecto:



---

# Representação gráfica dos dados

## Diagrama circular

É constituído por um círculo, em que se apresentam vários sectores circulares, tantos quantas as classes ou categorias consideradas na tabela de frequências.

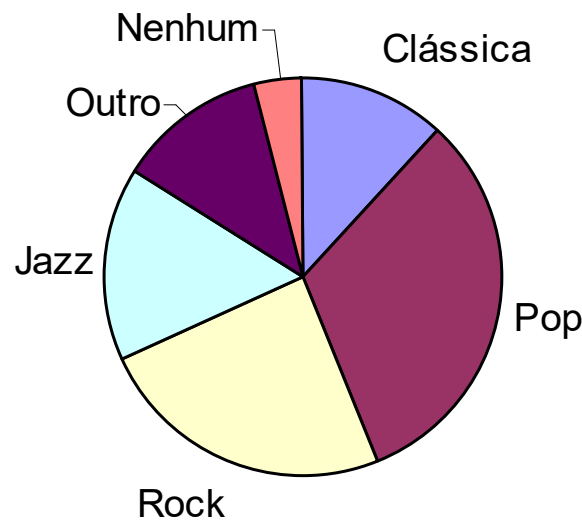
Os ângulos dos sectores são proporcionais às frequências das classes ou categorias.



# Representação gráfica dos dados

**Exemplo:** O gráfico circular que representa a distribuição do tipo de música ambiente dos funcionários da loja tem o seguinte aspecto:

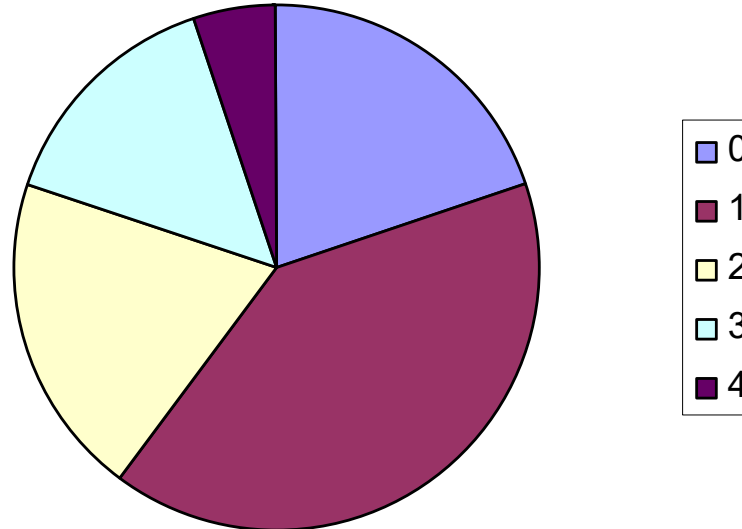
**Frequência Relativa**



# Representação gráfica dos dados

**Exemplo:** O gráfico circular que representa a distribuição do número de irmãos dos alunos da turma considerada tem o seguinte aspecto:

Frequência Absoluta



---

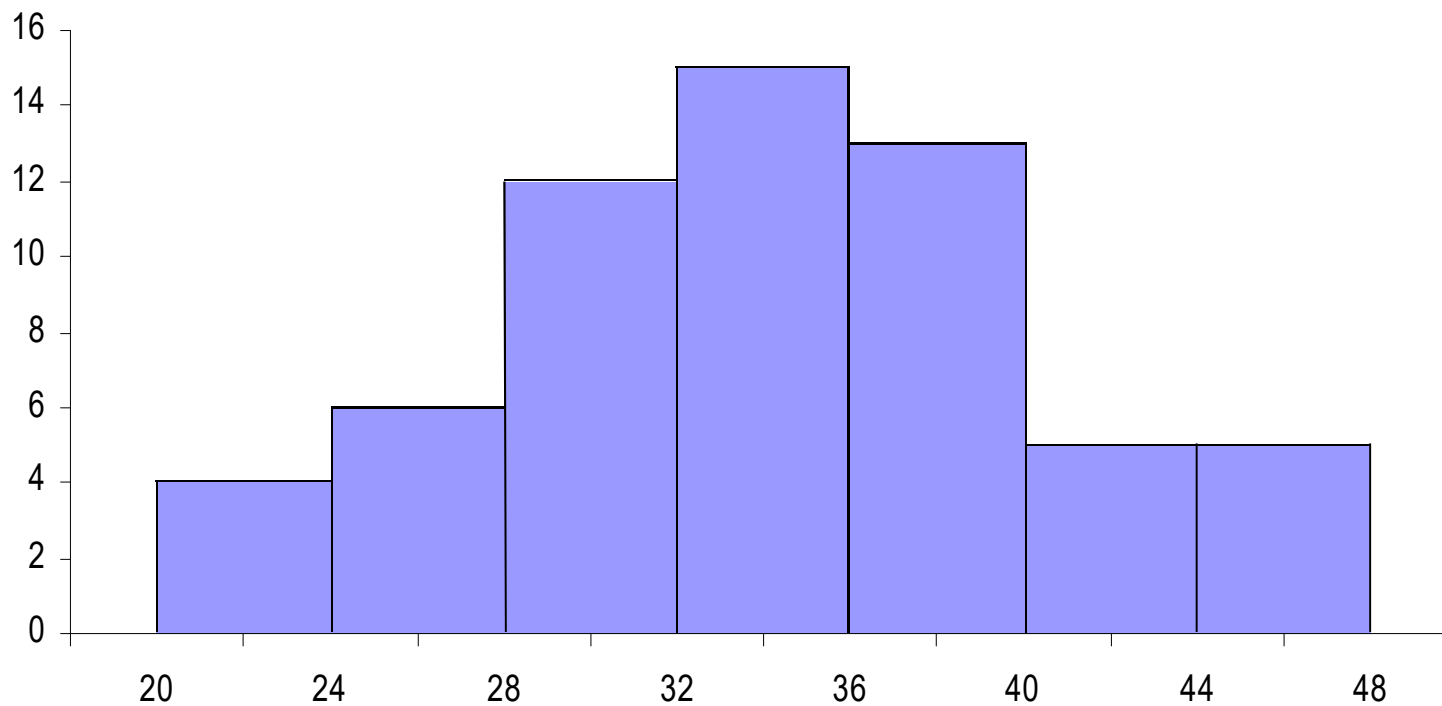
# Representação gráfica dos dados

## Histograma

É formado por rectângulos adjacentes, tendo por base um intervalo de classe e por altura a frequência absoluta ou relativa

# Representação gráfica dos dados

**Exemplo:** O histograma que representa a distribuição da frequência absoluta, do tempo de duração de uma dada cirurgia tem o seguinte aspecto:



---

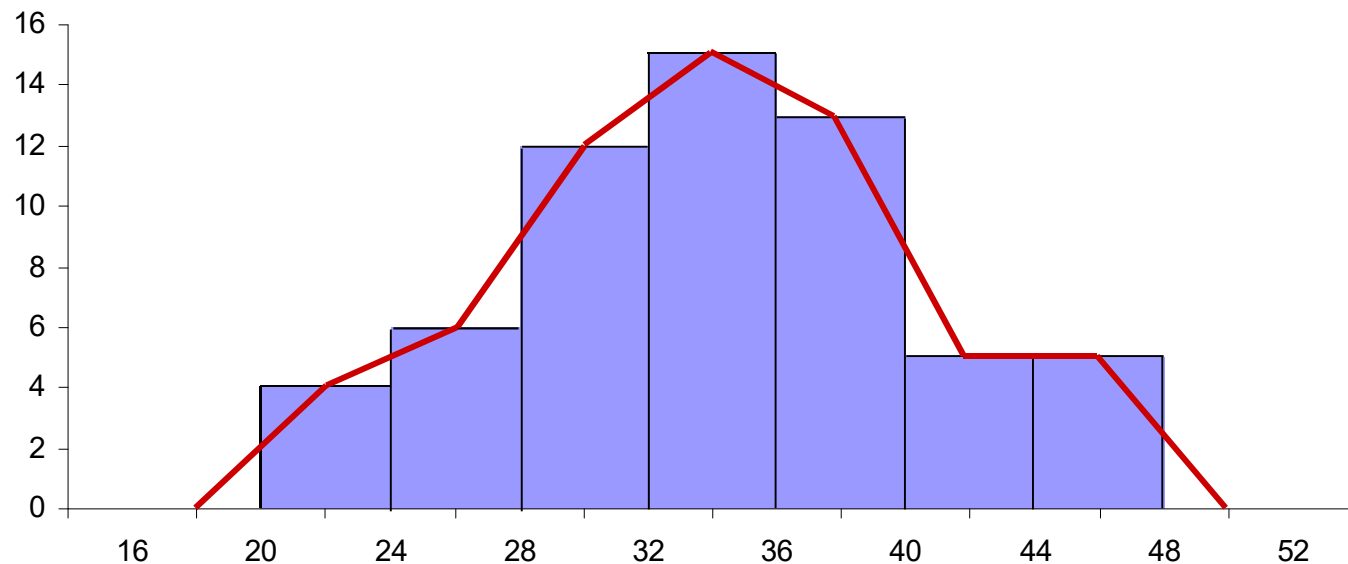
# Representação gráfica dos dados

## Polígono de frequências

É uma representação gráfica dos dados em que se une, através de segmentos de recta, todos os pontos cuja abcissa é o valor característico de uma classe e a ordenada é a frequência dessa classe. O polígono pressupõe, em cada extremo, a existência de uma classe adicional de frequência nula.

# Representação gráfica dos dados

**Exemplo:** Polígono de frequências do tempo de duração de uma cirurgia.



---

# Descrição dos dados – medidas

## O que são Medidas Descritivas?

São medidas que podem ser obtidas sobre conjuntos de dados numéricos, de forma a fornecer informações sobre esses dados.

---

# Descrição dos dados – medidas

## Resumindo numericamente

Para resumir numericamente dados quantitativos, o objetivo é escolher medidas apropriadas de:

- ❑ **Localização** – qual o tamanho dos números envolvidos
- ❑ **Dispersão** – quanta variação existe

para os tipos de dados

---



---

# Descrição dos dados – medidas

- Medidas de localização
  - Média
  - Mediana
  - Moda
  - Quartis e percentis
- Medidas de dispersão ou variabilidade
  - Amplitude
  - Desvio padrão e variância
  - Amplitude inter-quartil

# Descrição dos dados – média

## Definição:

Dado um conjunto de  $n$  valores observados  $x_1, x_2, \dots, x_n$ , a **média amostral** é dada por

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

# Descrição dos dados – média

Se os dados se encontram agrupados, a média amostral é dada por

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k f_i x'_i$$

onde

$k$  – número de classes

$f_i$  – frequência absoluta da classe  $i$

$x'_i$  – valor característico da classe  $i$

# Descrição dos dados – média

A média amostral será sempre uma medida representativa dos dados?

## Exemplo:

Ao determinar a média dos seguintes dados

12,4 13,5 13,6 11,2 113,5

obteve-se  $\bar{x} = 32,84$ .

## Obs.

- ❑ O exemplo revela a sensibilidade da média a observações extremas, isto é, observações muito grandes ou muito pequenas;
- ❑ A popularidade da média é consequência de ser uma medida de cálculo fácil.

# Descrição dos dados – média

## Exercício (dados não classificados):

Considere o exemplo do número de irmãos dos alunos da turma do 10º ano. Determine a média.

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

$$\bar{x} = \frac{1 + 2 + 2 + 1 + 3 + 0 + 0 + 1 + 1 + 2 + 1 + 1 + 1 + 0 + 0 + 3 + 4 + 3 + 1 + 2}{20}$$

$$\bar{x} = \frac{29}{20} = 1,45$$

# Descrição dos dados – média

## Exercício (dados classificados):

Considere o exemplo do número de irmãos dos alunos da turma do 10º ano. Determine a média.

nº de irmãos	0	1	2	3	4
$f_i$	4	8	4	3	1

$$\bar{x} = \frac{0 \times 4 + 1 \times 8 + 2 \times 4 + 3 \times 3 + 4 \times 1}{20}$$

$$\bar{x} = \frac{29}{20} = 1,45$$

# Descrição dos dados – média

## Exercício (dados não classificados):

Considere o exemplo dos tempos de duração de certa cirurgia. Calcule a média.

35 27 29 38 41 30 43 39 23 26 28 33 23 31 31  
29 31 35 33 35 24 35 25 38 41 41 35 36 34 35  
31 45 28 36 36 46 47 26 43 39 37 38 34 39 28  
30 22 20 47 34 47 35 34 38 33 37 30 35 26 39

$$\begin{aligned}\bar{x} &= \frac{35 + 27 + 29 + 38 + \dots + 35 + 26 + 39}{60} = \\ &= \frac{2044}{60} = 34,06666 \dots\end{aligned}$$

# Descrição dos dados – média

## Exercício (dados classificados):

Considere o exemplo dos tempos de duração de certa cirurgia. Calcule a média.

Tempo	$f_i$
[20,24[	4
[24,28[	6
[28,32[	12
[32,36[	15
[36,40[	13
[40,44[	5
[44,48[	5

$$\begin{aligned}\bar{x} &= \frac{22 \times 4 + 26 \times 6 + 30 \times 12 + 34 \times 15 + 38 \times 13 + 42 \times 5 + 46 \times 5}{60} \\ &= \frac{2048}{60} = 34,13333 \dots\end{aligned}$$



---

# Descrição dos dados – média

## Nota:

Sempre que possível deve-se calcular a média a partir dos dados originais.

No caso do exemplo anterior a média calculada a partir dos dados originais é 34,07 e a partir dos dados classificados é 34,13. No caso dos dados classificados há perda de informação.

# Descrição dos dados – mediana

## Definição:

Dado um conjunto de  $n$  valores observados  $x_1, x_2, \dots, x_n$ , ordenados, a **mediana amostral**, representada por **Me**, é o valor, pertencente ou não à amostra, que a divide ao meio.

Por outras palavras 50% dos elementos da amostra são menores ou iguais à mediana amostral e os restantes 50% são maiores ou iguais à mediana amostral.

# Descrição dos dados – mediana

Dado um conjunto de  $n$  valores observados  $x_1, x_2, \dots, x_n$ , ordenados, então

$$Me = \begin{cases} \frac{x_{\frac{n+1}{2}}}{2} & , se\ n\ \acute{e}\ \acute{im}par \\ \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} & , se\ n\ \acute{e}\ par \end{cases}$$

## Exercício (dados não classificados):

Obter as medianas dos seguintes conjuntos de observações:

$$10\ 10\ 12\ 7\ 8\ 5\ 4 \Rightarrow Me = 8$$

$$10\ 10\ 12\ 7\ 7\ 8\ 5\ 4 \Rightarrow Me = 7,5$$

---

# Descrição dos dados – mediana

## Exemplo:

A mediana dos dados

12,4   13,5   13,6   11,2   113,5

é  $Me = 13,5$ , o que nos dá uma melhor ideia sobre o conjunto dos dados.

## Obs.

Como vemos a mediana amostral é mais resistente a observações díspares.

---

## Descrição dos dados – mediana

No caso de termos dados agrupados em intervalos de classes, a identidade das observações perde-se; nesse caso, a mediana nunca pode ser calculada exatamente. Uma forma de o fazer é através da expressão:

$$Me = l_{cm} + \frac{0,5 - F_{r_{cm-1}}}{f_{r_{cm}}} \times (L_{cm} - l_{cm})$$

sendo  $CM_e = [l_{cm}, L_{cm}[$  a classe da mediana, ou seja, a primeira classe cuja frequência relativa acumulada é superior a 0,5

---

# Descrição dos dados – mediana

## Exemplo (dados classificados):

Considere o exemplo da duração de determinada cirurgia. Determine a mediana.

$$Me = 32 + \frac{0,5 - 0,37}{0,25} \times (36 - 32) = 34,08$$

---

# Descrição dos dados – moda

## Definição:

A **moda (Mo)** de um conjunto de  $n$  observações  $x_1, x_2, \dots, x_n$ , é a observação que ocorre com mais frequência na amostra, caso exista.

## Obs.

Quando os dados estão agrupados em intervalos de classes, a **classe modal** é a classe com maior frequência absoluta (relativa), podendo existir mais do que uma classe com esta propriedade.

---

# Descrição dos dados – moda

Uma distribuição pode classificar-se, quanto à moda por:

- ❑ Amodal – não tem moda
- ❑ Unimodal – tem uma moda
- ❑ Bimodal – tem duas modas
- ❑ Multimodal – tem mais de duas modas



---

# Descrição dos dados – moda

## Exercícios:

Determine a moda

**Dados não classificados**

Exemplo: N° de irmãos

$Mo = 1$

**Dados classificados**

Exemplo: Tempo de cirurgia

Classe modal = [ 32, 36 [

---

# Descrição dos dados – quartis

## Definição:

Dividem um conjunto de dados, ordenados por ordem crescente, em quatro partes iguais. Definem-se 3 quartis.

- **1º quartil** (ou quartil inferior) é o percentil correspondente à percentagem de 25%, o que significa que 25% dos elementos da amostra são menores ou iguais a ele, e os restantes 75 % são maiores ou iguais.

---

# Descrição dos dados – quartis

- **2º quartil** (ou Mediana) é o percentil correspondente à percentagem de 50%, o que significa que 50% dos elementos da amostra são menores ou iguais a ele, e os restantes 50% são maiores ou iguais
- **3º quartil** (ou quartil superior) é o percentil correspondente à percentagem de 75%, o que significa que 75% dos elementos da amostra são menores ou iguais a ele, e os restantes 25% são maiores ou iguais

# Descrição dos dados – quartis

Localização dos quartis (dados não classificados)

	$n$ é par	$n$ é ímpar
$Q_1$	$\frac{n + 2}{4}$	$\frac{n + 1}{4}$
$Q_3$	$\frac{3n + 2}{4}$	$3 \times \frac{n + 1}{4}$

---

# Descrição dos dados – quartis

## Exercício (dados não classificados):

Tendo-se decidido registrar os pesos dos alunos de uma determinada turma do 2º ano, obtiveram-se os seguintes valores (em kg):

52 56 62 54 52 70 51 60 61 56 55 56 54 57 67 61 49

Ordenar os dados:

49 51 52 52 54 54 55 56 56 56 57 60 61 61 62 67 70

$Q_1 = 53$  ;  $Q_2 = 56$  ;  $Q_3 = 61$

# Descrição dos dados – quartis

**No caso dos dados classificados tem-se:**

$$Q_1 = l_{cq1} + \frac{0,25 - F_{rcq1-1}}{f_{rcq1}} \times (L_{cq1} - l_{cq1})$$

sendo  $CQ_1 = [l_{cq1}, L_{cq1}[$  a classe do quartil 1, ou seja, a primeira classe cuja frequência relativa acumulada é superior a 0,25 e

$$Q_3 = l_{cq3} + \frac{0,75 - F_{rcq3-1}}{f_{rcq}} \times (L_{cq3} - l_{cq3})$$

sendo  $CQ_3 = [l_{cq3}, L_{cq3}[$  a classe do quartil 3, ou seja, a primeira classe cuja frequência relativa acumulada é superior a 0,75

# Descrição dos dados – quartis

## Exercício (dados classificados):

Considere o exemplo dos tempos de duração de certa cirurgia. Calcule os quartis.

$$Q_1 = 28 + \frac{0,25 - 0,17}{0,20} \times (32 - 28) = 29,6$$

$$Q_2 = Me = 34,08$$

$$Q_3 = 36 + \frac{0,75 - 0,62}{0,22} \times (40 - 36) = 38,36$$