

Distribuições de frequência e seus gráficos

Otaviano Francisco Neves

Distribuições de frequência

Minutos gastos ao telefone

102	124	108	86	103	82
71	104	112	118	87	95
103	116	85	122	87	100
105	97	107	67	78	125
109	99	105	99	101	92

Valores-chave:

Valor mínimo = 67

Valor máximo = 125

Faça uma tabela de distribuição de frequência com cinco classes.

Passos para construção da tabela de frequência

- ▶ 1. Decida o número de classes, que deve ficar entre 5 e 15.
 - ▶ (Para este problema use 5.)
- ▶ 2. Calcule a amplitude das classes (h)
 - ▶ Primeiro calcule: amplitude total = valor máximo - mínimo. Em seguida, divida o resultado pelo número de classes. Por fim, arredonde até o próximo número conveniente. $(125 - 67)/5 = 11,6$ (arredondado para 12)
- ▶ 3. Calcule os limites das classes.
 - ▶ O limite inferior da classe é o valor mais baixo que pertence a ela e o limite superior é o mais alto. Use o valor mínimo (67) como limite inferior da primeira classe.

Diagrama de Ramo e folhas

6	7
7	18
8	25677
9	25799
10	01233455789
11	268
12	245

Construa uma Tabela de Frequência

Mínimo = 67, Máximo = 125

Número de classes = 5

Amplitude de classe = 12

Classe	Limites	Contagens	f
67	79		3
79	91		5
91	103		8
103	115		9
115	127		5
			$\Sigma f = 30$

Faça primeiro todos os limites inferiores.

Tabela de Frequência

Limites	X_i	f_i	$f_i\%$	F_i	$F_i\%$
67 -79	73	3	10,00	3	10,00
79 -91	85	5	16,67	8	26,67
91 -103	97	8	26,67	16	53,33
103 -115	109	9	30,00	25	83,33
115 -127	121	5	16,67	30	100,00
Total	-	30	100,00	-	-

X_i = Ponto médio

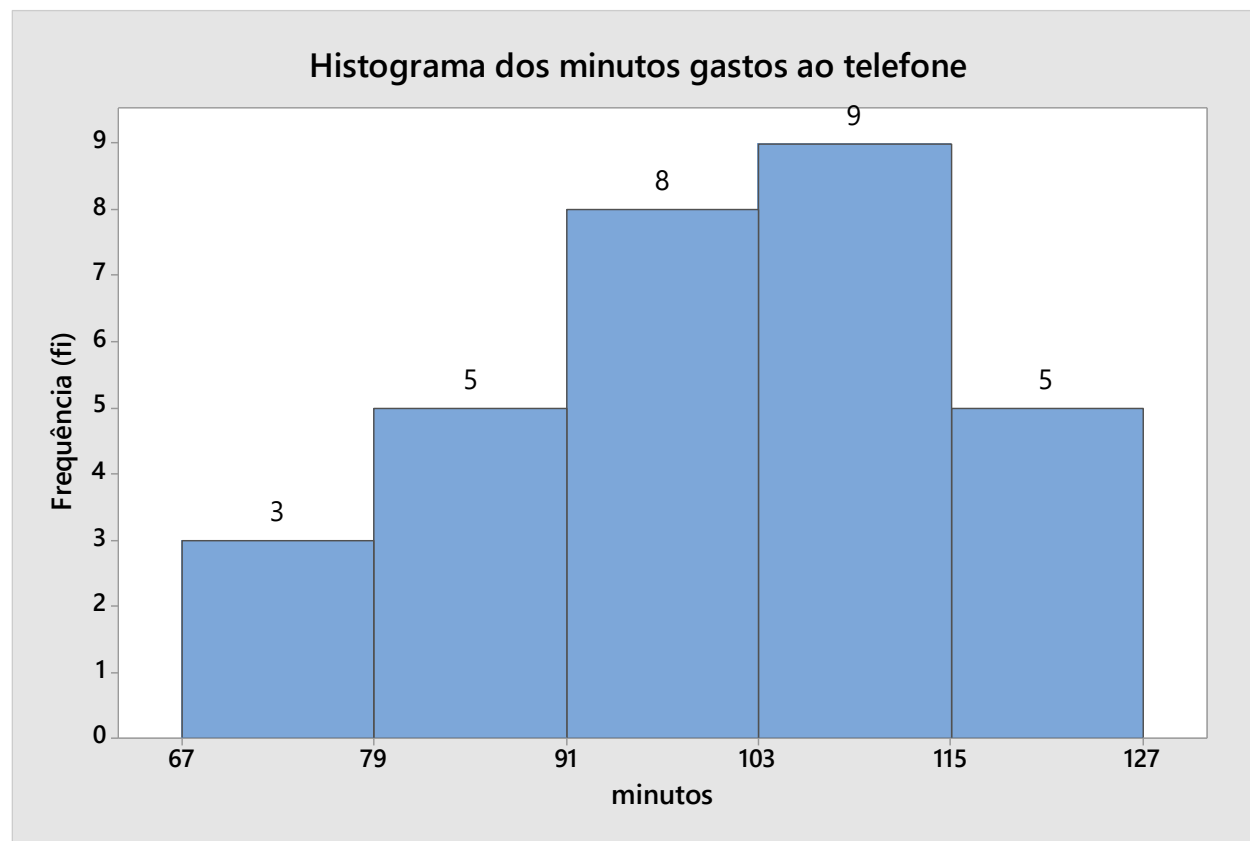
f_i = Frequência absoluta

$f_i\%$ = Frequência percentual

F_i = Frequência acumulada

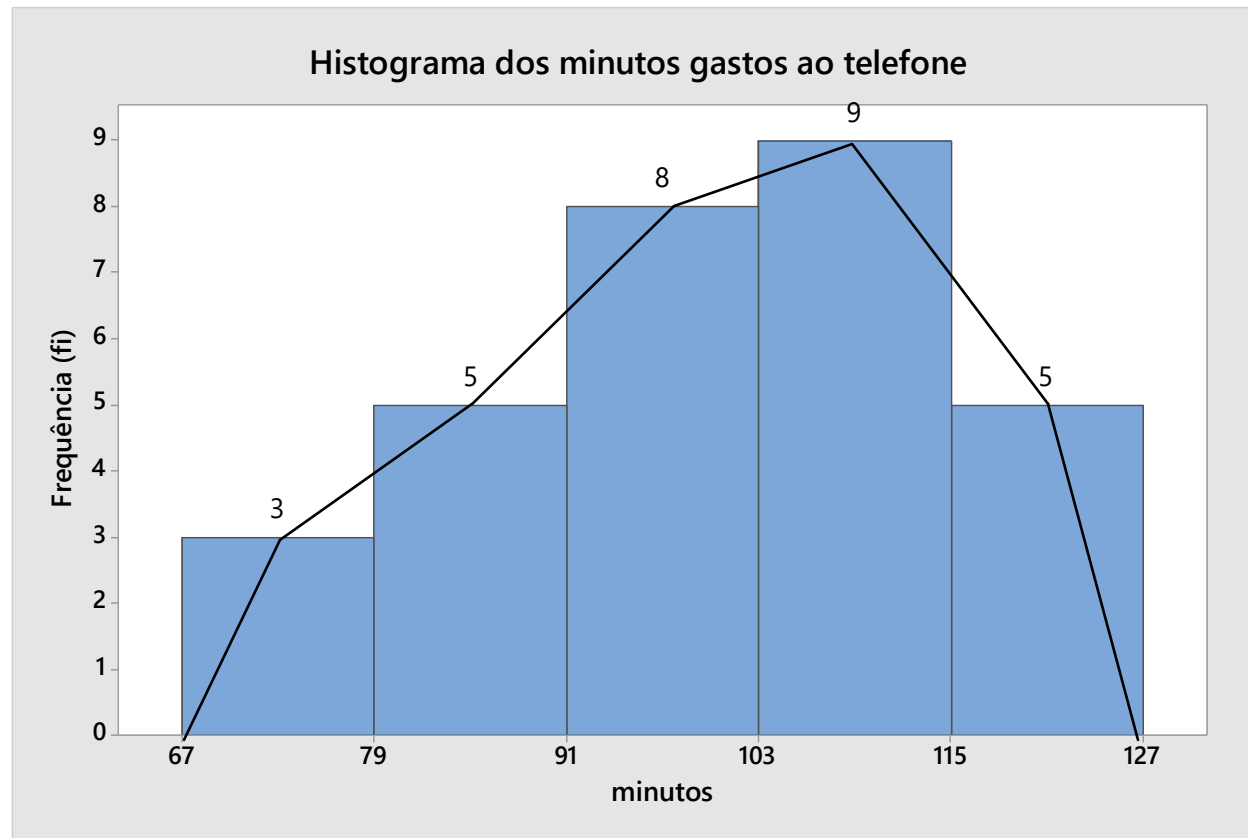
$F_i\%$ = Frequência acumulada percentual

Histograma



Limites	fi
67 -79	3
79 -91	5
91 -103	8
103 -115	9
115 -127	5

Histograma - com polígono de frequência

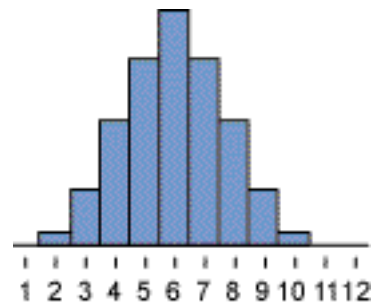


Limites	fi
67 -79	3
79 -91	5
91 -103	8
103 -115	9
115 -127	5

Marque o ponto médio no topo de cada barra. Conecte os pontos médios consecutivos. Estenda o polígono até os eixos.

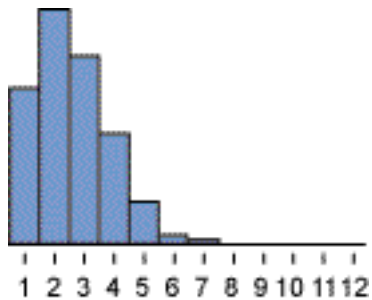
Histograma - Tipos de distribuições

Simétrica



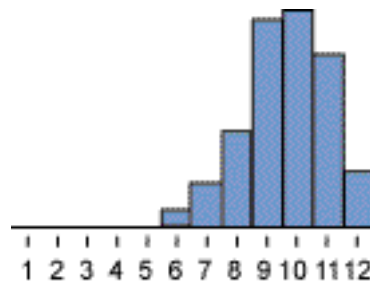
Média = Mediana

Assimétrica à direita



Média > Mediana

Assimétrica à esquerda



Média < Mediana

Cálculo de Média e Desvio Padrão para dados agrupados

- ▶ Quando os dados estão agrupados em tabelas de frequência temos que:

$$\bar{x} = \frac{\sum fixi}{n}$$

$$s = \sqrt{\frac{n \sum fixi^2 - (\sum fixi)^2}{n(n-1)}}$$

Exemplo

xi	fi	fixi	fixi^2
73	3	219	15987
85	5	425	36125
97	8	776	75272
109	9	981	106929
121	5	605	73205
soma	30	3006	307518

$$\bar{x} = \frac{3006}{30} = 100,2$$

$$s = \sqrt{\frac{30 * 307518 - 3006^2}{30(30 - 1)}} = 14,76$$

Tabela de Frequência

Limites	Xi	fi	fi%	Fi	Fi%
67 -79	73	3	10,00	3	10,00
79 -91	85	5	16,67	8	26,67
91 -103	97	8	26,67	16	53,33
103 -115	109	9	30,00	25	83,33
115 -127	121	5	17,67	30	100,00
Total	-	30	100,00	-	-

$$mediana = \tilde{x} = LScm - \frac{h}{f_{icm}} \left(F_{icm} - \frac{n}{2} \right) \quad cm = classe da mediana \quad Fi \geq \frac{n}{2}$$

$$mediana = \tilde{x} = 103 - \frac{12}{8} (16 - 15) = 103 - 1,5 = 101,5 \text{ minutos}$$

moda = xi da classe mais frequente = 109 minutos

$$Pm = \frac{LIT + LST}{2} = \frac{67 + 127}{2} = 97 \text{ minutos}$$

LIT = limite inferior da tabela ; LST = limite superior da tabela

Exercício

- Considere um estudo onde estamos interessados em avaliar o valor faturado bruto de 60 empresas do setor de calçados. Complete a tabela de frequência abaixo, e calcule a média, mediana, moda, ponto médio, desvio padrão e CV. Comente os resultados.

Faturado em milhões de Reais	xi	fi	fi%
0,2 ---- 0,6			13,33%
			21,67%
			26,67%
			20,00%
			10,00%
			8,33%
Total			100,00%

Exercício

- Considere um estudo onde estamos interessados em avaliar o valor faturado bruto de 60 empresas do setor de calçados. Complete a tabela de frequência abaixo, construa o histograma, calcule a média e o desvio padrão. Comente os resultados.

Faturado em milhões de Reais	xi	fi	fi%
0,2 ---- 0,6	0,4	8	13,33%
0,6 ---- 1,0	0,8	13	21,67%
1,0 ---- 1,4	1,2	16	26,67%
1,4 ---- 1,8	1,6	12	20,00%
1,8 ---- 2,2	2,0	6	10,00%
2,2 ---- 2,6	2,4	5	8,33%
Total	-	60	100,00%

Exercício

- Considere um estudo onde estamos interessados em avaliar o valor faturado bruto de 60 empresas do setor de calçados. Complete a tabela de frequência abaixo, construa o histograma, calcule a média e o desvio padrão. Comente os resultados.

x_i	f_i	fix_i	fix_i^2
0,4	8	3,20	1,28
0,8	13	10,40	8,32
1,2	16	19,20	23,04
1,6	12	19,20	30,72
2,0	6	12,00	24,00
2,4	5	12,00	28,80
	60	76,00	116,16

$$\bar{x} = \frac{76}{60} = 1,27 \text{ milhões}$$

$$S = \sqrt{\frac{60 * 116,16 - 76^2}{60(60 - 1)}} = 0,58 \text{ milhões}$$

Exercício

- Considere um estudo onde estamos interessados em avaliar o valor faturado bruto de 60 empresas do setor de calçados. Complete a tabela de frequência abaixo, construa o histograma, calcule a média e o desvio padrão. Comente os resultados.

Faturado em milhões de Reais	xi	fi	Fi
0,2 ---- 0,6	0,4	8	8
0,6 ---- 1,0	0,8	13	21
1,0 ---- 1,4	1,2	16	37 * cm
1,4 ---- 1,8	1,6	12	49
1,8 ---- 2,2	2,0	6	55
2,2 ---- 2,6	2,4	5	60
Total	----	60	

$$mediana = \tilde{x} = LScm - \frac{h}{ficm} (Ficm - \frac{n}{2})$$

$h = \text{amplitude de classe}$

$$\tilde{x} = 1,4 - \frac{0,4}{16} (37 - 30) = 1,4 - 0,175 = 1,23 \text{ milhões}$$

$\text{moda} = xi \text{ da classe mais frequente} = 1,2 \text{ milhões}$

$$Pm = \frac{LIT + LST}{2} = \frac{0,2 + 2,6}{2} = 1,4 \text{ milhões}$$

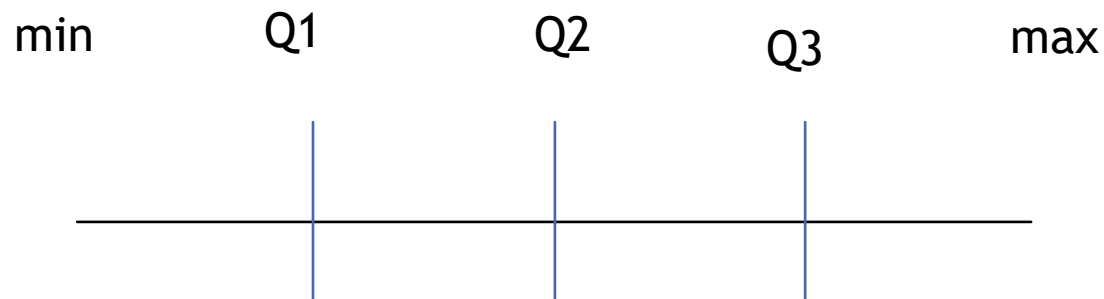
Medidas de Posição - Quartil (Q_1 , Q_2 e Q_3)

- ▶ Assim como a mediana divide o conjunto em 2 partes, temos os quartís (Q_1 , Q_2 e Q_3) que dividem o conjunto de dados em 4 partes.
- ▶ Assim:
 - ▶ 25 % dos dados estão abaixo de Q_1
 - ▶ 50 % dos dados estão abaixo de Q_2
 - ▶ 75 % dos dados estão abaixo de Q_3
- ▶ Para encontrar os quartís devemos:
 - ▶ 1 - ordenar dos dados de forma crescente
 - ▶ 2 - encontrar a localização do quartil k $LQ_k = k(n+1)/4$
 - ▶ Se LQ_k for inteiro o quartil k será o valor no posto LQ_k , caso contrário será a média ponderada dos dois valores mais próximos ao posto LQ_k .

Exemplo

- Você é gerente de uma loja. O volume de vendas (em salários mínimos - SM) em 27 dias corridos, selecionados aleatoriamente, em determinado ano, é dado abaixo. Determine Q_1 , Q_2 e Q_3 .

28 43 48 61 43 30 75 44 48 33 45 37 37 42 27 47 42 23 46 39 20 45
38 19 17 35 45



Resolução

- ▶ Ordene os dados de forma crescente:

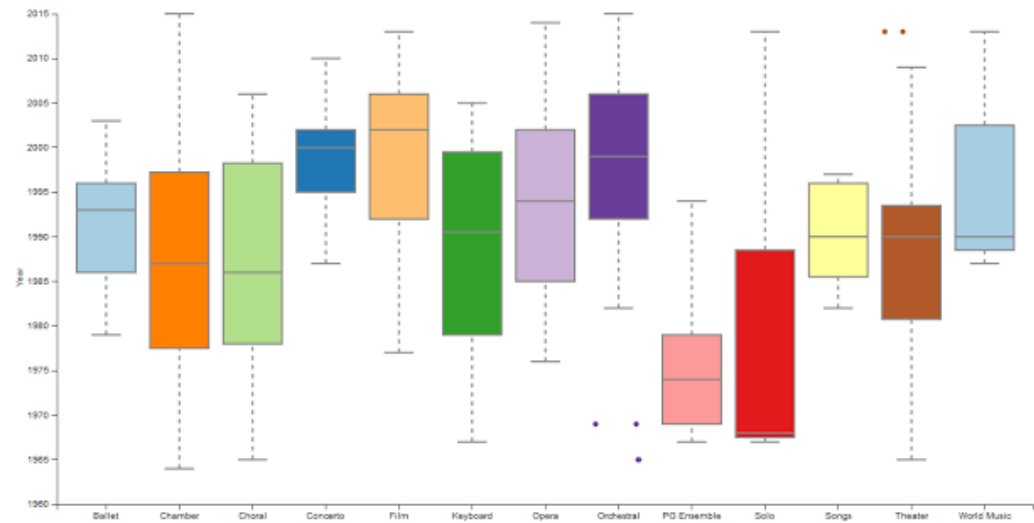
17 19 20 23 27 28 30 33 35 37 37 38 39 42 42 43 43 44 45 45 45 46
47 48 48 61 75.

- ▶ Encontre a localização dos Quartís : Q_1 , Q_2 e Q_3

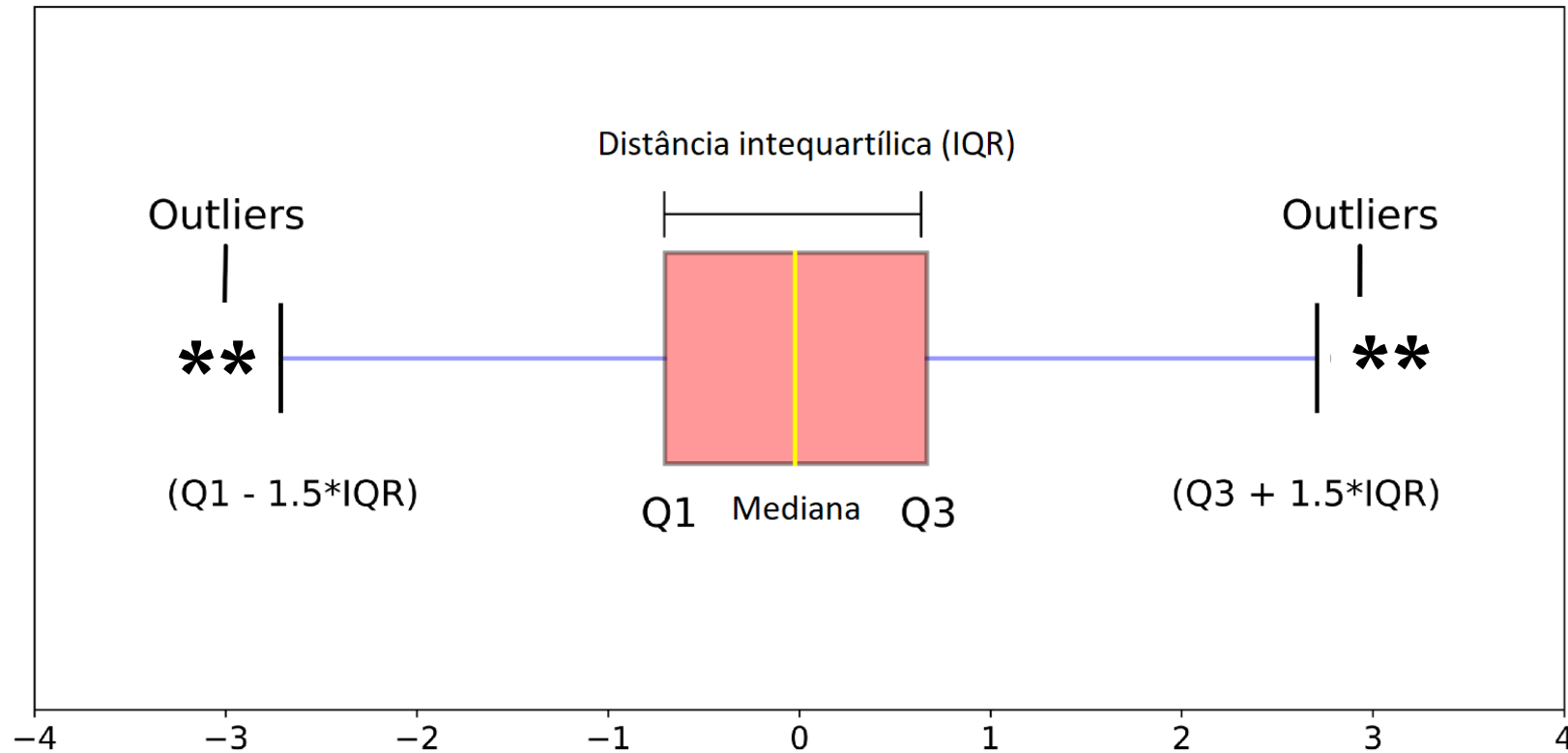
- ▶ $LQ1 = 1(27+1)/4 = 7$ “posto” - Assim $Q1 = 30$ SM
- ▶ $LQ2 = 2(27+1)/4 = 14$ “posto” - Assim $Q2 = 42$ SM
- ▶ $LQ3 = 3(27+1)/4 = 21$ “posto” - Assim $Q3 = 45$ SM

Box - plot

- ▶ É o gráfico que representa os quartís.



Box Plot - A caixa Mágica



Exemplo

- Para o conjunto de dados das vendas em salários mínimos, construa o box plot.

17 19 20 23 27 28 30 33 35 37 37 38 39 42 42 43 43 44 45 45 45 46
47 48 48 61 75.

$Q1=30$, $Q2=42$ e $Q3= 45$

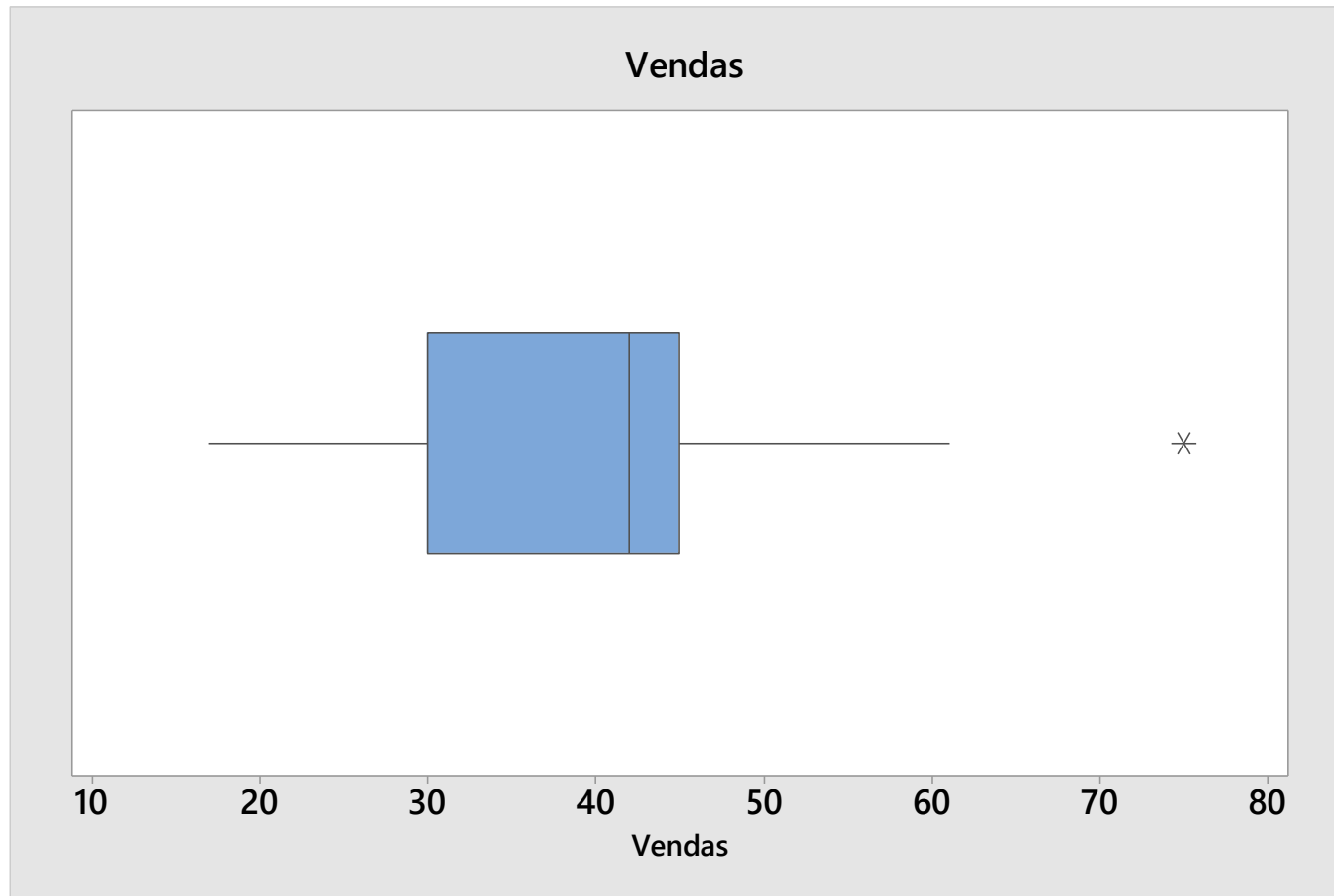
$IQR = 45-30 = 15$

$LI = 30 - 1,5*15 = 7,5 \text{ SM}$

$LS =45 + 1,5*15 = 67,5 \text{ SM}$

Outlier = 75 SM

Box - Plot



Exercícios

Os salários mensais em reais dos 25 funcionários da primeira turma de trainee de determinada empresa foram os seguintes (em mil reais):

2,90 2,90 2,95 2,95 3,10 3,10 3,15 3,20 3,20 3,25 3,30 3,40 3,45 3,45 3,50 3,65
3,65 3,80 3,90 3,90 4,00 5,00 5,20 5,50 6,40.

Construa o boxplot. Comente.