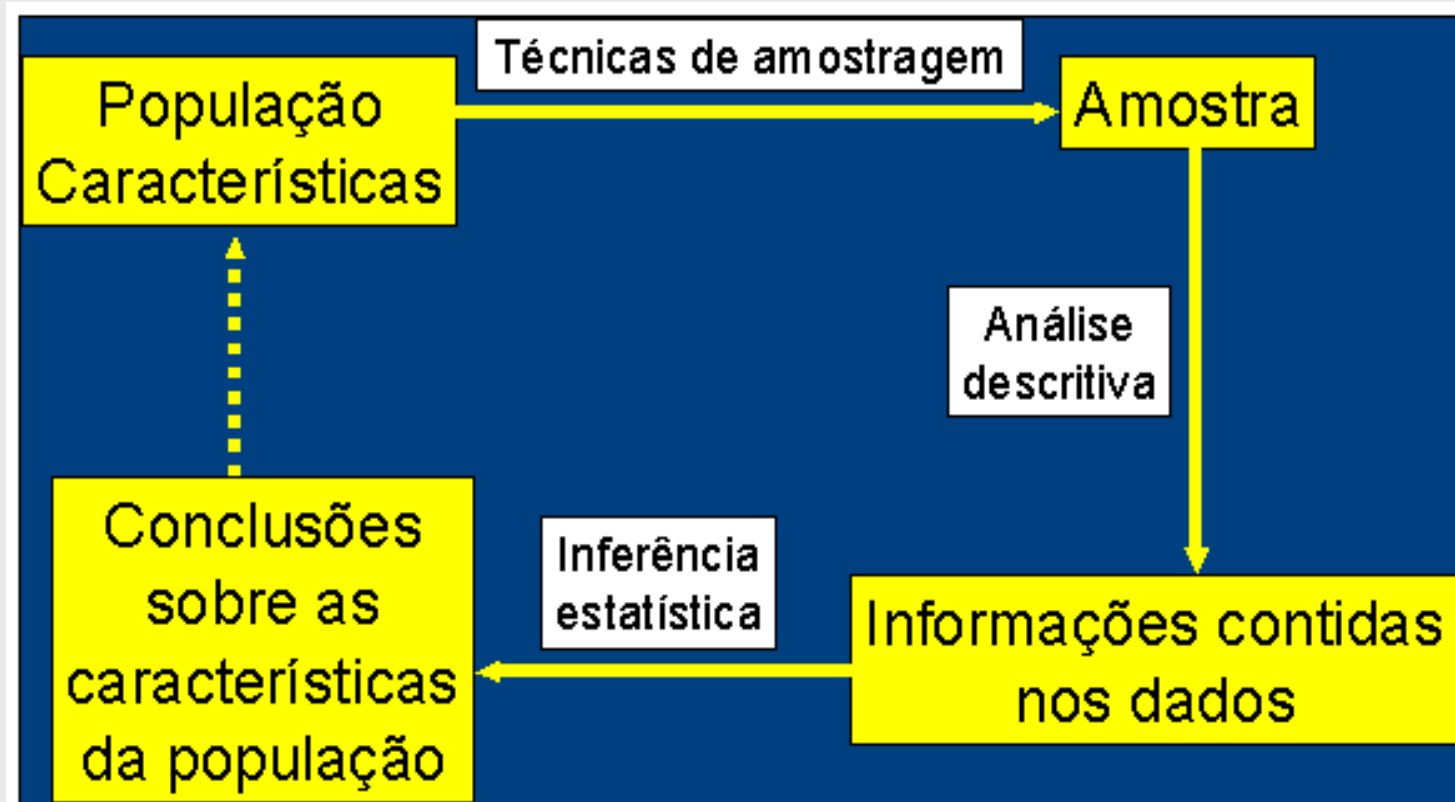


MQA2022

Medidas de Posição e Box-Plot

Professora Ana Amélia Benedito Silva
aamelia@usp.br

Etapas da Análise Estatística



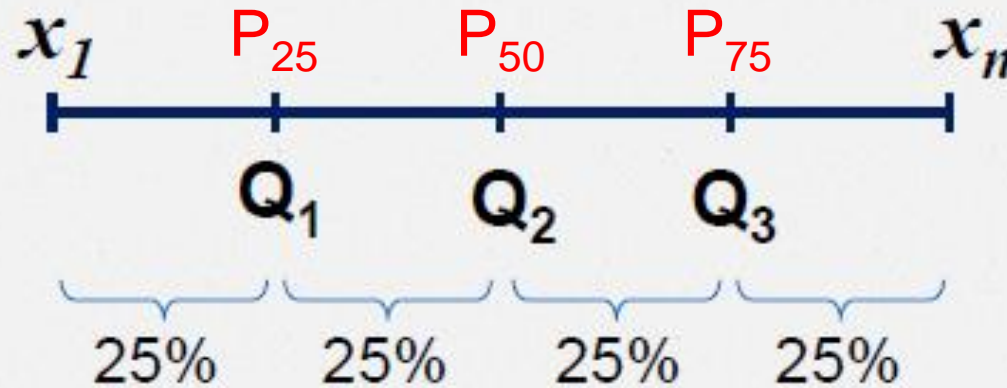
ANÁLISE DESCRITIVA

- tabelas
- gráficos
- medidas
 - média, mediana, moda
 - variância, desvio-padrão, coeficiente de variação
 - percentil, decil, quartil

MEDIDAS SEPARATRIZES

- medidas que dividem uma distribuição **ordenada** de dados em partes iguais
- servem para descrever **posições** numa distribuição de dados
 - Percentil ($P_1, P_2, P_3, \dots, P_{100}$)
 - Decil ($D_1, D_2, D_3, \dots, D_{10}$)
 - Quartil (Q_1, Q_2, Q_3)

Percentil



Os percentis dividem um conjunto de dados ordenados em 100 partes iguais

Percentiles

Percentile: the value below which a percentage of data falls.

Example: You are the fourth tallest person in a group of 20

80% of people are shorter than you:

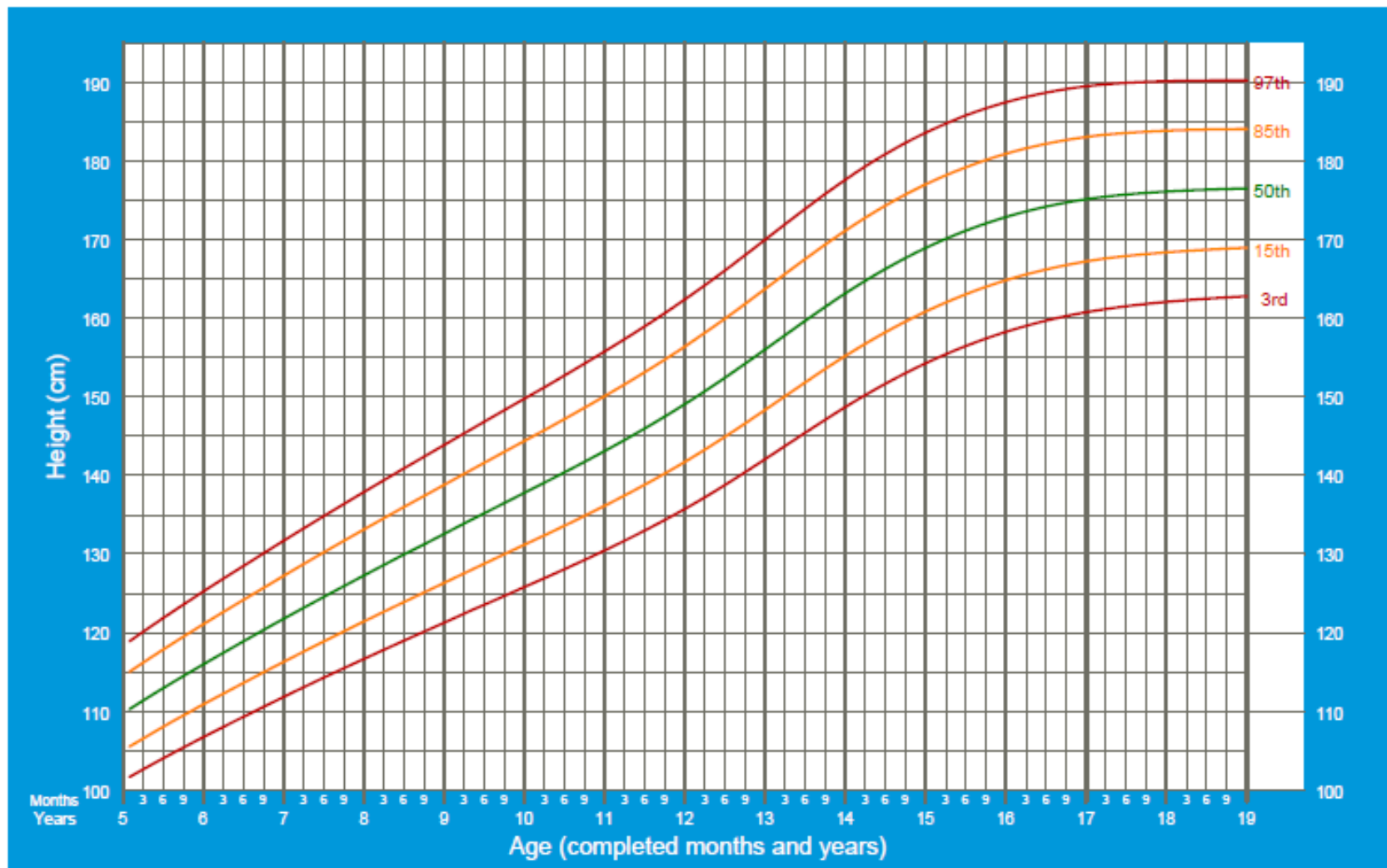


That means you are at the **80th percentile**.

If your height is 1,85m then "1,85m" is the 80th percentile height in that group.

Height-for-age BOYS

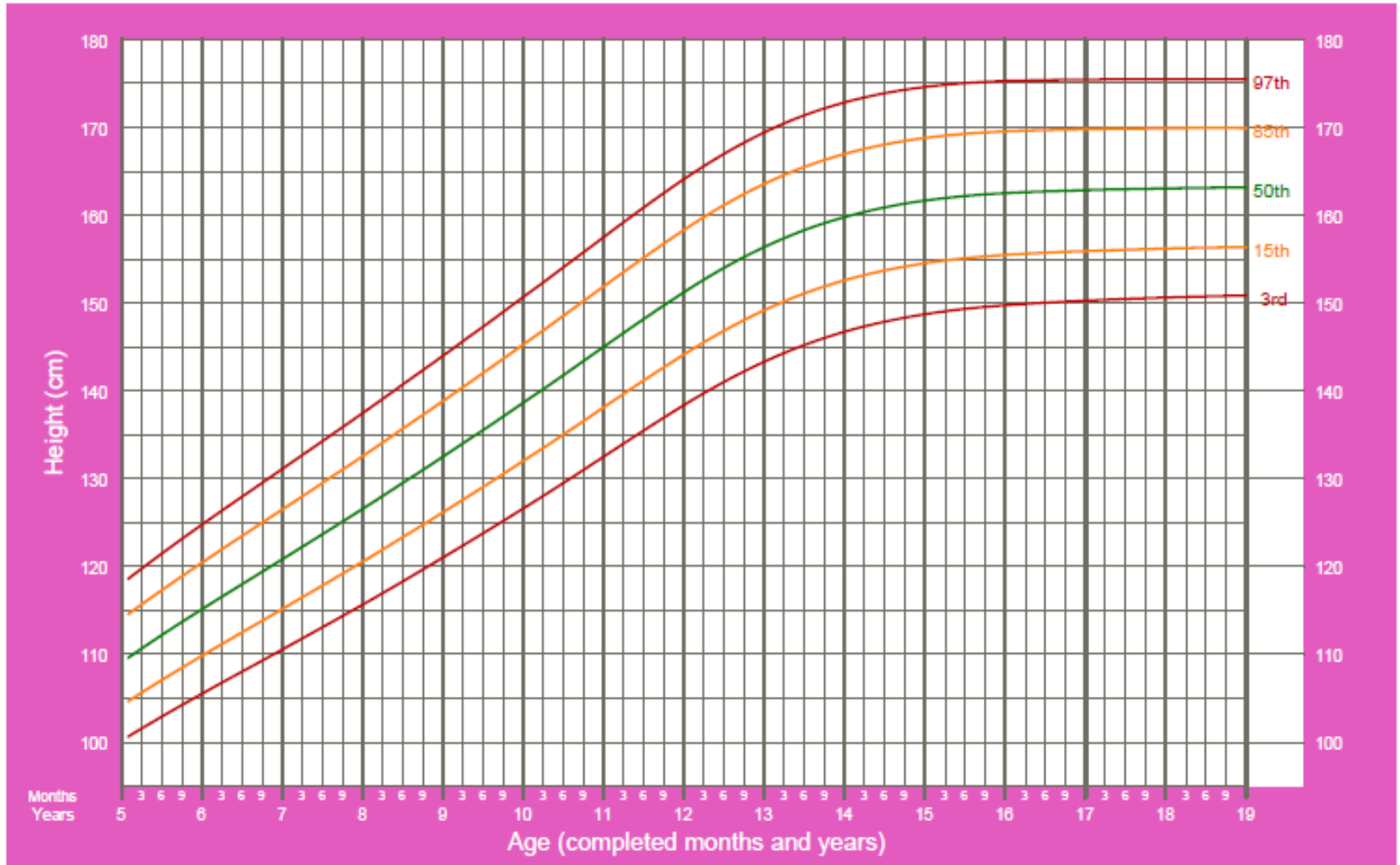
5 to 19 years (percentiles)



2007 WHO Reference

Height-for-age GIRLS

5 to 19 years (percentiles)



2007 WHO Reference

Deciles

Deciles are similar to Percentiles (sounds like decimal and percentile together), as they split the data into **10% groups**:

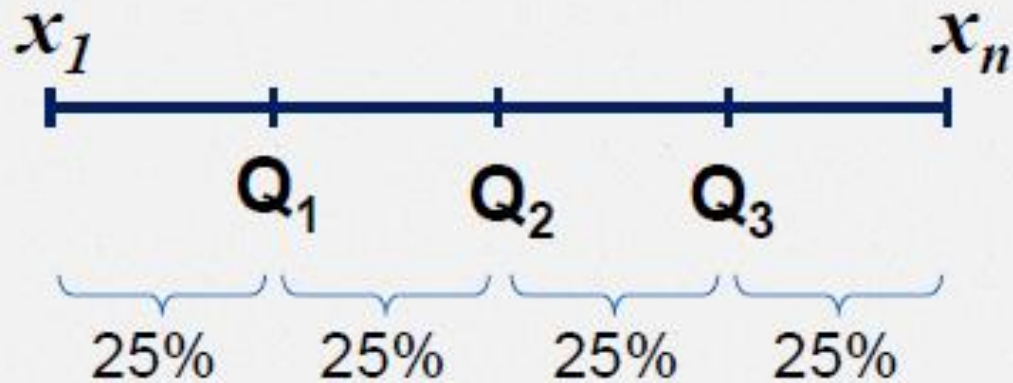
- The **1st decile** is the **10th percentile** (the value that divides the data so that **10%** is below it)
- The **2nd decile** is the **20th percentile** (the value that divides the data so that **20%** is below it)
- etc!

Example: (continued)



You are at the **8th decile** (the 80th percentile).

Quartil



- Os quartis dividem a distribuição em 4 partes iguais.

Quartis

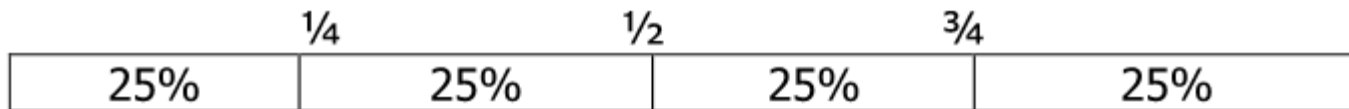
	$\frac{1}{4}$		$\frac{1}{2}$		$\frac{3}{4}$	
25%		25%		25%		25%

Q1: deixa abaixo 25% das observações

25%		75%
-----	--	-----

$$\text{Posição do } Q_1 \text{ (1}^\circ \text{ quartil)} = \frac{1}{4} (n+1)$$

Quartis



Q2: deixa abaixo 50% das observações



$$\text{Posição do } Q_2 \text{ (2}^\circ \text{ quartil)} = \frac{1}{2} (n+1)$$

Quartis

	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	
25%	25%	25%	25%	

Q3: deixa abaixo 75% das observações

75%	25%
-----	-----

$$\text{Posição do } Q_3 \text{ (3}^\circ \text{ quartil)} = \frac{3}{4} \cdot (n+1)$$

Exemplo

Quartis: posição e valor

9 10 17 29 31 53 55 122 129 140 336

→ n=11

Onde ficam os quartis ?

Qual o valor dos quartis ?

Exemplo Cálculo dos quartis

9 10 17 29 31 53 55 122 129 140 336 → n=11

Posição do Q1 = $\frac{1}{4} (11+1) = 3^{\text{a}}$ posição

Q1 = 17

Posição do Q2 = $\frac{1}{2} (11+1) = 6^{\text{a}}$ posição

Q2 = 53

Posição do Q3 = $\frac{3}{4} (11+1) = 9^{\text{a}}$ posição

Q3 = 129

Exemplo

Cálculo dos quartis

10 19 20 21 25 30 31 33 37 61 77 88 91 → $n=13$

Onde ficam os quartis?
Qual os seus valores?

Exemplo Cálculo dos quartis

10 19 20 21 25 30 31 33 37 61 77 88 91 → n=13

Posição do Q1 = $\frac{1}{4} (13+1) = 3,5^{\text{a}}$ posição

$$Q1 = 20 + 0,5*(21-20)$$

$$Q1 = 20,5$$

Posição do Q2 = $\frac{1}{2} (13+1) = 7^{\text{a}}$ posição

$$Q2 = 31$$

Posição do Q3 = $\frac{3}{4} (13+1) = 10,5^{\text{a}}$ posição

$$Q3 = 61 + 0,5*(77-61)$$

$$Q3 = 69$$

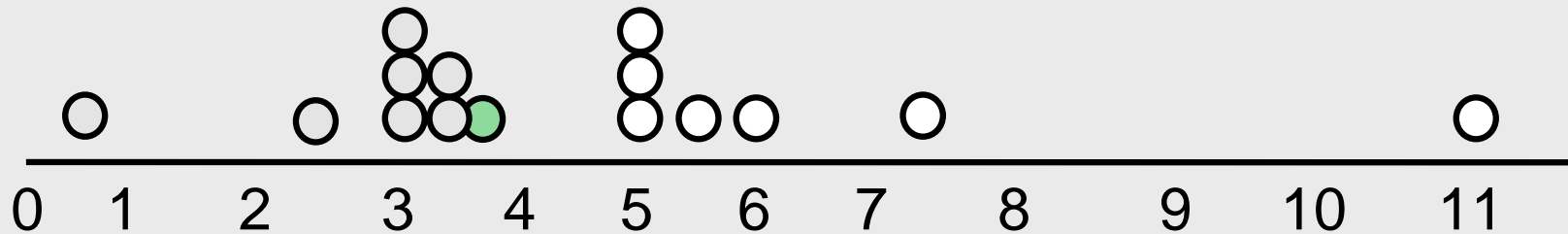
Exercício: Os dados abaixo referem-se aos salários (milhares de R\$/mês) de 15 pessoas com curso superior. Obtenha os quartis deste conjunto de dados.

11	2,5	5,0	5,0	5,5	3,0	3,5	3,0	0,4	3,2	5,0	3,0	3,2	7,4	6,0
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

colocando em ordem

0,4	2,5	3,0	3,0	3,0	3,2	3,2	3,5	5,0	5,0	5,0	5,5	6,0	7,4	11
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Gráfico de pontos dos salários das 15 pessoas com curso superior (milhares de R\$/mês):



Posição do Q1 = $\frac{1}{4} (15+1) = 4^{\text{a}}$ posição

$Q_1 = 3000,00$ reais

Posição do Q2 = $\frac{1}{2} (15+1) = 8^{\text{a}}$ posição

$Q_2 = 3500,00$ reais

Posição do Q3 = $\frac{3}{4} (15+1) = 12^{\text{a}}$ posição

$Q_3 = 5500,00$ reais

0,4	2,5	3,0	3,0	3,0	3,2	3,2	3,5	5,0	5,0	5,0	5,5	6,0	7,4	11
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Exercício: Os dados abaixo referem-se ao tempo de permanência em UTI de bebês nascidos prematuros (dias) .

Obtenha os quartis deste conjunto de dados.

Voce diria que há algum outlier?

18	21	21	23	23	25
27	29	30	31	32	32
32	34	35	36	38	41
42	42	43	44	45	46
46	47	48	50	54	56
57	58	60	61	98	116

1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°
18	21	21	23	23	25	27	29	30	31	32	32
13°	14°	15°	16°	17°	18°	19°	20°	21°	22°	23°	24°
32	34	35	36	38	41	42	42	43	44	45	46
25°	26°	27°	28°	29°	30°	31°	32°	33°	34°	35°	36°
46	47	48	50	54	56	57	58	60	61	98	116

$$n = 36$$

Posição do $Q_1 = (n+1)/4 = 37/4 = 9,25^\circ$ elemento

$$Q_1 = 30 + 0,25(31-30) = 30,25 \text{ dias}$$

Posição do $Q_2 = (n+1)/2 = 37/2 = 18,5^\circ$ elemento

$$Q_2 = 41 + 0,5(42-41) = 41,5 \text{ dias}$$

Posição do $Q_3 = (n+1) \cdot (3/4) = 27,75^\circ$ elemento

$$Q_3 = 48 + 0,75(50-48) = 49,5 \text{ dias}$$

Exercício:

Uma pesquisa foi realizada com o objetivo de verificar se o peso ao nascer está associado à ocorrência de óbito em bebês nascidos com problemas respiratórios.

Foram obtidos os pesos de 50 recém-nascidos, todos com problemas respiratórios, sendo que 23 sobreviveram e 27 morreram.

Qual foi a conclusão da pesquisa?

*Peso dos recém
nascidos que
sobreviveram
(n=23)*

1130
1410
1575
1680
1715
1720
1760
1930
2015
2040
2090
2200
2400
2550
2570
2600
2700
2830
2950
3005
3160
3400
3640

*Peso dos recém
nascidos que
morreram
(n=27)*

1030	2275
1050	2440
1100	2500
1175	2560
1185	2730
1225	
1230	
1262	
1295	
1300	
1310	
1500	
1550	
1600	
1720	
1750	
1770	
1820	
1890	
1940	
2200	
2270	

Cálculo dos quartis

Grupo dos Recém-nascidos que sobreviveram

Posição $Q_1 = \frac{1}{4} (23+1) = 6^{\circ}$ elemento $\rightarrow Q_1 = 1720g$

Posição $Q_2 = \frac{1}{2} (23+1) = 12^{\circ}$ elemento $\rightarrow Q_2 = 2200g$

Posição $Q_3 = \frac{3}{4} (23+1) = 18^{\circ}$ elemento $\rightarrow Q_3 = 2830g$

***Peso dos recém
nascidos que
sobreviveram (n=23)***

1130	
1410	
1575	
1680	
1715	
1720	Q1
1760	
1930	
2015	
2040	
2090	
2200	Q2
2400	
2550	
2570	
2600	
2700	
2830	Q3
2950	
3005	
3160	
3400	
3640	

Cálculo dos quartis

Grupo dos Recém-nascidos que morreram

Posição $Q_1 = \frac{1}{4} (27+1) = 7^{\circ}$ elemento $\rightarrow Q_1 = 1230g$

Posição $Q_2 = \frac{1}{2} (27+1) = 14^{\circ}$ elemento $\rightarrow Q_2 = 1600g$

Posição $Q_3 = \frac{3}{4} (27+1) = 21^{\circ}$ elemento $\rightarrow Q_3 = 2200g$

***Peso dos recém
nascidos que
morreram (n=27)***

1030	
1050	
1100	
1175	
1185	
1225	
1230	Q1
1262	
1295	
1300	
1310	
1500	
1550	
1600	Q2
1720	
1750	
1770	
1820	
1890	
1940	
2200	Q3
2270	
2275	
2440	
2500	
2560	
2730	

Bebês que sobreviveram

mínimo = 1130g

$Q_1 = 1720\text{g}$

$Q_2 = 2200\text{g}$

$Q_3 = 2830\text{g}$

máximo = 3640g

Bebês que morreram

mínimo = 1030g

$Q_1 = 1230\text{g}$

$Q_2 = 1600\text{g}$

$Q_3 = 2200\text{g}$

máximo = 2730g

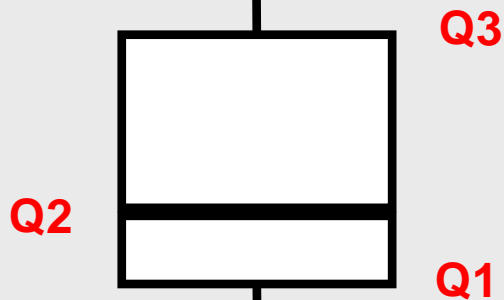
BOX-PLOT

BOX-PLOT

* outlier (pode ter mais que 1)

$$\text{Limite superior} = Q_3 + 1,5 \times (Q_3 - Q_1)$$

maior valor dentro dos limites (novo máximo)



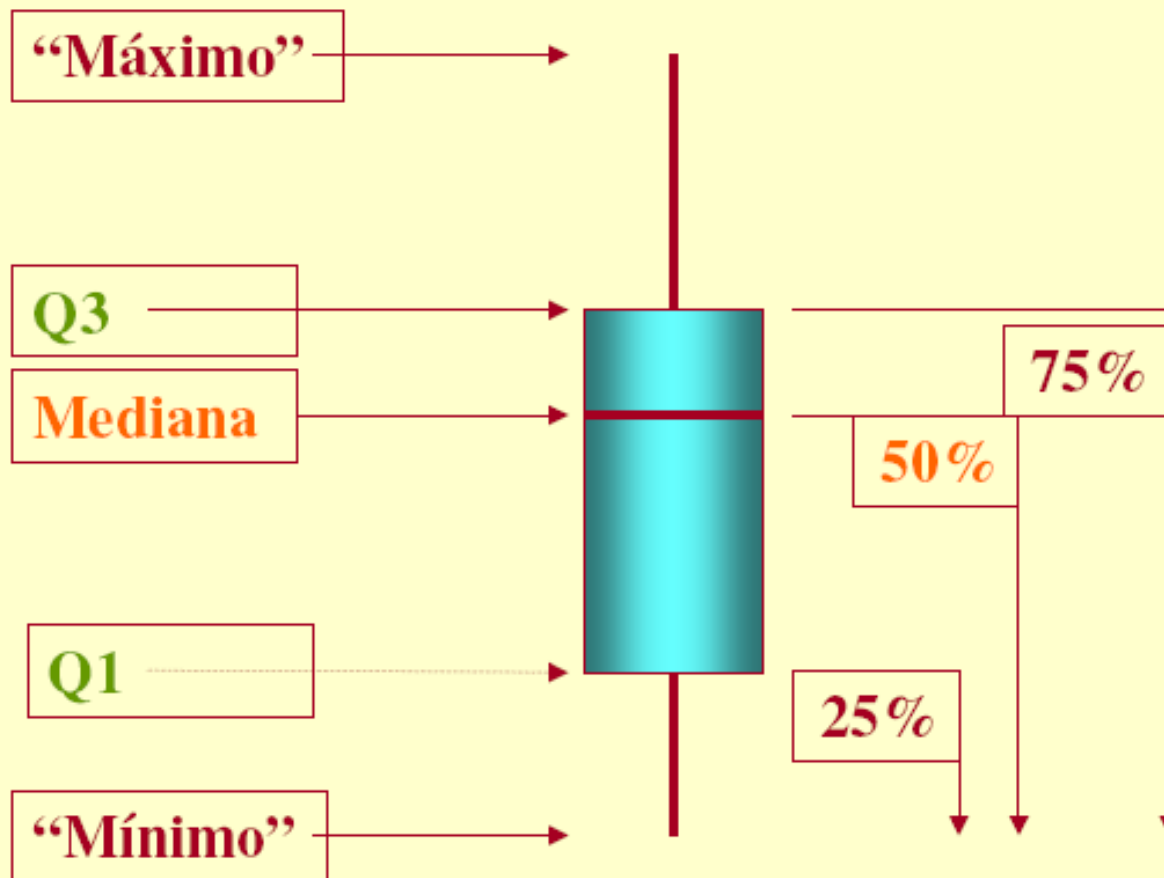
menor valor dentro dos limites (novo mínimo)

$$\text{Limite Inferior} = Q_1 - 1,5 \times (Q_3 - Q_1)$$

* outlier (pode ter mais que 1)

Construção

$$LS = Q3 + 1,5(Q3 - Q1)$$



$$LI = Q1 - 1,5(Q3 - Q1)$$

“Máximo” é o maior valor menor que LS ;

“Mínimo” é o menor valor maior que LI .

Exercício: Os dados abaixo referem-se ao tempo de permanência em UTI de bebês nascidos prematuros (dias) .

Obtenha os quartis deste conjunto de dados e construa um box-plot.

18	21	21	23	23	25
27	29	30	31	32	32
32	34	35	36	38	41
42	42	43	44	45	46
46	47	48	50	54	56
57	58	60	61	98	116

Q1= 30,25 dias

Q2= 41,5 dias

Q3= 49,5 dias

Box-plot

Q1= = 30,25 dias

Q2= 41,5 dias

Q3= 49,5 dias

Limite superior= $Q_3 + 1,5(Q_3 - Q_1)$

Limite superior = $49,5 + 1,5.(49,5-30,25)$

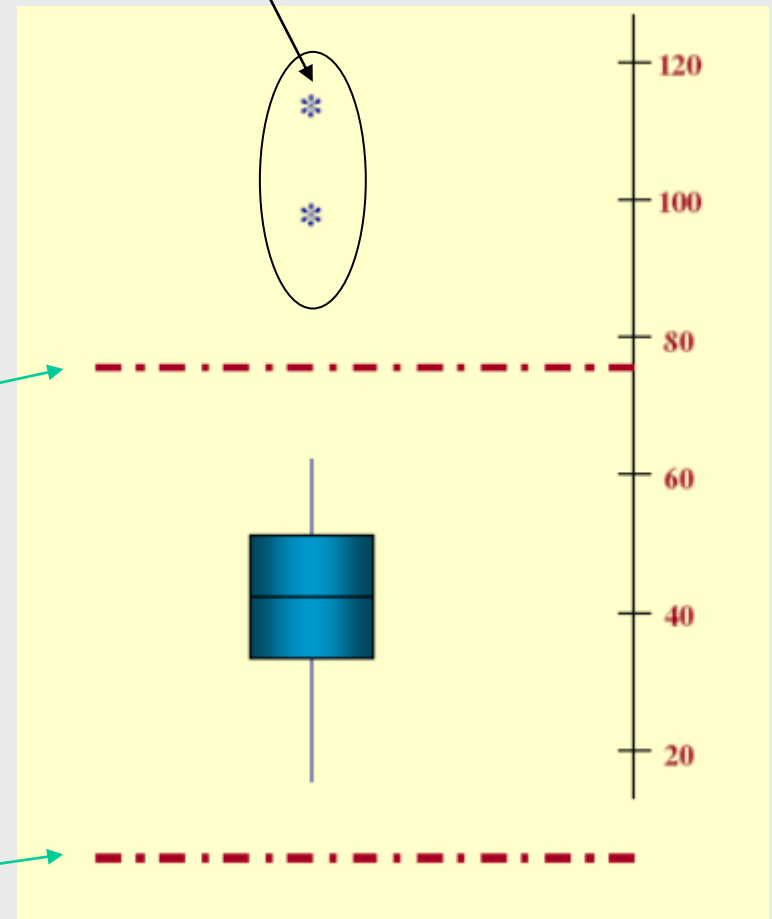
Limite superior = 78,375 dias

Limite Inferior= $Q_1 - 1,5(Q_3 - Q_1)$

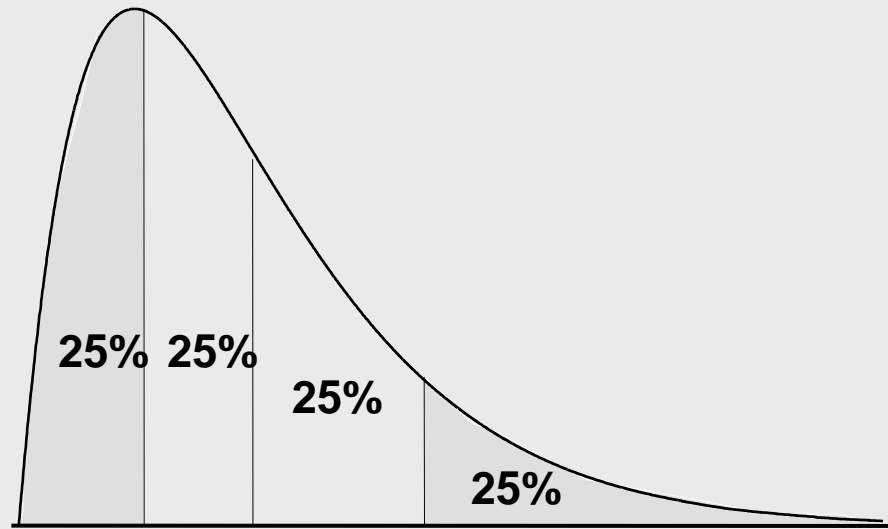
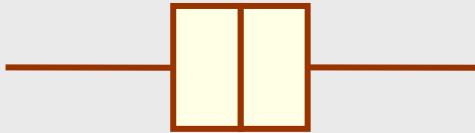
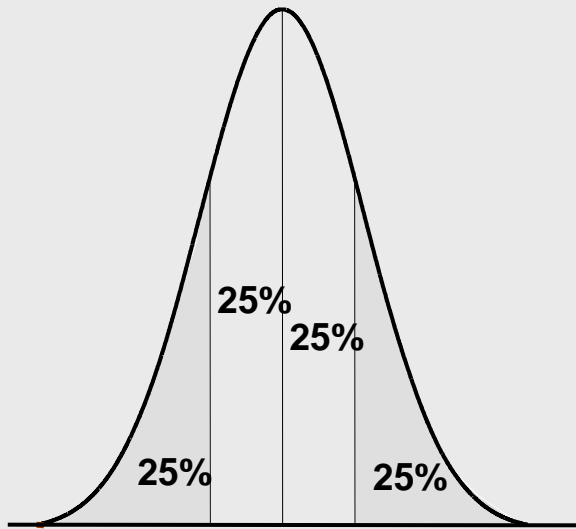
Limite inferior = $30,25 - 1,5.(49,5-30,25)$

Limite inferior = 1,375 dias

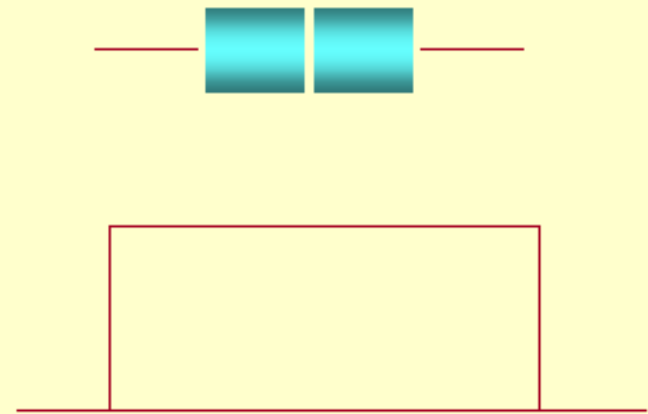
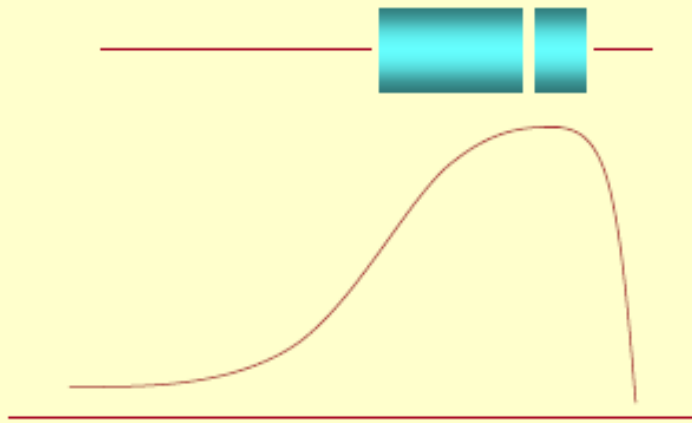
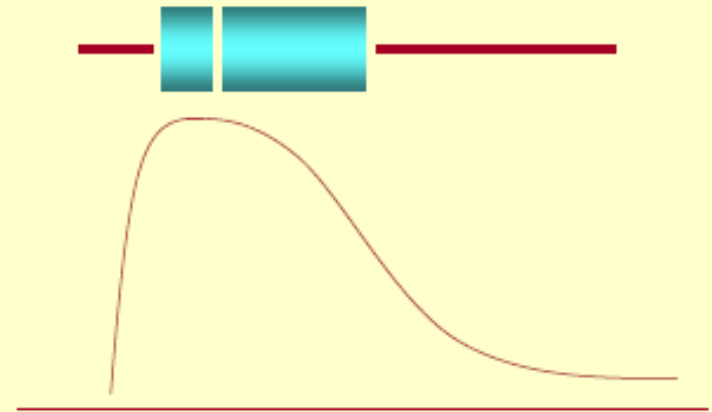
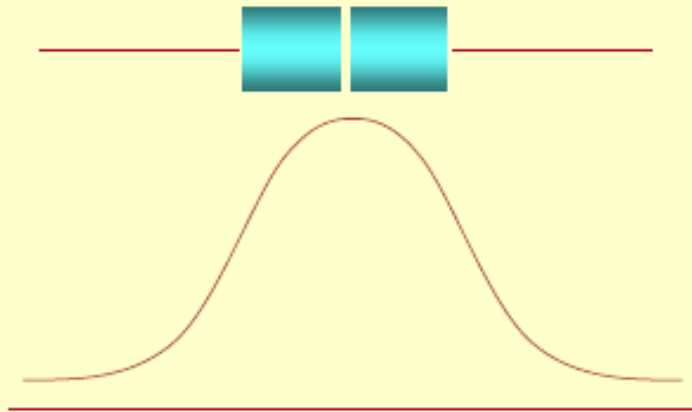
Conclusão: há 2 outliers.



BOX-PLOT x HISTOGRAMA

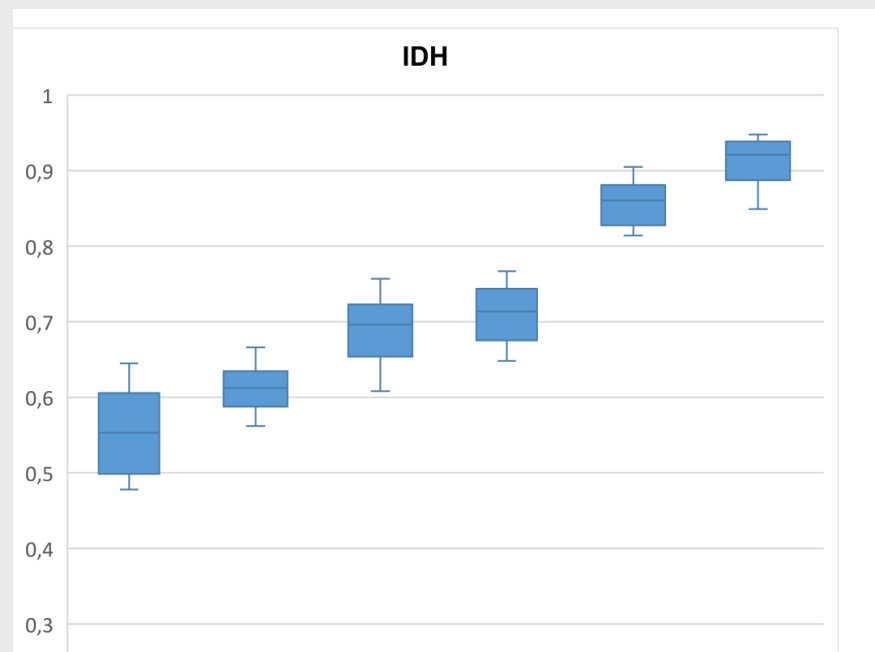


Forma da Distribuição



COUNTRY	MÉDIA	VALOR MIN	QUARTIL 1	MEDIANA	QUARTIL 3	VALOR MAX
Guatemala	0,559	0,478	0,514	0,562	0,605	0,645
Kyrgyzstan	0,611	0,562	0,589	0,612	0,633	0,666
Brazil	0,691	0,608	0,662	0,698	0,718	0,757
Mexico	0,71	0,648	0,678	0,713	0,742	0,767
Japan	0,861	0,814	0,842	0,865	0,882	0,905
Norway	0,913	0,849	0,889	0,921	0,938	0,948

IDH – Índice de Desenvolvimento Humano

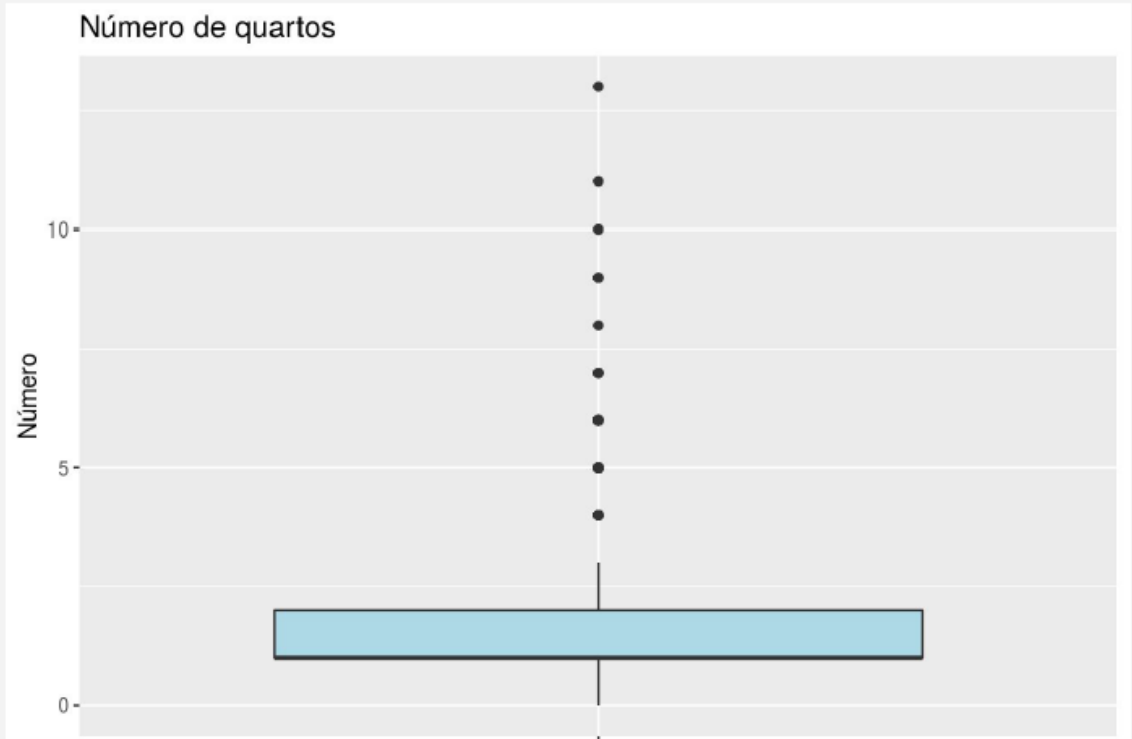


Arbnb

Número de quartos

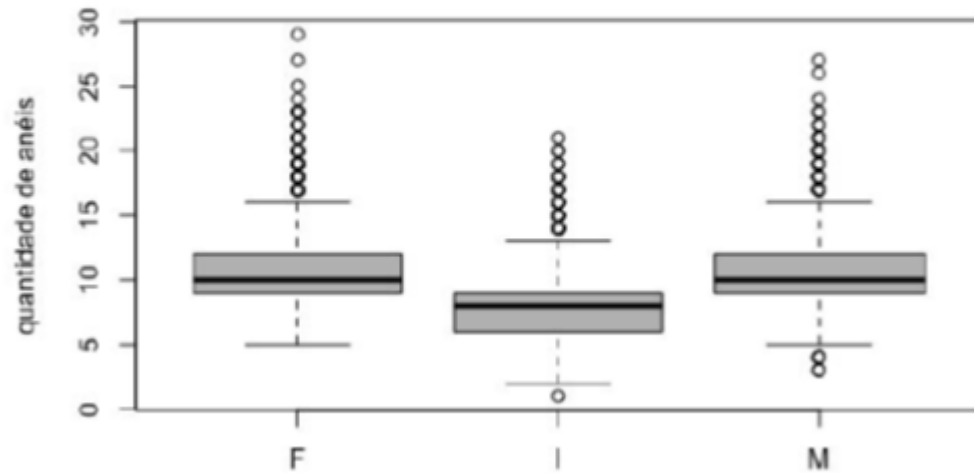
MÍNIMO	0,00
Q1	1,00
MEDIANA	1,00
MÉDIA	1,51
Q3	2,00
MÁXIMO	13,00
VARIÂNCIA	0,91
D.P.	0,95

No de imóveis = 17979





Abalones - Variável Rings



Exercício

As rendas mensais em reais dos 25 ingressantes num certo curso de pós-graduação em finanças de uma universidade foram as seguintes:

2200, 2200, 2200, 2200, 2300, 2300, 2400, 2400, 2400,
2400, 2500, 2500, 2600, 2600, 2600, 2700, 2700, 2800,
3500, 3600, 3600, 4000, 4200, 4400, 5400.

Construa o “boxplot”.

Etapas

1. Colocar em ordem
2. Obter quartis
3. Obter mínimo e máximo
4. Obter limite inferior = $Q_1 - 1,5(Q_3 - Q_1)$
5. Obter limite superior = $Q_3 + 1,5(Q_3 - Q_1)$
6. Fazer o box-plot

Exercício

As rendas mensais em reais dos 25 ingressantes num certo curso de pós-graduação em finanças de uma universidade foram as seguintes:

2200, 2200, 2200, 2200, 2300, 2300, 2400, 2400, 2400, 2500, 2500, 2600, 2600, 2700, 2700, 2800, 3500, 3600, 3600, 4000, 4200, 4400, 5400.

Solução

$$n = 25$$

$$Q1 = 2300 + 0,5 * (2400 - 2300) = 2350 \text{ reais (posição } 6,5^{\circ}\text{)}$$

$$Q2 = 2600 \text{ reais (posição } 13^{\circ}\text{)}$$

$$Q3 = 3500 + 0,5 * (3600 - 3500) = 3550 \text{ reais (posição } 19,5^{\circ}\text{)}$$

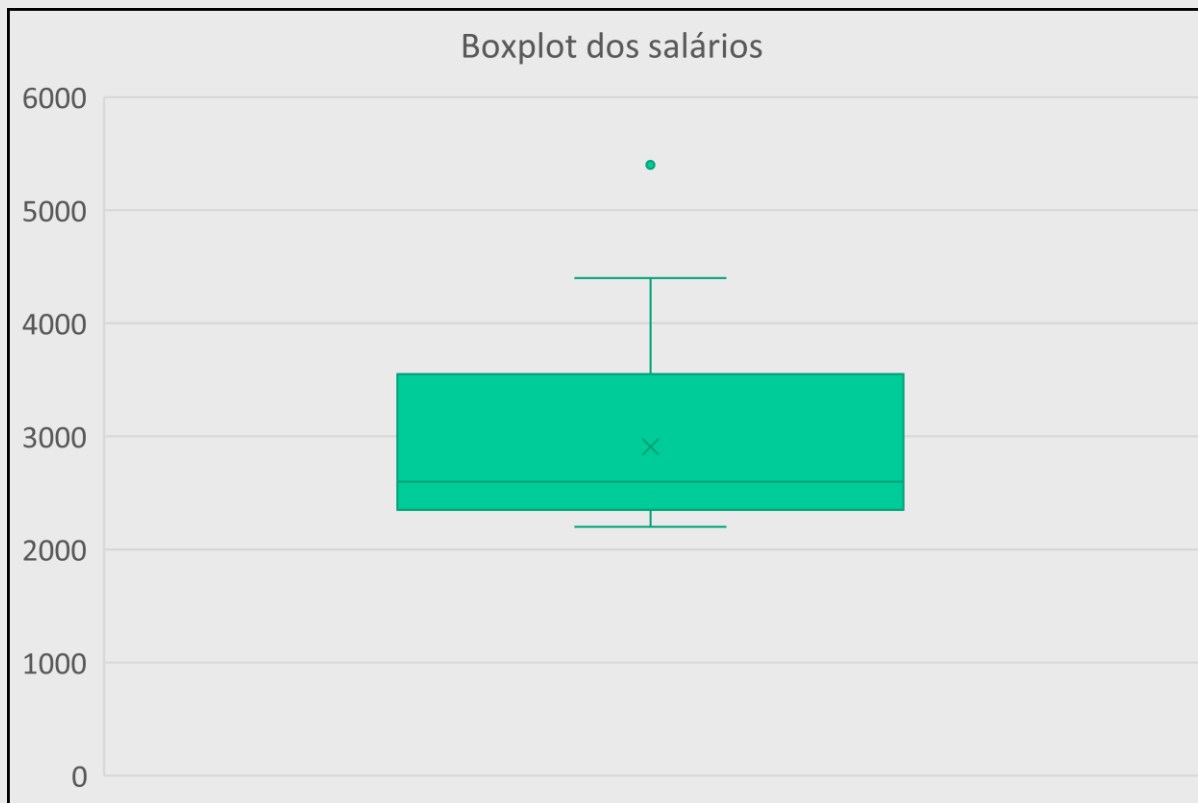
Mínimo = 2200 reais; Máximo = 5400 reais

$$\text{Limite inferior} = 2350 - 1,5 * (3550 - 2350) = 550 \text{ reais}$$

$$\text{Limite superior} = 3550 + 1,5 * (3550 - 2350) = 5350 \text{ reais}$$

Logo há um outlier -> 5400 reais

Boxplot dos salários



OBRIGADA!!