

Universidade de São Paulo

FFCLRP: Introdução a Estatística

Aula – Separatrizes

Prof_a: Fernando FF

Ribeirão Preto

Maio - 2020

MEDIDAS SEPARATRIZES

São números que dividem a sequência ordenada de dados em partes que contém a mesma quantidade de elementos da serie.

MEDIDAS SEPARATRIZES

- Para entender bem uma distribuição, você pode conhecer valores acima ou abaixo dos quais se encontra uma determinada porcentagem dos dados através das medidas separatrizes.
- Separatriz de uma série de n elementos colocados em ordem crescente, é valor que divide em duas partes quaisquer.
- As principais separatrizes são: mediana, quartis, decis e percentis.

MEDIDAS SEPARATRIZES

- A mediana divide uma sequência **ordenada** em dois grupos, cada um deles contendo 50% dos valores da sequência.
- Os quartis dividem a sequência **ordenada** em quatro grupos contendo 25% dos dados da sequência.
- Os quintis dividem a sequência **ordenada** em cinco grupos contendo 20% dos dados da sequência
- Os decis dividem a sequência **ordenada** em dez partes iguais. Cada qual contendo 10% dos dados da sequência
- Os percentis dividem a sequência **ordenada** em cem partes iguais

Quartis

Se dividirmos a série ordenada em quatro partes, cada uma ficará com seus 25% de seus elementos. Os elementos que separam estes grupos são chamados de **quartis**.

Assim, o primeiro quartil, que indicaremos por Q_1 , separa a sequência ordenada deixando 25% de seus valores à esquerda e 75% de seus valores à direita. Q_3 é o contrário.

O segundo quartil, que indicaremos por Q_2 , separa a sequência ordenada deixando 50% de seus valores à esquerda e 50% de seus valores à direita. É a **mediana**

Quintis

Se dividirmos a série ordenada em cinco partes, cada uma ficara com seus 20% de seus elementos. Os elementos que separam estes grupos são chamados de **quintis**.

Assim, o primeiro quintil, que indicaremos por K_1 , separa a sequência ordenada deixando 20% de seus valores à esquerda e 80% de seus valores a direita.

De modo análogo são definidos os outros quintis.

Quintis

Se dividirmos a série ordenada em dez partes, cada uma ficara com seus 10% de seus elementos. Os elementos que separam estes grupos são chamados de **decis**.

Assim, o primeiro quintil, que indicaremos por D1, separa a sequência ordenada deixando 10% de seus valores à esquerda e 90% de seus valores a direita.

De modo análogo são definidos os outros decis.

Percentis

Se dividirmos a série ordenada em cem partes, cada uma ficara com seus 1% de seus elementos. Os elementos que separam estes grupos são chamados de **percentis ou centis**.

Assim, o primeiro quintil, que indicaremos por P1, separa a sequência ordenada deixando 1% de seus valores à esquerda e 99% de seus valores a direita.

De modo análogo são definidos os outros percentis.

P1	P2 A P100
1%	99%

Percentis

Note que:

QUARTIS- PERCENTIS

$$Q_1 = P_{25}$$

$$Q_2 = P_{50}$$

$$Q_3 = P_{75}$$

QUINTIS-PERCENTIS

$$K_1 = P_{20}$$

$$K_2 = P_{40}$$

$$K_3 = P_{60}$$

$$K_4 = P_{80}$$

DECIS - PERCENTIS

$$D_1 = P_{10}$$

$$D_2 = P_{20}$$

$$D_3 = P_{30}$$

$$D_4 = P_{40}$$

$$D_5 = P_{50}$$

$$D_6 = P_{60}$$

$$D_7 = P_{70}$$

$$D_8 = P_{80}$$

$$D_9 = P_{90}$$

Dados Não Agrupados

Identificamos a medida que queremos obter com o percentil correspondente, P_i . Calculamos $i\%$ de n para localizar a posição do percentil i no Rol, ou seja:

$$\frac{i \times n}{100}$$

Em seguida, identificamos o elemento que ocupa esta posição. Note que se o elemento for um número inteiro, então P_i que estamos procurando identificar é um dos elementos da sequência ordenada. Se não for um número inteiro, isto significa que P_i é um elemento intermediário entre os elementos que ocupam as posições aproximadas por falta ou por excesso do valor calculado. Neste caso, P_i é definido como sendo a média dos valores que ocupam estas posições aproximadas.

Dados Agrupados sem Intervalos de Classe

Identificamos a medida que queremos obter com o percentil correspondente, P_i .

Calculamos $i\%$ de $n(\sum f_j)$ para localizar a posição do percentil i no Rol, ou seja:

$$\frac{i \times \sum_j f_j}{100}$$

Exemplo: Calcule o D_4 para a série

x_i	f_i	F_i
2	3	3
4	5	8
5	8	16
7	6	22
10	2	24
	$\Sigma f_i = 24$	

Solução

- Montando a tabela com Frequência Acumulada, tem-se :

x	f_i	F_i
2	3	3
4	5	8
5	8	16
7	6	22
10	2	24
Total	24	

D4 = QUARTO DECIL EQUIVALE A P40 (4X10)
D4= P40

USANDO
$$\frac{i \times \sum_j f_j}{100} = \frac{40 \times 24}{100} = 9,6$$

D4=5

Solução

$$\frac{i \times \sum_j f_j}{100} = \frac{40 \times 24}{100} = 9,6$$

- Montando a tabela com Frequência Acumulada, tem-se :

x	f _i	F _i
2	3	3
4	5	8
5	8	16
7	6	22
10	2	24
Total	24	

Interpretacao: 40% dos valores desta sequênciã são valores menores ou iguais a 5 e 60% dos valores desta sequênciã são valores Maiores ou iguais a 5.

Dados Agrupados sem Intervalos de Classe

Exemplo: Calcule o Q3, K2, D7 e P88 para a série

Tabela 1 – Valores de aptidão para os indivíduos da população.

x	f_i
50	6
100	14
200	20
270	10
Total	50

Solução

$$\frac{i \times \sum_j f_j}{100}$$

- Montando a tabela com Frequência Acumulada, tem-se :

x	f_i	F_i
50	6	6
100	14	20
200	20	40
270	10	50
Total	50	50

$$Q3 = P75$$

$$75 \times 50 / 100 = 37,5$$

$$Q3 = 200$$

$$K2 = P40$$

$$40 \times 50 / 100 = 20$$

$$K2 = 100$$

$$D7 = P70$$

$$70 \times 50 / 100 = 35$$

$$D7 = 100$$

$$P88$$

$$88 \times 50 / 100 = 44$$

$$P88 = 270$$

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D8 para a série

classe	f_i
0 ----- 200	6
200 ----- 400	14
400 ----- 600	20
600 ----- 800	10
Total	50

Dados Agrupados com Intervalos de Classe

Para obtermos a fórmula geral para o cálculo dos percentis, vamos generalizar a fórmula de mediana:

$$P_i = li + \frac{\left[\frac{i \times n}{100} - F(\text{ant}) \right] \times h}{f_i}$$

Sendo:

P_i – Percentil i (1, 2, 3, ..., 99);

li - limite inferior da classe que contém o percentil;

n – número de elementos da serie ($\sum f_i$);

$F(\text{ant})$ – frequência acumulada da classe anterior a classe que contém o percentil;

f_i - frequência simples da classe que contém o percentil;

h - amplitude do intervalo da classe mediana

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, K4, D8

Intervalo de classe	f_i	F_i	
0 ----- 200	6	6	
200 ----- 400	14	20	
400 ----- 600	20	40	
600 ----- 800	10	50	
Total	50		

$$P_i = li + \frac{\left[\frac{i \times n}{100} - F(\text{ant}) \right] \times h}{f_i}$$

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D8

Intervalo de classe	f_i	F_i	
0 ----- 200	6	6	
200 ----- 400	14	20	
400 ----- 600	20	40	
600 ----- 800	10	50	
Total	50		

Vamos achar a posição ou classe de cada quantil a ser calculado

$$\frac{i \times \sum_j f_j}{100}$$

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, K4, D8

$$P_i = li + \frac{\left[\frac{i \times n}{100} - F(\text{ant}) \right] \times h}{f_i}$$

classe	fi	Fi	
0 ----- 200	6	6	
200 ----- 400	14	20	
400 ----- 600	20	40	
600 ----- 800	10	50	
Total	50		

Vamos achar a posição ou classe de cada quantil a ser calculado

$$\frac{i \times \sum_j f_j}{100} = \frac{i \times n}{100} = \frac{i \times 50}{100} = \frac{i}{2}$$

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant})\right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)
0 ----- 200	6	6	0
200 ----- 400	14	20	6
400 ----- 600	20	40	20
600 ----- 800	10	50	40
Total	50		

Vamos achar a posição ou classe de cada quantil a ser calculado

$$Q1 = P25 \quad \longrightarrow \quad 25/2 = 12.5$$

$$K4 = P80 \quad \longrightarrow \quad 80/2 = 40$$

$$D9 = P90 \quad \longrightarrow \quad 90/2 = 45$$

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant}) \right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	K4
600 ----- 800	10	50	40	D9
Total	50			

Vamos achar a posição ou classe de cada quantil a ser calculado

Q1 = P25 \longrightarrow **25/2 = 12.5**

K4 = P80 \longrightarrow **80/2 = 40**

D9 = P90 \longrightarrow **90/2 = 45**

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant}) \right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	K4
600 ----- 800	10	50	40	D9
Total	50			

Vamos achar a posição ou classe de cada quantil a ser calculado

Q1 limite inferior \longrightarrow **li=200** e **fi=14**
K4 limite inferior \longrightarrow **li=400** e **fi=20**
D9 limite inferior \longrightarrow **li=600** e **fi=10**

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant}) \right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	K4
600 ----- 800	10	50	40	D9
Total	50			

Vamos achar a posição ou classe de cada quantil a ser calculado

Q1 = P25=

$$Q1 = P_{25} = li + \frac{\left[\frac{25}{2} - F(\text{ant}) \right] \times h}{f_i}$$

Dados Agrupados com Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant})\right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	K4
600 ----- 800	10	50	40	D9
Total	50			

Vamos achar Q1 a ser calculado

$$Q1 = P_{25} = 200 + \frac{\left[\frac{25}{2} - 6\right] \times 200}{14} \sim 292,85$$

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant})\right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	K4
600 ----- 800	10	50	40	D9
Total	50			

Vamos achar K4 a ser calculado

$$K4 = P_{80} = 400 + \frac{\left[\frac{80}{2} - 20\right] \times 200}{20} = 600$$

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, K4, D9

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant})\right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	K4
600 ----- 800	10	50	40	D9
Total	50			

Vamos achar D9 a ser calculado

$$D9 = P_{90} = 600 + \frac{\left[\frac{90}{2} - 40\right] \times 200}{10} = 700$$

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, Q2 e Q3

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant})\right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	
600 ----- 800	10	50	40	
Total	50			

Vamos achar a posição na tabela acima

Q2 = P50 \longrightarrow $i/2$ corresponde a 25

Q3 = P75 \longrightarrow $i/2$ corresponde a 37,5

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, Q2 e Q3

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant}) \right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	Q2, Q3
600 ----- 800	10	50	40	
Total	50			

Vamos achar a posição na tabela acima

Q2 = P50 \longrightarrow **i/2** corresponde a 25

Q3 = P75 \longrightarrow **i/2** corresponde a 37,5

Dados Agrupados em Intervalos de Classe

Exemplo: Calcule o Q1, Q2 e Q3

$$P_i = li + \frac{\left[\frac{i}{2} - F(\text{ant})\right] \times h}{f_i}$$

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	Q2 ,Q3
600 ----- 800	10	50	40	
Total	50			

Vamos o valor dos quartis usando a fórmula

$$Q2 = P_{50} = 400 + \frac{\left[\frac{50}{2} - 20\right] \times 200}{20} = 450$$

$$Q3 = P_{75} = 400 + \frac{\left[\frac{75}{2} - 20\right] \times 200}{20} = 575$$

BOXPLOT

- Uma aplicação interessante para os quartis é a construção do chamado gráfico de caixa (ou boxplot).

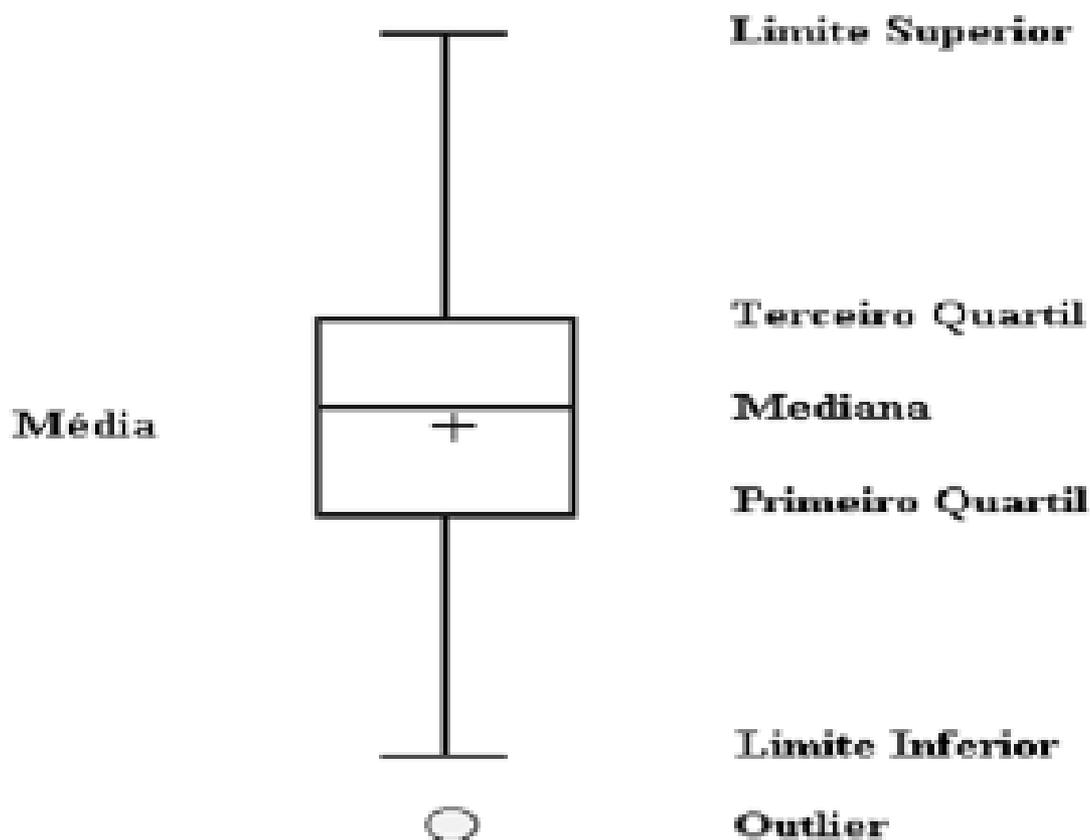


Figura 1: Esboço do gráfico de caixas (BoxPlot).

BOXPLOT

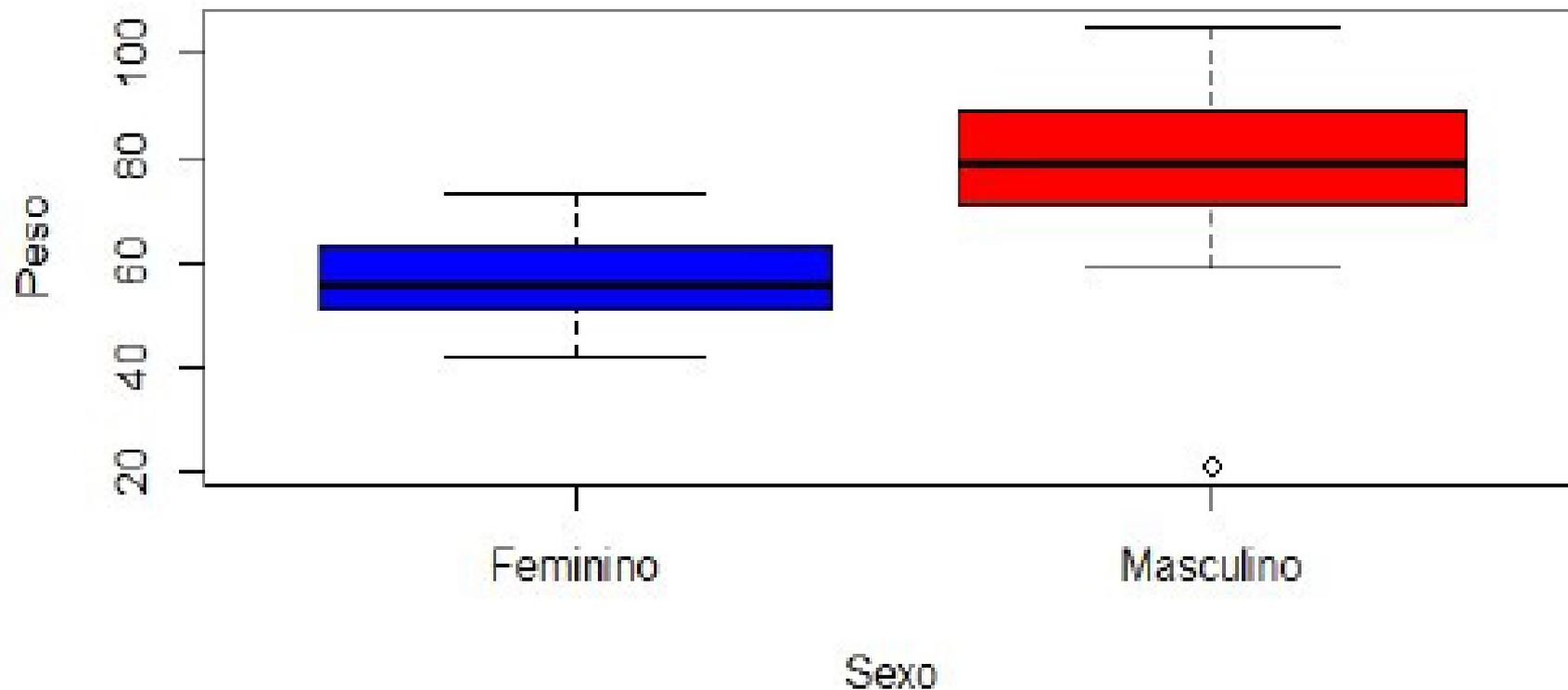


Figura 2: Esboço do BoxPlot dos pesos dos alunos da disciplina Estatística Econômica A de 2017 por sexo.

Faça o boxplot para o exemplo abaixo

Exemplo: dado o Q1, Q2 e Q3

Intervalo de classe	fi	Fi	F(ant)	
0 ----- 200	6	6	0	
200 ----- 400	14	20	6	Q1
400 ----- 600	20	40	20	Q2 ,Q3
600 ----- 800	10	50	40	
Total	50			

$$Q1 = P_{25} = 200 + \frac{\left[\frac{25}{2} - 6\right] \times 200}{14} \sim 292,85$$

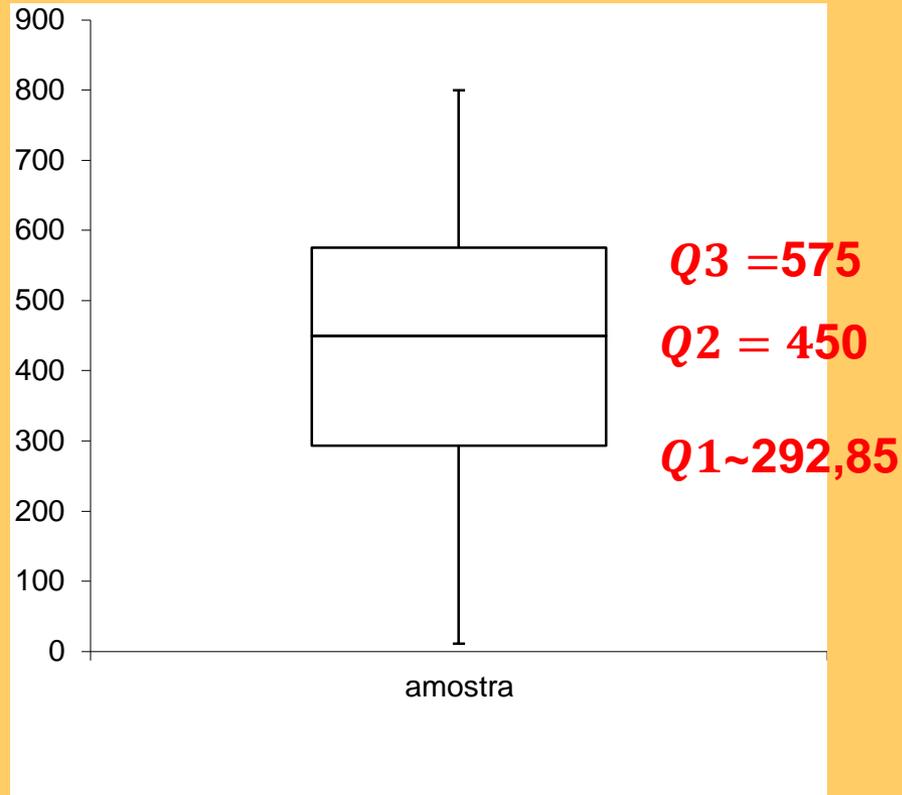
$$Q2 = P_{50} = 400 + \frac{\left[\frac{50}{2} - 20\right] \times 200}{20} = 450$$

$$Q3 = P_{75} = 400 + \frac{\left[\frac{75}{2} - 20\right] \times 200}{20} = 575$$

boxplot

Intervalo de classe

0 |----- 200
200 |----- 400
400 |----- 600
600 |----- 800



Referências

- CRESPO, Antônio Arnot. **Estatística fácil**. Editora Saraiva, 2017.
- VIEIRA, Sonia. Estatística básica. **São Paulo: Cengage Learning**, v. 9, 2012..
- DA CUNHA, Sonia Baptista; CARVAJAL, Santiago Ramírez. **Estatística Básica-a Arte de Trabalhar com Dados**. Elsevier Brasil, 2009.