

Ficha 7 — Dados univariados: representação e leitura

10.º Ano Matemática A 4 páginas 12 exercícios

Objetivo: Construir e ler tabelas de frequências (absolutas, relativas e acumuladas), diagramas de barras, caule-e-folhas, diagramas de extremos-e-quartis e histogramas (com classes).

Revisão rápida — da amostra ao gráfico

- Frequência absoluta** f_i = nº de observações na classe · **relativa** f_i/N (fração ou %) · **acumulada** = soma progressiva (só chega a 100% / N no fim).
- Barras:** altura = frequência (→ relativa compara «regularidade» entre amostras diferentes). **Histograma:** as classes estão *ordenadas e contíguas* → barras coladas, área proporcional (frequência = altura × largura quando a frequência é por unidade de amplitude).
- Caule-e-folhas:** cada valor em folha, linhas = caule (ordem de grandeza). **Extremos-e-quartis (boxplot):** min–Q1–Q2(media)–Q3–max; Q2 divide em metades; Q1/Q3 dividem cada metade em dois quartis.
- Amplitude total (AT)** = max – min · **AIQ** = Q3 – Q1 — a «caixa» do boxplot.

Horas de estudo (semana) — 40 alunos	f_i	f_i rel.	Acumulada
[0 – 2[	4	10%	4
[2 – 4[	6	15%	10
[4 – 6[	14	35%	24
[6 – 8[	10	25%	34
[8 – 10]	6	15%	40

Exercícios

1. Verifica a tabela: N = 40 e relativas = 100% (foi preenchida). Depois classifica o gráfico ideal: a) notas de 50 valores → melhor: __ b) horas de estudo (5 classes) → __ c) 40 pesos ordenados → __ d) comparar % de aprovados em 3 escolas → __

2. Frequências acumuladas: com a tabela: a) alunos com < 6 h (acumulada até [4–6[): __ b) com ≥ 6 h: __ c) explica porque a acumulada é sempre crescente e só chega ao total na última classe.

3. Caule-e-folhas: caule 4 | 0 5, caule 5 | 0 0 2 5, caule 6 | 0 5 — a) lê 3 valores: ___ b) nº de observações: ___ c) reconstitui os 7 valores por ordem e calcula a amplitude total: ___

4. Extremos-e-quartis (figura): O diagrama abaixo tem extremos em 6 e 34, caixa de 12 a 28 e mediana a 20:



Coloca os 5 números-resumo: min ___ Q1 ___ Q2 ___ Q3 ___ max ___ → AT ___ AIQ ___. Se 75% dos dados são ≤ 28 , o que se pode dizer dos 25% restantes? ___

5. Classes e histograma: dados 7, 14, 18, 5, 12, 9, 20, 11, 16, 6 — amplitude total ___ com 4 classes → largura de classe ___ → completa: [4–8[$f_i =$ ___ · [8–12[___ · [12–16[___ · [16–20] ___ · soma ___.

6. Barras vs. relativas: escola A: 120/400 aprovados; escola B: 30/80. a) % de cada ___ b) qual «melhor» em %? c) em que situação uma só interessa mesmo a absoluta?

7. Histograma leitura: da tabela da Revisão: a) classe modal __ b) onde cai a mediana ($N=40 \rightarrow$ posição 20,5) __ c) a distribuição tem assimetria? — qual? __

8. Frequência por unidade de amplitude: classe $[4-6[$ com $f_i = 14 \rightarrow$ densidade $= 14/2 =$ __ · classe $[0-2[$ $f_i = 4 \rightarrow$ __ · a barra da $[4-6[$ é __× mais alta — explica porque é que usa-se isto (classes de larguras diferentes, 1 frase).

9. Algoritmo (Python):

```
dados=[7,14,18,5,12,9,20,11,16,6]
classe=lambda x: min(int(x//4),3)
f=[0,0,0,0]
for x in dados: f[classe(x)]+=1
print(f)
```

a) lista impressa __ b) que intervalos? __ c) substitui lambda por passo a passo (o que faz $\text{int}(x//4)$)?

10. Boxplot comparativo: caixa 12–28 (mediana 20) vs. caixa 14–26 (mediana 19): a) qual tem maior AIQ? __ b) qual é mais «homogéneo»? — argumenta com AIQ
c) ambos têm mesma mediana? __

11. Um gráfico enganoso: desenho (ou descreve) um eixo das ordenadas a começar em 35 que faz parecer dobrar os valores de 40 para 80. a) porque é que «80 parece 4×40 »? b) 2 regras de honestidade do eixo zero.

12. Produção: 20 pesos de uma turma (a fornecer pelo professor em folha de cálculo) $\rightarrow$ constrói tabela com 4 classes, histograma e extremos-e-quartis; comenta: classe modal, mediana vs. média, outliers (se 90 kg surgir, porque é outlier?).

Soluções — Ficha 7: Dados univariados

1. $N = 40$ ✓ · relativas = $10+15+35+25+15 = 100\%$ ✓ · a) caule-e-folhas (muitos valores, ver precisão) b) histograma (classes contíguas) c) extremos-e-quartis (resumo 5 números) d) barras *relativas* (tamanhos de turma diferentes).

2. a) acumulada de $[4-6[$ = **24 alunos** b) ≥ 6 h: $40 - 24 = 16$ c) a acumulada soma progressivamente as frequências — nunca pode diminuir (não se «descontam» observações) e na última classe contém todos os N dados.

3. a) ex.: 45, 52, 55 b) **7** c) 40, 45, 50, 52, 55, 60, 65 → $AT = 65 - 40 = 25$.

4. $\min = 6 \cdot Q1 = 12 \cdot Q2 = 20 \cdot Q3 = 28 \cdot \max = 34 \rightarrow AT = 34 - 6 = 28 \cdot AIQ = 28 - 12 = 16$. Se $75\% \leq 28$ (é $Q3$): os **25% restantes** ficam espalhados entre 28 e 34 — apenas um quarto ocupa a faixa superior, enquanto metade (entre $Q1$ e $Q3$) vive dentro da caixa de largura 16.

5. $AT = 20 - 5 = 15 \rightarrow$ com 4 classes: largura = $15/4 = 3,75 \rightarrow$ arredonda-se a **4** (classes $[4-8[$, $[8-12[$, $[12-16[$, $[16-20[$): fi: $[4-8[= \{7,5,6\} = 3 \cdot [8-12[= \{11,9\} = 2 \cdot [12-16[= \{14,12,16\} = 3 \cdot [16-20[= \{18,20\} = 2 \cdot$ soma = **10** ✓.

6. a) $A: 120/400 = 30\%$ · B: $30/80 = 37,5\%$ b) **B** em % ($37,5\% > 30\%$) c) a absoluta interessa quando importa a *escala/impacto* (ex.: n° de vagas, alunos afetados) — 30% de 400 = 120 alunos é mais gente que $37,5\%$ de 80.

7. a) classe modal: **$[4-6[$ (14)** b) posição 20,5 → acumulada 24 na $[4-6[$, 10 na anterior → **dentro de $[4-6[$** c) as frequências sobem depressa até ao pico (4, 6, 14) e descem mais devagar à direita (10, 6 — 16 observações à direita do pico contra 10 à esquerda) → **assimetria à direita (positiva)**: a cauda estica-se em direção aos valores altos.

8. $14/2 = 7 \cdot 4/2 = 2 \cdot 3,5\times$ mais alta — com classes de larguras diferentes, a área (não só a altura) deve ser proporcional à frequência, senão o gráfico mente.

9. a) **$[0, 3, 2, 5]$** b) as 4 classes de largura 4 seriam $[0-4[$, $[4-8[$, $[8-12[$, $[12-16[$ — mas, por causa do $\min(\dots, 3)$, a última posição junta tudo o que é ≥ 12 c) $\text{int}(x//4) =$ parte inteira de $x/4 \rightarrow$ o índice da classe de largura 4 (0,1,2,3,...) e o $\min(\dots, 3)$ **prende tudo o que é ≥ 12 na 4.ª posição** (12, 14, 16, 18, 20 → 5) para não estourar o índice da lista — é uma simplificação: os dados 16–20 ficariam numa classe errada se fosse histograma a sério.

10. a) AIQ caixa 1: $28 - 12 = 16$ · caixa 2: $26 - 14 = 12 \rightarrow$ **caixa 1 tem AIQ maior** b) o 2.º é mais homogéneo (menor AIQ = metade central mais concentrada) c) 20 vs 19 — **não**: quase iguais, mas $20 > 19$ — mediana parecida não significa distribuições parecidas (dispersões diferentes).

11. a) se o eixo começa em 35, a barra de 40 mede 5 unidades e a de 80 mede 45 → parece $9\times$ maior embora o valor seja só $2\times$ — a área/altura mensura *distância ao zero artificial*, não o valor b) (1) eixo das ordenadas com **zero** (ou origem explicitamente marcada como truncada) (2) **mesclas** nunca lado a lado com escalas diferentes — se truncar, assinalar.

12. (orientação) com dados fornecidos: tabela 4 classes (largura = $AT/4$ arredondada), fi + relativas, histograma com barras coladas, 5 números ($Q2$ = médias dos valores centrais quando par), classe modal; comenta: se 90 kg num conjunto 55–75 é **outlier** $> Q3 + 1,5 \cdot AIQ$ — registar como anomalia/erro ou dado real a negociar com a amostra.