

Exercício 4

1)

a)

Lesões:

1: N = 131 (35.7%)

2: N = 123 (33.5%)

3: N = 113 (30.8%)

N = 367

15 valores omissos

Lesão	Frequência
1	131
2	123
3	113

Qualitativa nominal

HA:

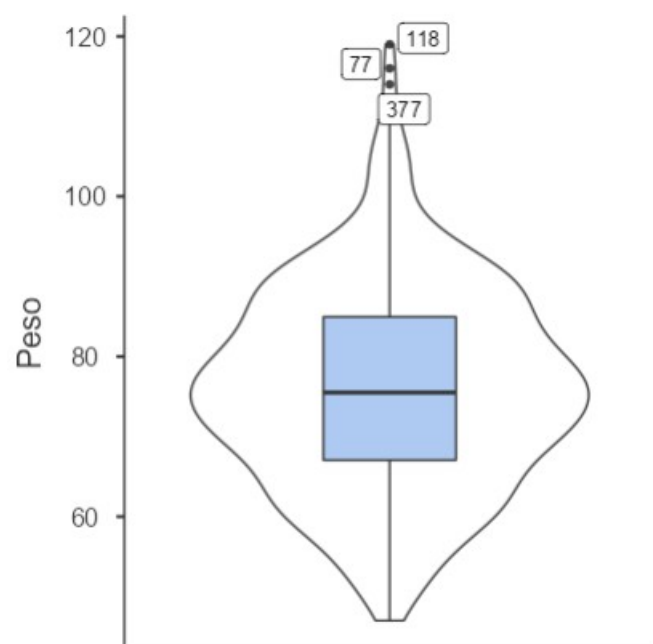
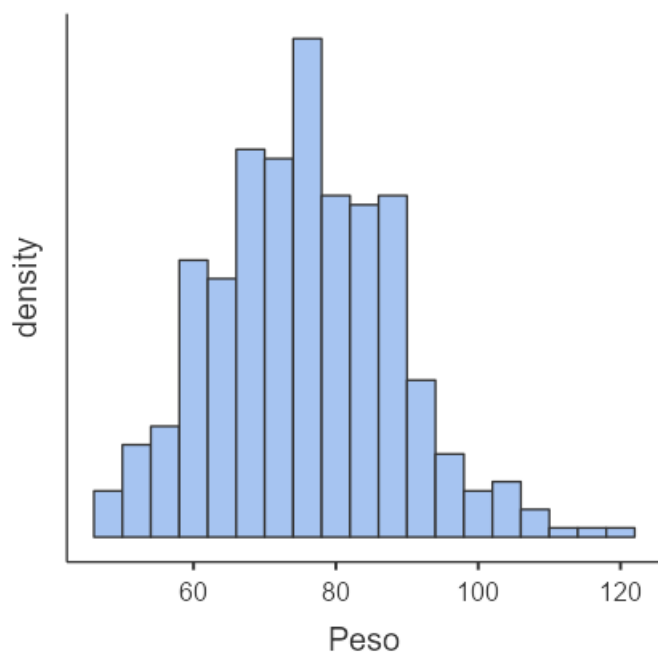
0: N = 160 (43.2%)

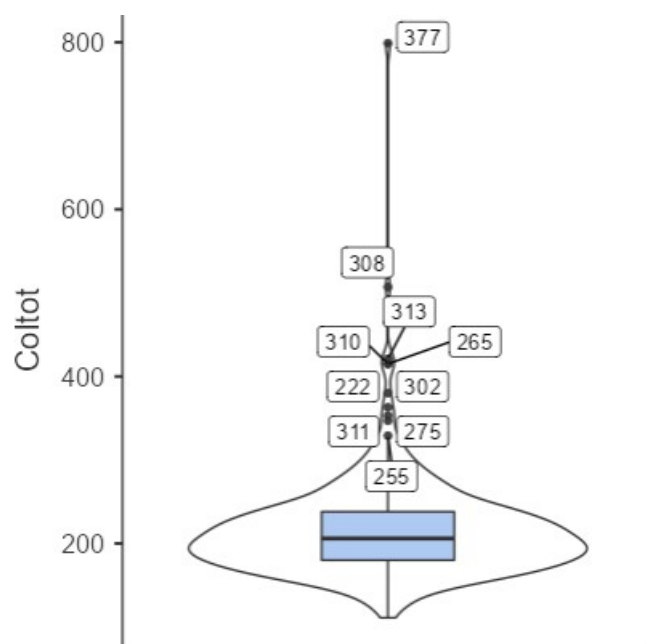
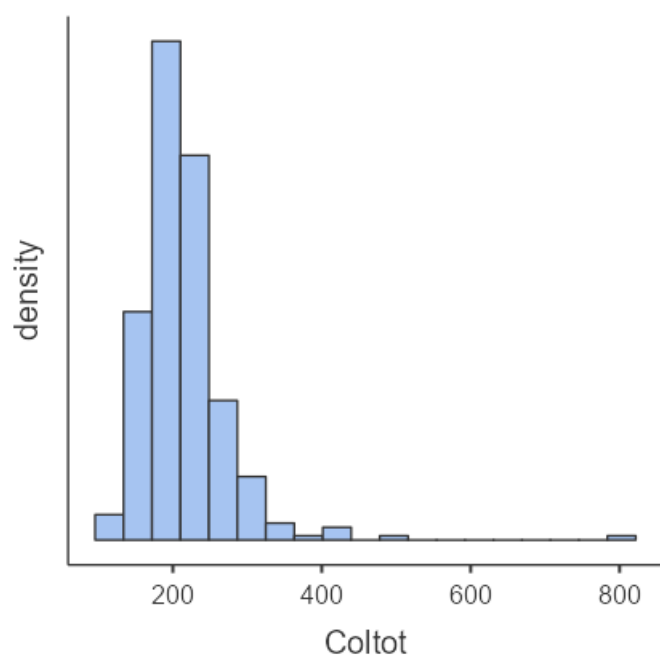
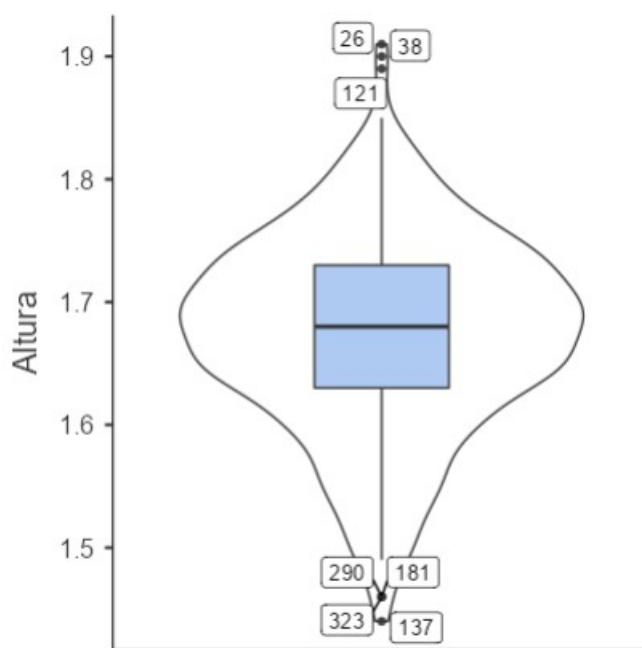
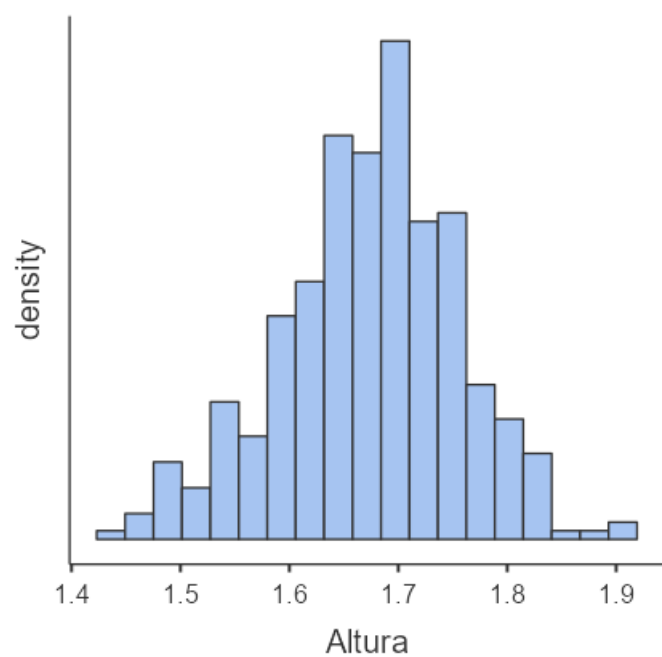
1: N = 210 (56.8%)

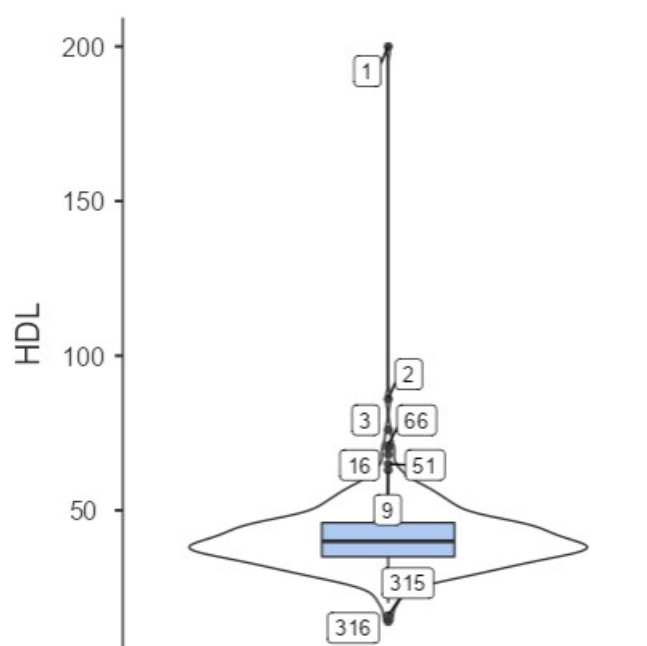
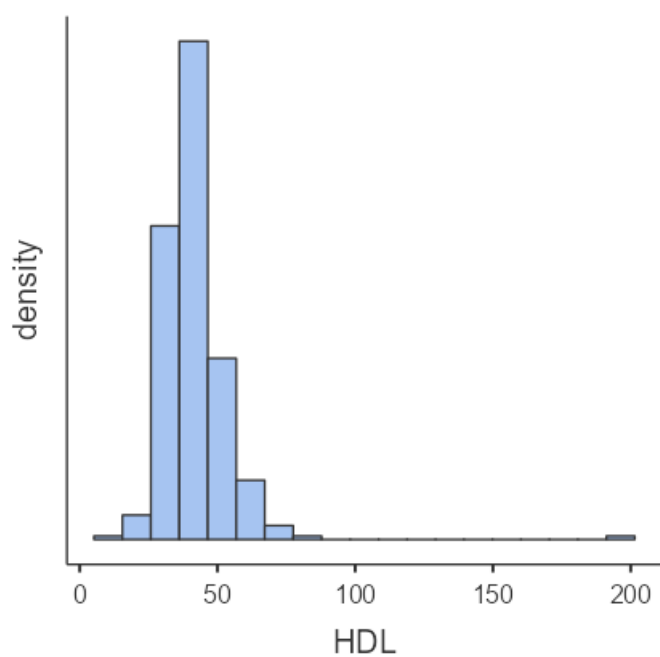
Qualitativa nominal

HA	Frequência
0	160
1	210

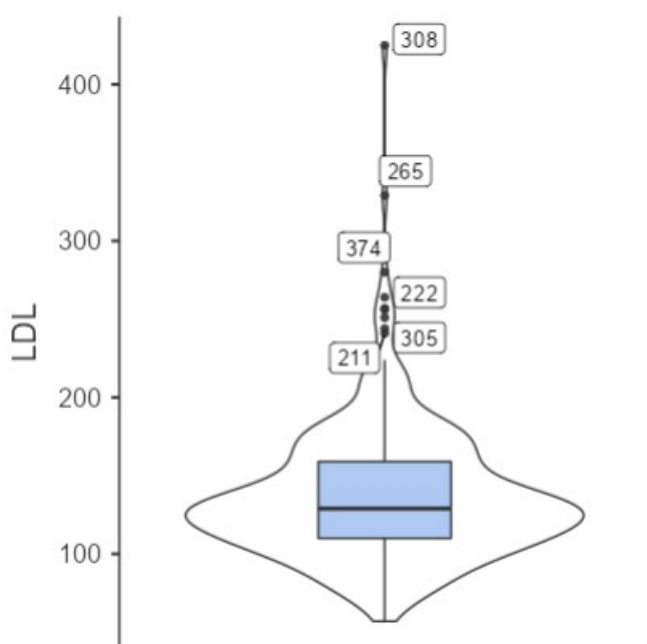
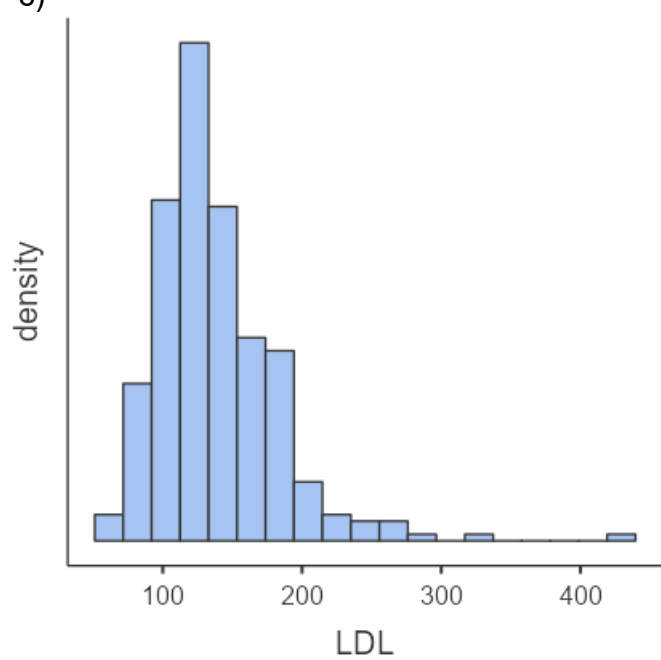
b)







c)



d)

A variável altura segue a distribuição normal de forma satisfatória. A variável peso se distancia do normal em seus menores e maiores valores. O restante das variáveis seguem pouco a reta normal, se destoando bastante nos menores e maiores valores.

2)

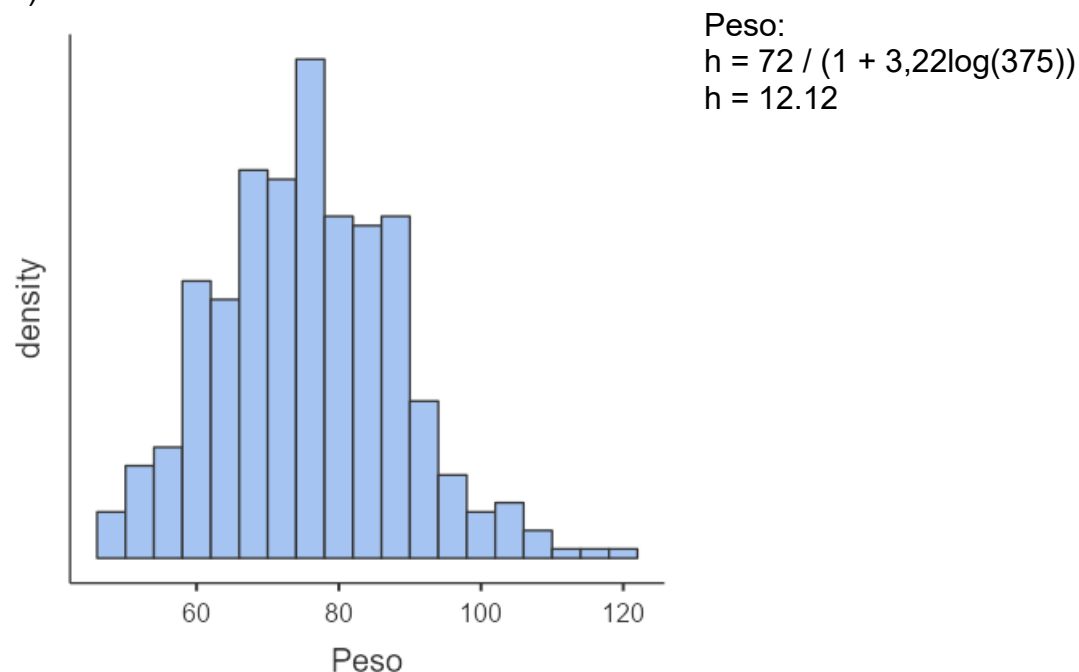
a)

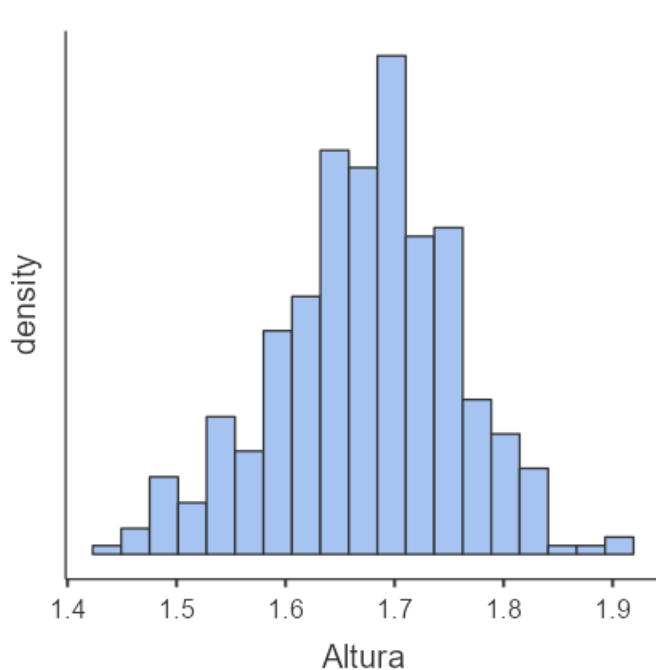
Variável	Média	Mediana	Moda	Desvio-padrão	Mínimo	Máximo	Amplitude
idade	68.4	69	59	15.1	19	110	91
tmunic	49.3	50	50	20.6	0	103	103
htransp	1.78	0.83	0	2.54	0	12	12
cargatabag	25.6	1.5	0	39.3	0	256	256
antracose	0.202	0.17	0	0.16	0	0.829	0.829
ses	-0.238	-0.310	-0.310	0.362	-1	1	2
densid	0.0176	0.018	0.017	0.0069	0	0.037	0.037
distmin	171	114	16	191	0	1964	1964

b)

Aumentar a dispersão dos valores e reduzir a quantidade de outliers ao invés de concentrar grande parte da densidade em valores próximos

3)





Altura:
 $h = 0.47 / (1 + 3,22\log(376))$
 $h = 0.79$

4)

Com o valor atípico:

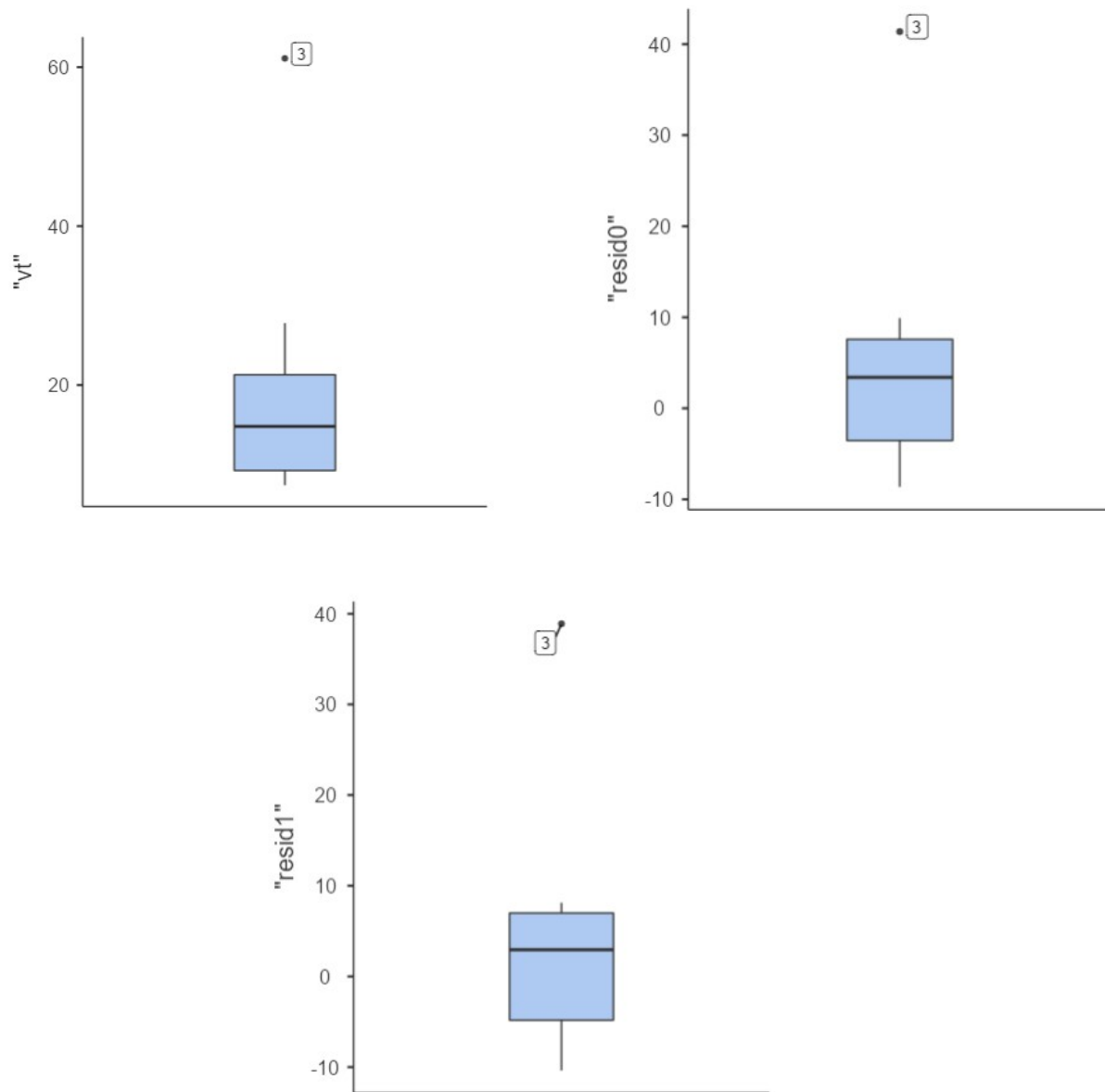
Variável	Média	Moda	Q1	Q2	Q3	Variância	Desvio-padrão	Amplitude
vt	18.4	7.4	9.25	14.8	21.3	183	13.5	53.7

Sem o valor atípico:

Variável	Média	Moda	Q1	Q2	Q3	Variância	Desvio-padrão	Amplitude
vt	15.3	7.4	8.3	13.9	20.4	46.7	6.84	20.4

Nota-se que após a retirada do valor a moda não sofre impacto, os valores de média, primeiro, segundo e terceiro quartis sofrem uma pequena alteração, no entanto percebe-se uma diferença massiva nos novos valores das medidas de dispersão, com destaque para a variância.

5)



6)

Variável	P = 1/4	P = 1/3	P = 1/2	P = 3/4
vt	9.25	11.5	14.8	21.3
resid0	-3.54	-0.908	3.4	7.58
resid1	-4.83	-1.46	2.94	6.99

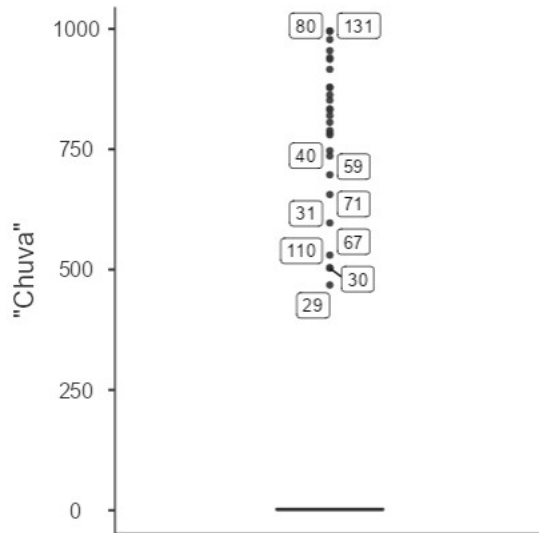
7)

Quantitativa contínua.

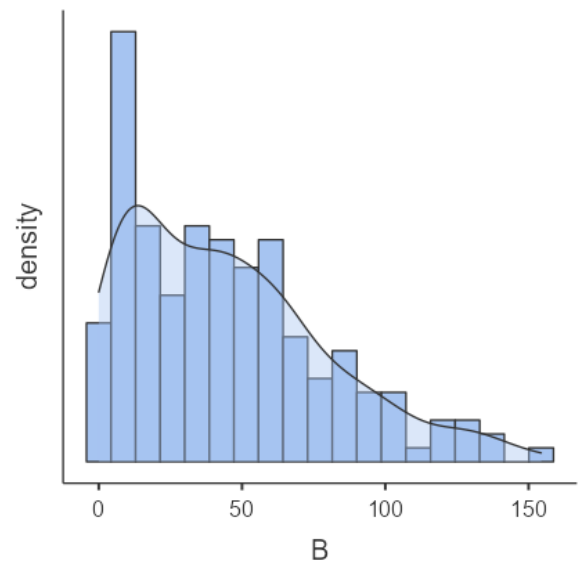
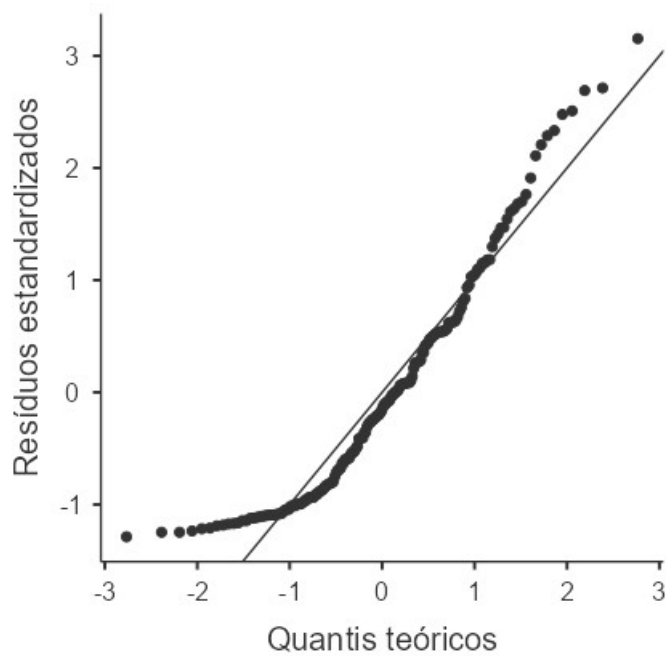
8)

Quantitativa discreta

9)



10)



11)

vt:

$Q1 = 9.25$

$Q2 = 14.8$

$Q3 = 21.3$

$$A = ((21.3 - 14.8) - (14.8 - 9.25)) / (21.3 - 9.25)$$

$$A = 0.8$$

resid0

$$Q1 = -3.54$$

$$Q2 = 3.4$$

$$Q3 = 7.58$$

$$A = ((7.58 - 3.4) - (3.4 - -3.54)) / (7.58 - -3.54)$$

$$A = -0.22$$

resid1

$$Q1 = -4.83$$

$$Q2 = 2.94$$

$$Q3 = 6.99$$

$$A = ((6.99 - 2.94) - (2.94 - -4.83)) / (6.99 - -4.83)$$

$$A = -0.31$$

12)

a)

Pode-se observar que nenhuma paciente sente dor intensa com nível 0 da doença e nos níveis 3 e 4 percebe-se uma concentração de pacientes que sentem dor intensa ao menstruar.

b)

Desconsidere o valor. Talvez considerando a média desses valores.

c)

Grupo	Gestações	Contagem
Controle	37	15
Doente	72	35

13)

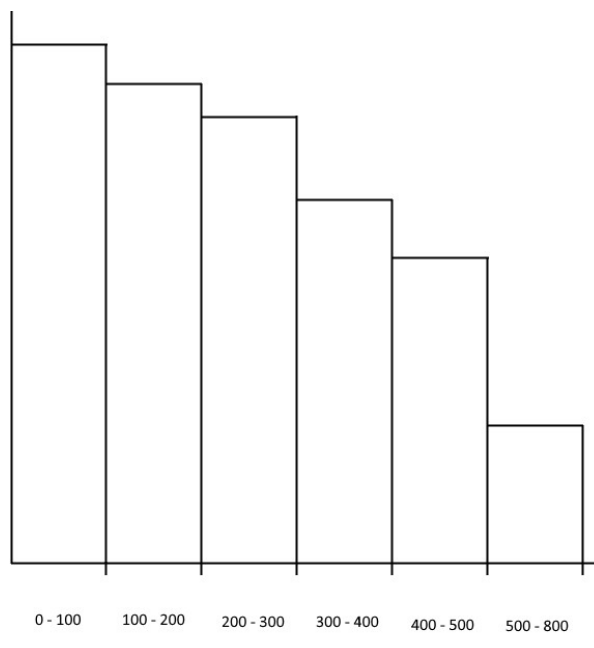
a)

Diversos pacientes são discrepantes em PCR, no pico, poucos são, e no restante, menos que no PCR.

b)

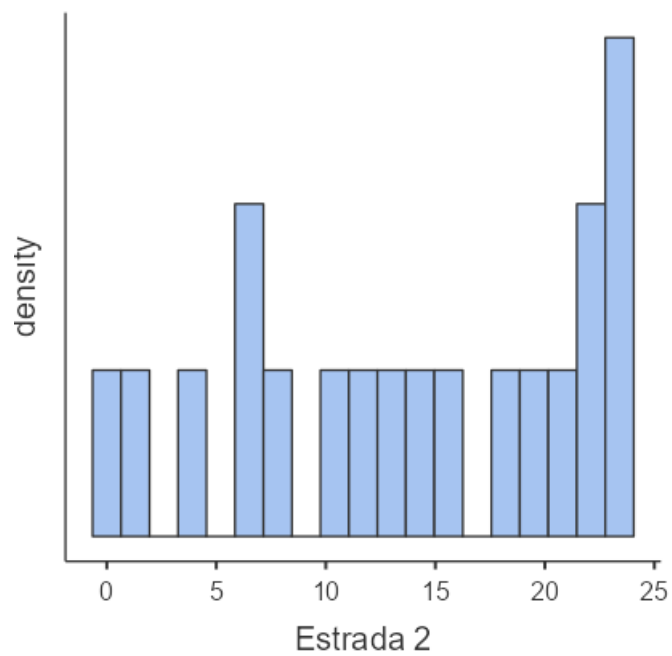
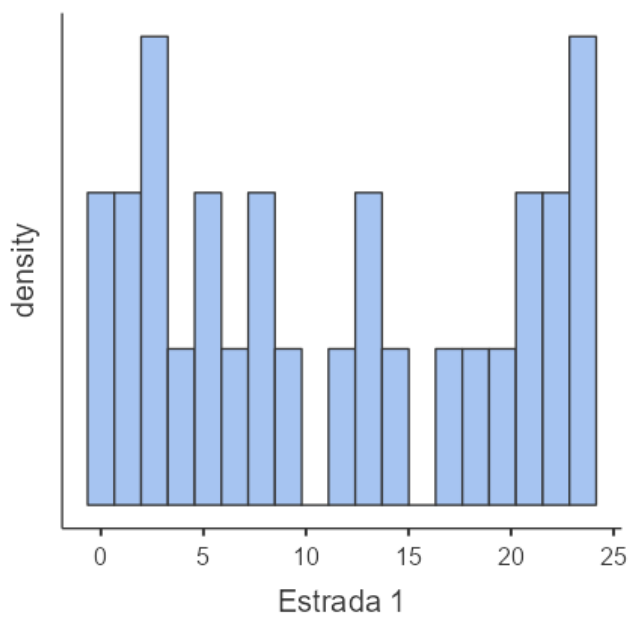
NYHA	Frequência
0	40
1	27
2	23
3	24
4	14

14)
a)

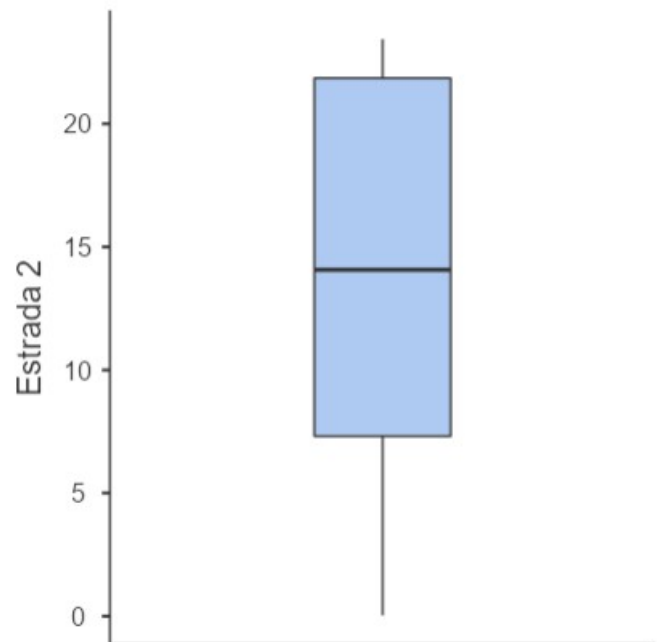
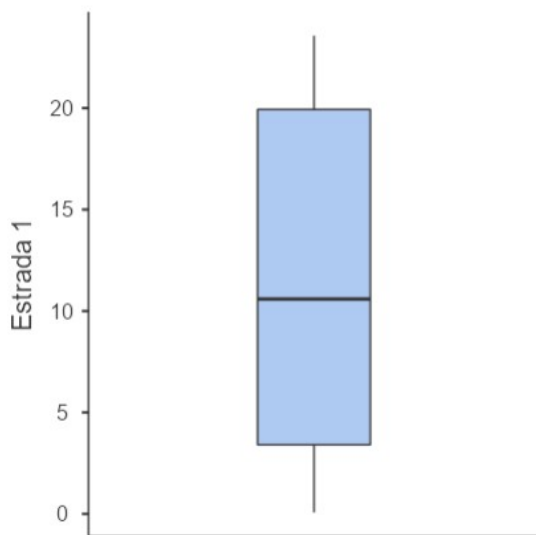


b)
0-100, 100-200, 200-300, 300-400, 400-500.

15)
a) Estradas em sistema de 24 horas.



b)



É sensato concluir que a estrada 2 trabalha com horários mais tardes, enquanto a estrada 1 trabalha mais cedo.

c)
Parece haver intervalos maiores durante a manhã em relação à noite na estrada 2, enquanto nota-se o contrário em estrada 1.

16)

a)

Mediana = 5

Média = 5.12

Q1 = 4

Q2 = 5

Q3 = 6

b)

A variância dos aprovados é de 1.05.

A dos reprovados é igual a 0.194.

A diferença é de 0.856.

17)

a)

Incorreto, 25% e 50%, na verdade.

b)

Incorreto, é a mesma proporção.

c)

Incorreto, o grupo mais ativo fisicamente possui uma média cardíaca superior.

d)

Incorreto, pouco menos que a metade.

18)

b) B e D

19)

a)

Incorreto, possuem a mesma média (14), mas variâncias diferentes (11.2 e 8)

b)

Incorreto, a diferença entre as variâncias é de 3.2.

c)

Incorreto, a média é a mesma para R1 e R2.

d)

Correto.

20)

a)

75% de chance de ganhar menos que 5000 unidades. Bem mais provável de acontecer que o contrário.

b)

Na companhia B, a chance de receber um salário igual ou maior na empresa A é baixíssima, onde na B há uma quase garantia de 7000 unidades de salário.

21)

a)

Incorreto, a distância interquartil é 10.

b)

Incorreto, apenas valores a partir de 35

c)

Incorreto, a mediana se manteria inalterada.

d)

Não há como afirmar.

22)

a)

Incorreto, há uma diferença no desvio-padrão

b)

Incorreto, a probabilidade é maior com o medicamento B.

c)

Correto

23)

Todos incorretos, a média é 2.4, a mediana 2 e a moda 2.

24)

d) 73,3%.

25)

b) A mediana está entre 10 e 15.

26)

c) A média estimada é 12,8.

27)

d) Ambos os estudos sugerem que a distribuição do colesterol na população é simétrica.

28)

a)

$$12 \pm 3 \cdot 4 / 100$$

Erro padrão: 0,04

Média: 11,88 - 12,12

b)

$$M \pm Z \cdot DP / N^{1/2}$$

$$12 \pm 3 \cdot 4 / 2$$

Erro padrão: 2

A média está compreendida no intervalo entre 18 e 6.

c)

$$M \pm Z \cdot DP / N^{1/2}$$

Com N = 9:

$$12 \pm 3 \cdot 4 / 3$$

Erro padrão: 1,33

A média está compreendida no intervalo entre 16 e 8

Com N = 100:

$$12 \pm 3 \cdot 4 / 10$$

Erro padrão: 0,4

A média está compreendida no intervalo entre 13,2 e 10,8

Quanto maior o N da amostra, menor o erro padrão, concluindo numa média menos abrangente (mais precisa).