

CASO PRÁTICO 1

Estatística Descritiva – Tópicos de Resolução

Exercício 1:

Classifique as seguintes variáveis:

Nacionalidade (Portuguesa/Outra) Qualitativa nominal	Cor dos automóveis (Branco/Preto/Vermelho/Cinza/...) Qualitativa nominal	Número de erros num ditado Quantitativa discreta
Número de idas a um restaurante por mês Quantitativa discreta	Idade de um indivíduo (anos) Quantitativa contínua	Duração de uma bateria de telemóvel (horas) Quantitativa contínua
Rendimento familiar mensal (Baixo/Médio/Alto) Qualitativa ordinal	Marcas de automóveis (A/B/C) Qualitativa nominal	Peso de um indivíduo (kg) Quantitativa contínua
Rendimento familiar mensal (euros) Quantitativa contínua	Altura de um indivíduo (m) Quantitativa contínua	Salário mensal (euros) Quantitativa contínua
Escalão etário (menos de 18 anos/18-30 anos/mais de 30 anos) Qualitativa ordinal		

Exercício 2:

Numa empresa recolheu-se informação sobre algumas características de alguns dos seus colaboradores. Alguns dos outputs obtidos foram os seguintes.

Sexo (Feminino/Masculino)

Feminino	Masculino	Feminino	Masculino
32	6	84.2	15.8

Idade (anos)

25	26	27	28	29	32	44	25	26	27	28	29	32	44
15	12	6	1	2	1	1	39.5	31.6	15.8	2.6	5.3	2.6	2.6

2.1 Classifique as variáveis.

Sexo (Feminino/Masculino) → qualitativa nominal

Idade (anos) → Quantitativa contínua

2.2 Analise os resultados.

Participaram no estudo 38 colaboradores, 32 (84,2%) são do sexo feminino e 6 (15,8%) do sexo masculino.

A idade dos colaboradores inquiridos situa-se entre 25 e 44 anos.

- **15 (39,5%) dos colaboradores têm 25 anos;**
- **12 (31,6%) têm 26 anos;**
- **6 (15,8%) têm 27 anos;**
- **2 (5,3%) têm 29 anos;**
- **existe um colaborador com 28 anos, um com 32 anos e outro com 44 anos.**

Verifica-se maior concentração das idades nos menores valores, 71,1% (27) dos colaboradores têm 25 ou 26 anos (assimetria à direita).

A idade mais frequente entre os colaboradores é 25 anos.

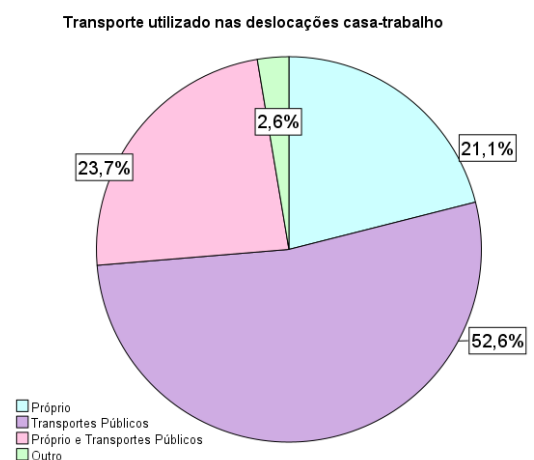
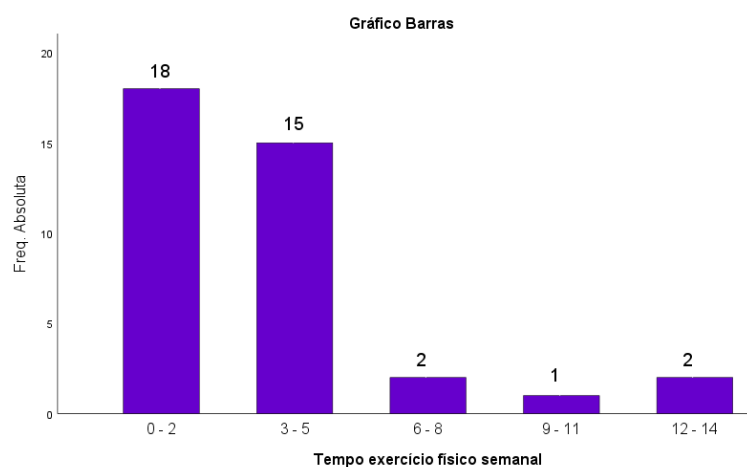
Exercício 3:

Além das características indicadas no exercício anterior, também se recolheu informação sobre as seguintes variáveis:

Transporte utilizado pelos colaboradores nas deslocações casa-trabalho (Próprio/Transportes Públicos/Próprio e Transportes Públicos/ Outro)

Tempo semanal dedicado à prática de exercício físico (0-2/3-5/6-9/9-11/12-14/15 ou mais)

Número de irmãos





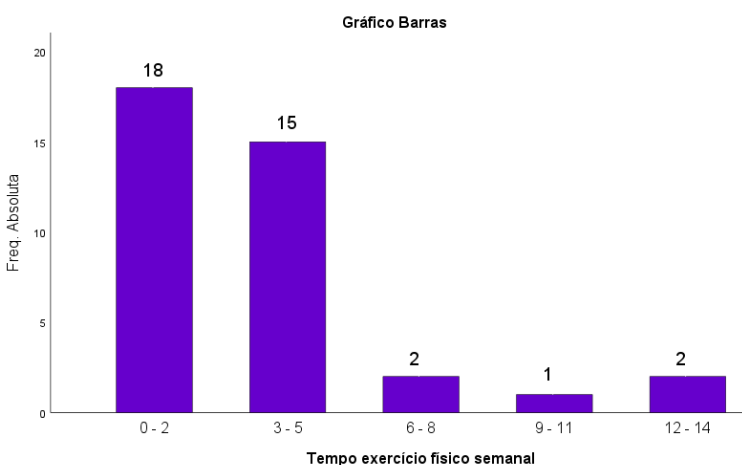
3.1 Classifique as variáveis.

Transporte utilizado pelos colaboradores nas deslocações casa-trabalho (Próprio/Transportes Públicos/Próprio e Transportes Públicos/ Outro) → qualitativa nominal

Tempo semanal dedicado à prática de exercício físico (0-2 horas/3-5 horas/6-9 horas/9-11 horas/12-14 horas /15 ou mais horas) → qualitativa ordinal

Número de irmãos → quantitativa discreta

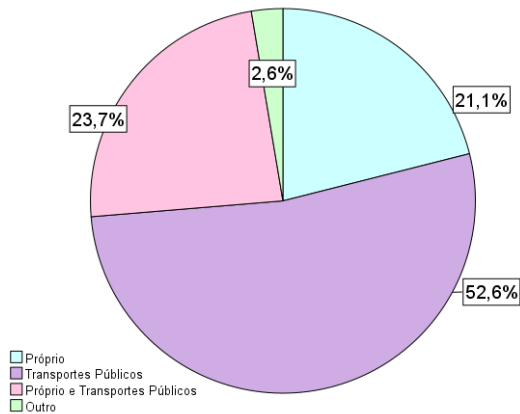
3.2 Analise os outputs.



Dos 38 colaboradores que participaram no estudo:

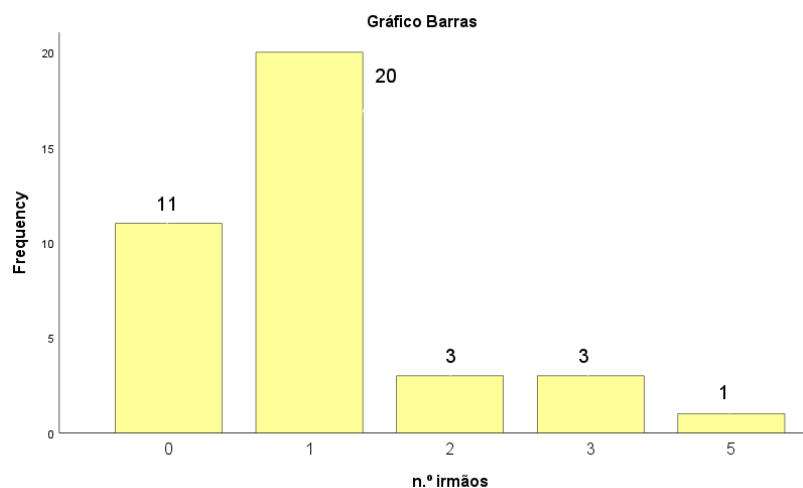
- 18 indicaram que praticam até duas horas de exercício físico por semana;
- 15 indicaram que praticam 3 a 5 horas de exercício físico por semana;
- 2 indicaram que praticam 6 a 8 ou 12 a 14 horas de exercício físico por semana;
- 1 indicou que pratica 9 a 11 horas de exercício físico por semana;
- nenhum indicou praticar 15 ou mais horas de exercício físico por semana por semana;
- 86,8% dos colaboradores indicaram que praticam no máximo 5 horas de exercício físico por semana (assimetria à direita);
- O intervalo mais frequente é 0 a 2 horas.

Transporte utilizado nas deslocações casa-trabalho



A maioria dos colaboradores inquiridos (52,6%) utiliza transportes públicos nas deslocações casa-trabalho (meio de transporte utilizado mais frequentemente).

Quase 24% dos colaboradores inquiridos conjugam transporte próprio e transportes públicos, 21,1% utilizam transporte próprio e 2,6% utilizam outro meio de transporte nas deslocações casa-trabalho.



O número de irmão dos colaboradores inquiridos da empresa situa-se entre 0 e 5.

- 11 colaboradores não têm irmãos;
- 20 colaboradores têm 1 irmão;
- 3 colaboradores têm 2 ou 3 irmãos;
- 1 colaborador tem 5 irmãos;
- É evidente a concentração das observações nas duas primeiras classes, 31 (81,6%=31/33) colaboradores têm no máximo 1 irmão (assimetria à direita).

Exercício 4:

As notas dos alunos de uma turma são:

9, 15, 15, 17, 14, 16, 19, 10, 14, 15, 18, 16, 13, 17

Calcule a média, mediana e o desvio padrão e interprete os valores obtidos.

$n=14$

$$\bar{x} = \frac{9 + 15 + 15 + \dots + 16 + 13 + 17}{14} = 14,9$$

A nota média dos alunos da turma é 14,9 valores (quase 15 valores).

me=?

Ordenar a amostra e identificar os elementos centrais

9 10 13 14 14 15 15 16 16 17 17 18 19

9 10 13 14 14 15 15 15 15 16 16 17 17 18 19

$$me=(15+15)/2=15$$

50% dos alunos da turma têm no máximo 15 valores e os restantes alunos têm no mínimo 15 valores.

ou

50% dos alunos têm entre 9 e 15 valores e os restantes 50% têm entre 15 e 19 valores.

$S'=?$

$$s' = \sqrt{\frac{(9 - 14,9)^2 + (10 - 14,9)^2 + \dots + (19 - 14,9)^2}{14 - 1}} = 2,9$$

$9 - 14,9 \rightarrow$ desvio à média

$10 - 14,9$

A nota média dos alunos da turma é 14,9 valores com uma dispersão de 2,9 valores.

Exercício 5:

Analizando os seguintes *outputs* caracterize o Peso (kg), a Altura (cm) e a Idade (anos) dos indivíduos que participaram num estudo.

Peso (kg) → quantitativa contínua

Altura (cm) → quantitativa contínua

Idade (anos) → quantitativa contínua

Peso

N	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd
33	46.00	57.00	64.00	65.18	69.00	96.00	11.485

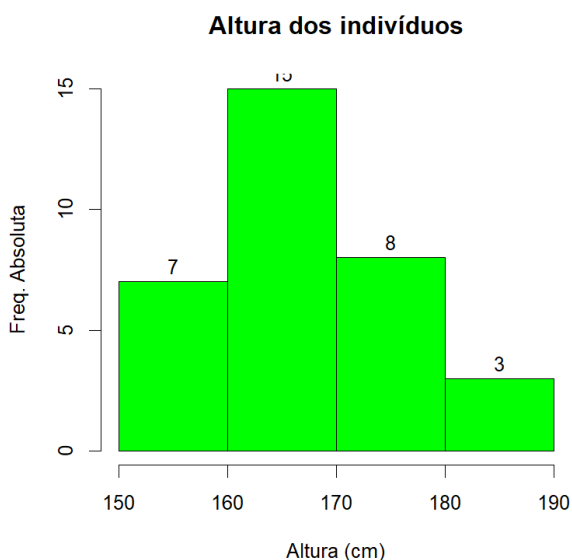
O peso dos 33 inquiridos situa-se entre 46 e 96 Kg.

Em média, os inquiridos pesam cerca de 65 kg (65,18) com uma dispersão de quase 11,5 kg (11,485).

50% (ou metade) dos inquiridos pesa entre 46 e 64 kg e os restantes 50% pesam entre 64 e 96 kg.

25% dos inquiridos pesam entre 46 e 57 kg e os restantes 75% pesam entre 57 e 96 kg.

75% dos inquiridos pesam entre 46 e 69 kg e os restantes 25% pesam entre 69 e 96 kg.



As alturas dos inquiridos situam-se entre 150 e 190 cm.

- **Verifica-se uma maior concentração das alturas entre 160 e 170 cm (15 ou 45,5%)→ alturas mais frequentes situam-se entre 160 e 170 cm;**
- **8 indivíduos têm 170 a 180 cm;**
- **7 indivíduos têm 150 a 160 cm;**
- **3 indivíduos têm 180 a 190 cm;**

Idade

vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Idade	33	21.03	1.59	21	20.7	1.48	20	28	8	2.75	8.76	0.28

A idade dos 33 inquiridos situa-se entre 20 e 28 anos (a amplitude das idades é igual a $28-20=8$). Em média, os inquiridos têm cerca de 21 anos com uma dispersão de 1,6 anos, aproximadamente. O valor da trimédia (20,79) é semelhante ao da média (21,03) o que indica que “não existem idades extremas” (muito altas ou muito pequenas).

AIQ=1 → comprimento do intervalo que contém 50% das idades (centrais), entre o Q1 e o Q3.

Coef. Assimetria (skewness)=2,75 > 0 → distribuição assimétrica à direita (maior concentração nas menores idades).

Coef. Curtose (achatamento)=8,76 > 0 → distribuição leptocúrtica

Nota:

standard error (útil para obter IC para a Idade média)

$$se = \frac{1,59}{\sqrt{33}} = 0,28$$

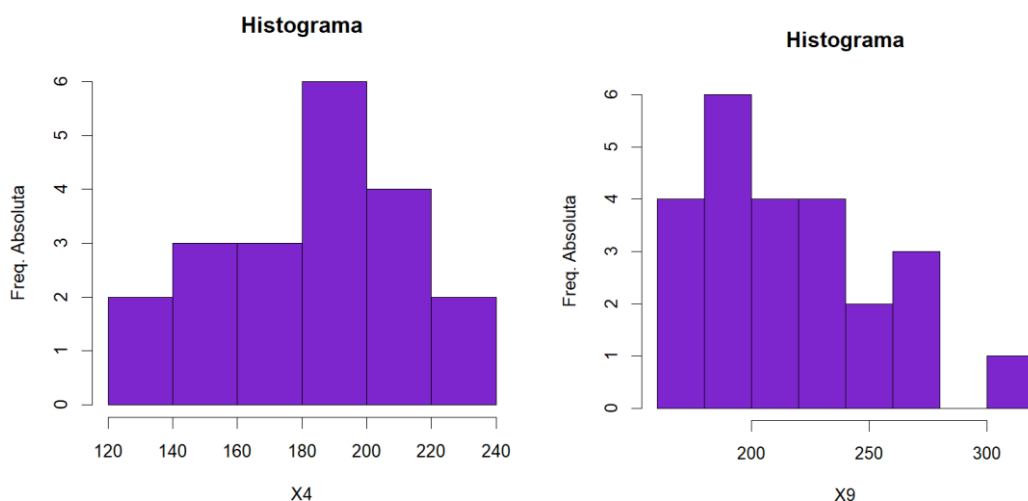
Exercício 6:

Que ferramenta(s) estatística(s) pode utilizar para analisar a simetria/assimetria da distribuição de frequências?

Pode utilizar-se representações gráficas (gráfico de barras/histograma/Box-Plot), comparar medidas de localização central (média, moda e mediana) e o coeficiente de assimetria.

Exercício 7:

Analisando os outputs seguintes estude a simetria/assimetria dos dados.



vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis
x4	20	186	31.23	192.5	188.44	25.95	120	230	110	-0.52	-0.78
se											
x4	6.98										

vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis
x9	24	218.33	40.45	212.5	216.25	37.06	160	310	150	0.52	-0.74
se											
x9	8.26										

X4 e X9 quantitativas contínuas

Analisando o histograma da variável X4 é evidente uma maior concentração das observações nos maiores valores → assimetria à esquerda

Média=186 < mediana=192,5 → assimetria à esquerda

$g_1 = -0,52 < 0$ → assimetria à esquerda

Analisando o histograma da variável X9 é evidente uma maior concentração das observações nos menores valores → assimetria à direita

Média=218,33 > mediana=212,5 → assimetria à direita

$g_1 = 0,52 > 0$ assimetria à direita

Exercício 8:

- a. Considere a seguinte amostra

12, 24, 37, 42, 55, 62, 72, 77, 246, 592

Qual(ais) a(s) medida(s) de tendência central e de dispersão adequadas para caracterizar a amostra? Justifique.

- b. Que medida(s) é(são) adequada(s) para caracterizar as variáveis Sexo, Curso e Estado civil?

a. $n=10$

Média da amostra=121,9

Mediana=(55+62)/2=58,5

média > mediana → assimetria à direita

Como a distribuição é assimétrica, a mediana é mais adequada do que a média e a amplitude interquartil é mais adequada que o desvio padrão.

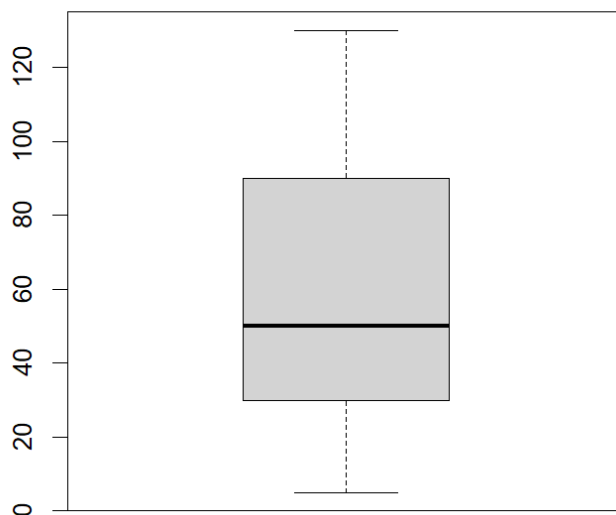
b. Sexo, Curso e Estado Civil são variáveis qualitativas nominais → moda

Exercício 9:

Considere as box-plot seguintes e identifique (aproximadamente) os 5 números.

5 números → mínimo, máximo, Q1, mediana e Q3

Box-Plot



Mínimo → entre 0 e 10

Máximo → superior a 120

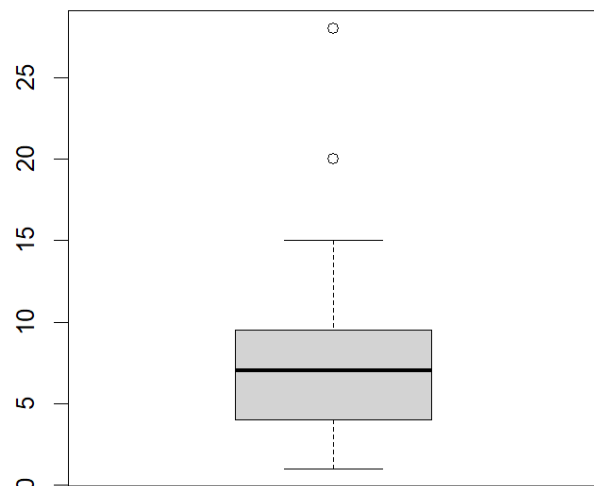
Q1 → entre 20 e 40

Mediana → entre 40 a 60

Q3 → entre 80 e 100

Assimetria à direita (maior concentração nos menores valores).

Box-Plot



Mínimo → entre 0 e 2,5

Máximo → superior a 25

Q1 → quase 5

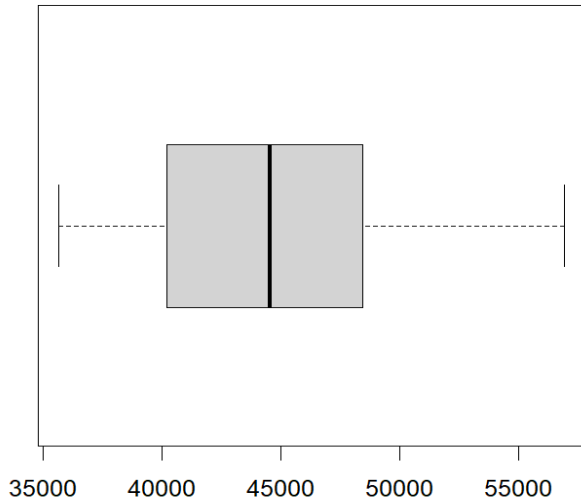
Mediana → entre 5 a 10

Q3 → quase 10

Existem 2 outliers

Assimetria à direita (maior concentração nos menores valores).

Box-Plot



Mínimo → ligeiramente superior a 35000

Máximo → superior a 55000

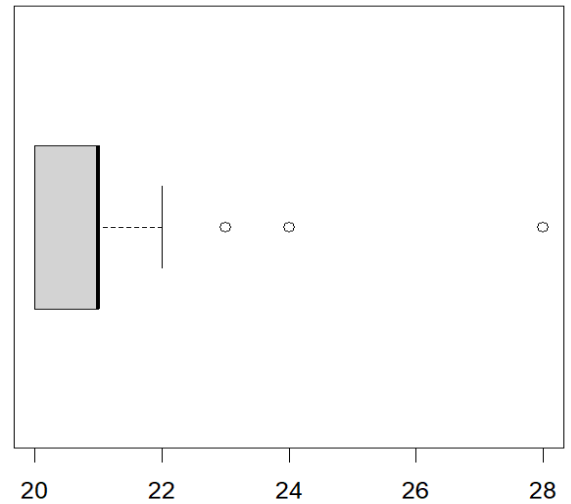
Q1 → cerca de 40000

Mediana → quase 45000

Q3 → entre 45000 e 50000

Assimetria à direita.

Box-Plot



Mínimo=Q1=20

Máximo=28

Mediana=Q3 → entre 20 e 22

Existem 3 *outliers*

Evidente assimetria à direita.