

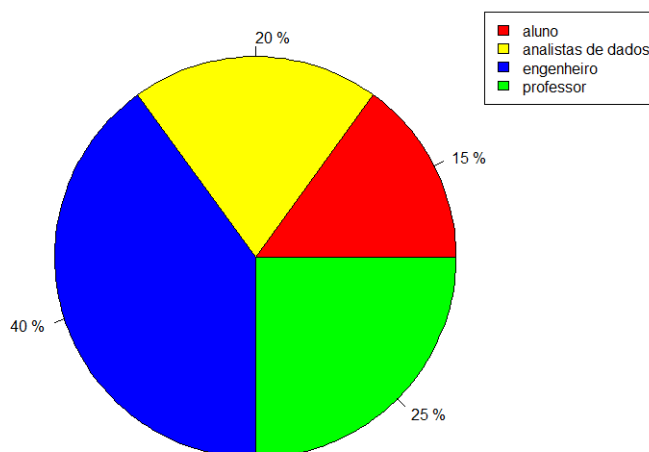
- A ficha de trabalho foi resolvida recorrendo ao software R: ver script ficha de trabalho 3.

1. (a) -

(b) Tabela de frequências:

i	x_i	n_i	f_i
1	aluno	12	$\frac{12}{80} = 0.15$
2	analista de dados	16	$\frac{16}{80} = 0.20$
3	engenheiro	32	$\frac{32}{80} = 0.40$
4	professor	20	$\frac{20}{80} = 0.25$
		$n = 80$	1

Representação gráfica:



Medidas:

como os dados são qualitativos, a única medida possível de calcular é a moda

moda = engenheiro

2. (a) Variável Género

como os dados são qualitativos, a única medida possível de calcular é a moda

moda = Masculino

Variável Idade

como os dados são quantitativos, é possível calcular todas as medidas

i. medidas de localização central:

$$\begin{aligned}\text{moda} &= 18 \text{ anos} \\ \text{média} &= \bar{x} = 24.3126 \text{ anos} \\ \text{mediana} &= \tilde{x} = Q_2 = Q_{0.50} = 22.7779 \text{ anos}\end{aligned}$$

ii. quartis:

$$\begin{aligned}1^\circ \text{quartil} &= Q_{0.25} = 19.9462 \text{ anos} \\ 2^\circ \text{quartil} &= Q_{0.50} = \text{mediana} = \tilde{x} = 22.7779 \text{ anos} \\ 3^\circ \text{quartil} &= Q_{0.75} = 26 \text{ anos}\end{aligned}$$

iii. $D_9 = Q_{0.90} = 33.2268 \text{ anos}$

iv. $P_3 = Q_{0.03} = 17.0385 \text{ anos}$

v. extremos:

$$\begin{aligned}\text{mínimo} &= 14 \text{ anos} \\ \text{máximo} &= 61 \text{ anos}\end{aligned}$$

vi. medidas de dispersão absoluta:

$$\begin{aligned}\text{Amplitude Total} &= 47 \text{ anos} \\ \text{Amplitude Interquartil} &= \text{AIQ} = 6.0538 \text{ anos} \\ \text{variância} &= s^2 = 40.2713 \text{ anos}^2 \\ \text{desvio padrão} &= s = 6.346 \text{ anos}\end{aligned}$$

vii. medida de dispersão relativa:

$$\text{coeficiente de variação} = \text{CV} = \frac{6.346}{24.3126} \times 100 = 26.1016\%$$

viii. $b_1 = 1.5269$

Variável Altura

como os dados são quantitativos, é possível calcular todas as medidas

i. medidas de localização central:

$$\begin{aligned}\text{moda} &= 1.70 \text{ metros} \\ \text{média} &= \bar{x} = 1.7017 \text{ metros} \\ \text{mediana} &= \tilde{x} = Q_2 = Q_{0.50} = 1.7005 \text{ metros}\end{aligned}$$

ii. quartis:

$$1^{\circ}\text{quartil} = Q_{0.25} = 1.63 \text{ metros}$$

$$2^{\circ}\text{quartil} = Q_{0.50} = \text{mediana} = \tilde{x} = 1.7005 \text{ metros}$$

$$3^{\circ}\text{quartil} = Q_{0.75} = 1.7685 \text{ metros}$$

iii. $D_9 = Q_{0.90} = 1.8238 \text{ metros}$

iv. $P_3 = Q_{0.03} = 1.5299 \text{ metros}$

v. extremos:

$$\text{mínimo} = 1.45 \text{ metros}$$

$$\text{máximo} = 1.98 \text{ metros}$$

vi. medidas de dispersão absoluta:

$$\text{Amplitude Total} = 0.53 \text{ metros}$$

$$\text{Amplitude Interquartil} = \text{AIQ} = 0.1385 \text{ metros}$$

$$\text{variância} = s^2 = 0.0087 \text{ metros}^2$$

$$\text{desvio padrão} = s = 0.0933 \text{ metros}$$

vii. medida de dispersão relativa:

$$\text{coeficiente de variação} = \text{CV} = \frac{0.0933}{1.7017} \times 100 = 5.4831\%$$

viii. $b_1 = -0.0128$

Variável Peso

como os dados são quantitativos, é possível calcular todas as medidas

- medidas de localização central:

$$\text{moda} = 80 \text{ kg}$$

$$\text{média} = \bar{x} = 86.5861 \text{ kg}$$

$$\text{mediana} = \tilde{x} = Q_2 = Q_{0.50} = 83 \text{ kg}$$

- quartis:

$$1^{\circ}\text{quartil} = Q_{0.25} = 65.4239 \text{ kg}$$

$$2^{\circ}\text{quartil} = Q_{0.50} = \text{mediana} = \tilde{x} = 83 \text{ kg}$$

$$3^{\circ}\text{quartil} = Q_{0.75} = 107.4827 \text{ kg}$$

- $D_9 = Q_{0.90} = 120.9215 \text{ kg}$

- $P_3 = Q_{0.03} = 44.2601 \text{ kg}$

- extremos:

$$\text{mínimo} = 39 \text{ kg}$$

$$\text{máximo} = 173 \text{ kg}$$

- medidas de dispersão absoluta:

$$\text{Amplitude Total} = 134 \text{ kg}$$

$$\text{Amplitude Interquartil} = \text{AIQ} = 42.0587 \text{ kg}$$

$$\text{variância} = s^2 = 685.9775 \text{ kg}^2$$

$$\text{desvio padrão} = s = 26.1912 \text{ kg}$$

- medida de dispersão relativa:

$$\text{coeficiente de variação} = \text{CV} = \frac{26.1912}{86.5861} \times 100 = 30.2487\%$$

- $b_1 = 0.255$

Variável FAVC

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = 1 = \text{Sim}$$

Variável FCVC

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = 2 = \text{Às vezes}$$

Variável NCP

como os dados são quantitativos, é possível calcular todas as medidas

- medidas de localização central:

$$\text{moda} = 3 \text{ refeições principais}$$

$$\text{média} = \bar{x} = 2.6878 \text{ refeições principais}$$

$$\text{mediana} = \tilde{x} = Q_2 = Q_{0.50} = 3 \text{ refeições principais}$$

- quartis:

$$1^\circ \text{quartil} = Q_1 = Q_{0.25} = 3 \text{ refeições principais}$$

$$2^\circ \text{quartil} = Q_{0.50} = \text{mediana} = \tilde{x} = 3 \text{ refeições principais}$$

$$3^\circ \text{quartil} = Q_{0.75} = 3 \text{ refeições principais}$$

- $D_9 = Q_{0.90} = 3 \text{ refeições principais}$

- $P_3 = Q_{0.03} = 1 \text{ refeição principal}$

- extremos:

mínimo = 1 refeição principal
 máximo = 4 refeições principais

- medidas de dispersão absoluta:

Amplitude Total = 3 refeições principais
 Amplitude Interquartil = AIQ = 0 refeições principais
 variância = $s^2 = 0.6556$ refeições principais²
 desvio padrão = $s = 0.8097$ refeições principais

- medida de dispersão relativa:

$$\text{coeficiente de variação} = CV = \frac{0.8097}{2.6878} \times 100 = 30.124\%$$

- $b_1 = Q_{0.90} = -1.0667$

Variável CAEC

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = S = \text{Às vezes}$$

Variável Fumar

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = 0 = \text{Não}$$

Variável CH2O

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = 2 = \text{entre 1 e 2 litros}$$

Variável FAF

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = 1 = 1 \text{ ou } 2 \text{ dias}$$

Variável CALC

como os dados são qualitativos, a única medida possível de calcular é a moda

$$\text{moda} = S = \text{Às vezes}$$

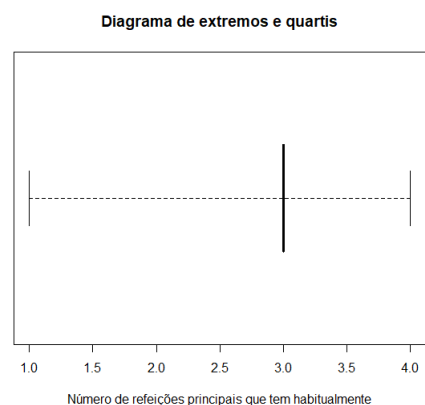
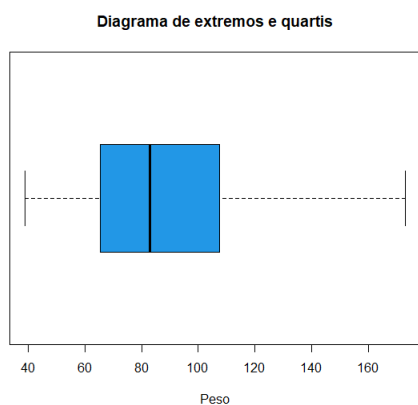
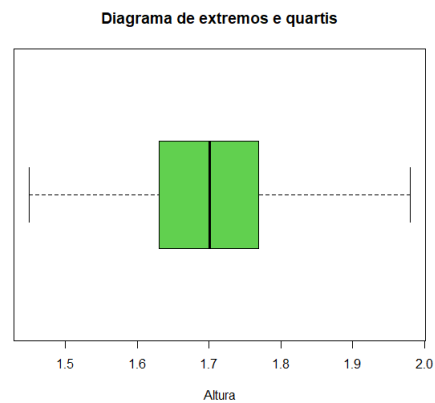
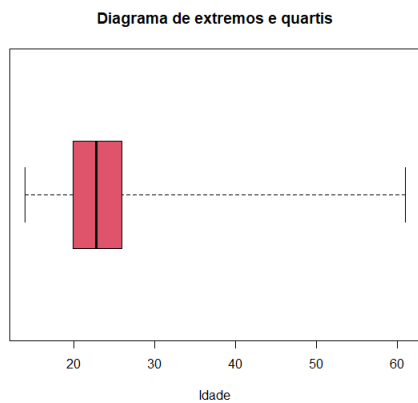
Variável MTRANS

como os dados são qualitativos, a única medida possível de calcular é a moda

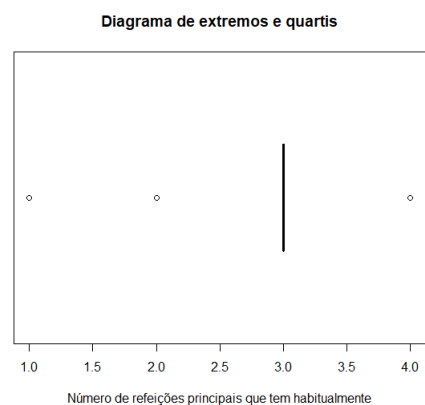
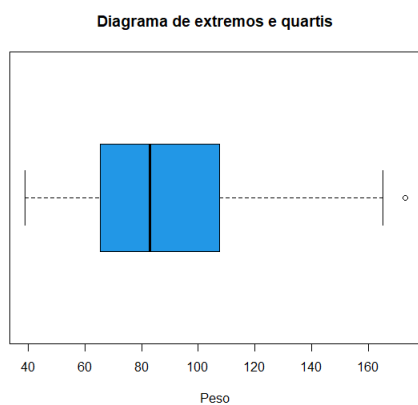
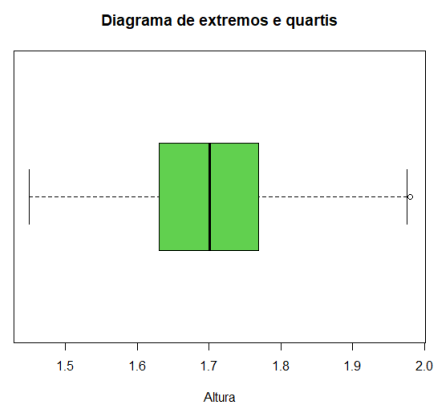
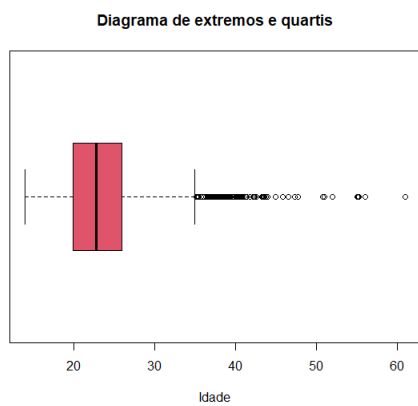
$$\text{moda} = \text{Transportes públicos}$$

(b) O diagrama de extremos e quartis só pode ser calculado para as variáveis quantitativas:

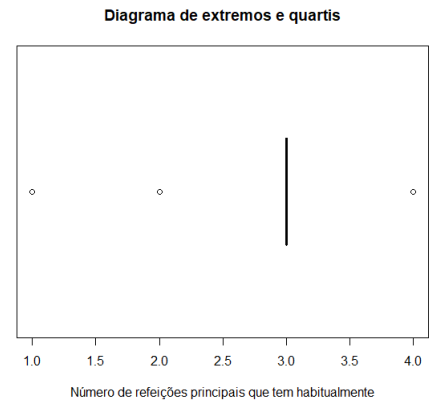
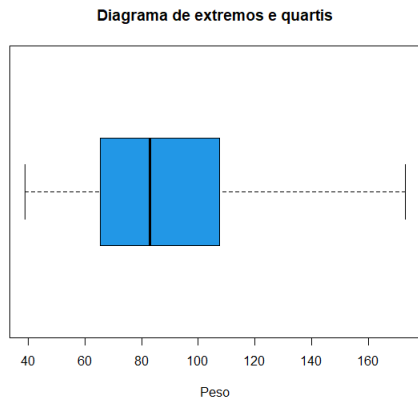
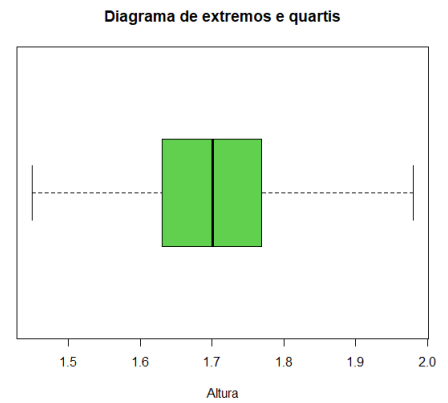
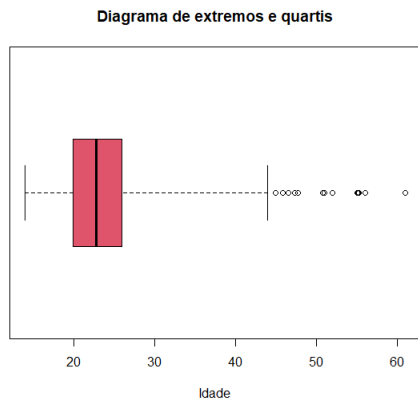
i. sem indicação de outliers



ii. com indicação de outliers a partir dos moderados



iii. com indicação de outliers a partir dos severos



(c) Variável Idade

- Como $\text{moda} < \text{mediana} < \text{média}$, os dados aparentam distribuir-se de forma assimétrica positiva.
- Como $b_1 > 0$, os dados aparentam distribuir-se de forma assimétrica positiva.

Variável Altura

- Como $\text{moda} \approx \text{mediana} \approx \text{média}$, os dados aparentam distribuir-se de forma simétrica.
- Como $b_1 \approx 0$, os dados aparentam distribuir-se de forma simétrica.

Variável Peso

- Como $\text{moda} < \text{mediana} < \text{média}$, os dados aparentam distribuir-se de forma assimétrica positiva.
- Como $b_1 > 0$, os dados aparentam distribuir-se de forma assimétrica positiva.

Variável NCP

- Como $\text{moda} = \text{mediana} > \text{média}$, os dados aparentam distribuir-se com uma ligeira assimetria negativa, não parecendo afastar-se muito da simetria.
- Como $b_1 < 0$, os dados aparentam distribuir-se de forma assimétrica negativa.

(d) Feito no R.

- (e) Em relação à idade, o género masculino apresenta as medidas de localização (excepto a moda) ligeiramente superiores e as medidas de dispersão (exceto AIQ) ligeiramente inferiores do que o género feminino. Em relação à altura, o género masculino apresenta as medidas de localização superiores e as medidas de dispersão ligeiramente inferiores do que o género feminino. Em relação ao peso, o género masculino apresenta as medidas de localização superiores e as medidas de dispersão inferiores do que o género feminino. Em relação à variável "Se come habitualmente vegetais nas refeições", no género feminino a maioria respondeu "Sempre" enquanto no género masculino a maioria respondeu "Às

vezes". Em relação à variável "Com que frequência pratica atividade física por semana", no género feminino a maioria respondeu "não pratica" enquanto no género masculino a maioria respondeu "1 ou 2 dias". Nas restantes variáveis não há diferenças.

(f) i. Classe modal =]20, 26] anos

pois é a classe com a maior frequência absoluta

Classe dos quartis

- Classe do 1.^o quartil = Classe do $Q_{0.25}$ = [14, 20] anos
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.25
- Classe do 2.^o quartil = Classe do $Q_{0.50}$ =]20, 26] anos
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.50
- Classe do 3.^o quartil = Classe do $Q_{0.75}$ =]20, 26] anos
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.75

ii. Classe modal =]1.6, 1.8] metros

pois é a classe com a maior frequência absoluta

Classe dos quartis

- Classe do 1.^o quartil = Classe do $Q_{0.25}$ =]1.6, 1.8] metros
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.25
- Classe do 2.^o quartil = Classe do $Q_{0.50}$ =]1.6, 1.8] metros
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.50
- Classe do 3.^o quartil = Classe do $Q_{0.75}$ =]1.6, 1.8] metros
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.75

iii. Classe modal = [77.3, 96.4[kg

pois é a classe com a maior frequência absoluta

Classe dos quartis

- Classe do 1.^o quartil = Classe do $Q_{0.25}$ = [58.1, 77.3[kg
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.25
- Classe do 2.^o quartil = Classe do $Q_{0.50}$ = [77.3, 96.4[kg
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.50
- Classe do 3.^o quartil = Classe do $Q_{0.75}$ = [96.4, 116[kg
pois é a classe onde se encontra a primeira frequência relativa acumulada que ultrapassa 0.75