Tratamento e Análise de Dados e Informações

Medidas de Posição

Crespo (2009) – Capítulo 5

- **As Distribuições de Frequência** nos permite descrever de modo geral os grupos de valores que a variável pode assumir.

- Ressaltam também as tendências características de cada distribuição. Essas tendências podem ser traduzidos por conceitos denominados **elementos típicos da distribuição**.

Os principais elementos típicos da distribuição são:

1- As Medidas de Posição

2- As Medidas de Dispersão

3- Medidas de Assimetria

4- Medidas de Curtose

As medidas de posição mais importantes são as medidas de tendência central:

1- Média

2- Moda

3- Mediana

1- Média

É utilizada quando:

a. Desejamos obter a medida que possui maior estabilidade;

b. Houver necessidade de um tratamento algébrico posterior;

c. Uma medida simples de inferência: pode ser uma medida que representam todos os dados;

i. **Aritmética Simples:** é o quociente da divisão da soma dos valores da variável pelo número deles:

$$\bar{X} = \frac{\sum x_i}{n}$$

$\bar{X}$: média aritmética
x_i: valores da variável
n: número de valores

X	Frequência - fi
3,15	1
3,16	1
4,21	1
4,25	1
4,75	1
5,81	1
5,90	1
6,00	1

Média simples: $\bar{X} = \frac{\sum x_i}{n} = \frac{3{,}15+3{,}16+4{,}21+4{,}25+4{,}75+5{,}81+5{,}9+6}{8} = \mathbf{4{,}65}$

ii. Média Aritmética Ponderada: de uma maneira simplificada, os valores observados são multiplicados pela frequência que ocorrem e depois divididos pela soma das frequências. A frequência denota uma intensidade, um peso, uma ponderação para um intervalo ou variável.

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i}$$

X	Frequência Simples - fi	Frequência Relativa - fri
3	2	$\frac{2}{8} = 0{,}250$
4	3	$\frac{3}{8} = 0{,}375$
5	2	$\frac{2}{8} = 0{,}250$
6	1	$\frac{1}{8} = 0{,}125$
TOTAL	8	1,0

Média ponderada: $\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{3\times2+4\times3+5\times2+6\times1}{8} = 3 \times 0{,}25 + 4 \times 0{,}375 + 5 \times 0{,}25 + 6 \times 0{,}125 =$ **4,25**

Classe	X	Frequência - fi
1	3 ⊢ 4	2
2	4 ⊢ 5	3
3	5 ⊢ 6	2
4	6 ⊢ 7	1

Média ponderada: $\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{3{,}5 \times 2 + 4{,}5 \times 3 + 5{,}5 \times 2 + 6{,}5 \times 1}{8} = \frac{7 + 13{,}5 + 11 + 6{,}5}{8} = 4{,}75$

iii. Média Geométrica: resultado da raiz da multiplicação de todos valores da variável.

$$G = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} = \sqrt[n]{x_1 \times x_2 \times x_3 \ldots x_n}$$

X	Frequência - fi
3,15	1
3,16	1
4,21	1
4,25	1
4,75	1
5,81	1
5,90	1
6,00	1

Média geométrica: $G = \sqrt[n]{x_1 \times x_2 \times x_3 \ldots x_n} = \sqrt[8]{3{,}15 \times 3{,}16 \times 4{,}21 \times 4{,}25 \times 4{,}75 \times 5{,}81 \times 5{,}90 \times 6} =$ **4,51**

iv. Média Harmônica: número total de valores que a variável assume n, dividido pela soma do inverso de cada um.

$$\mathrm{H} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

X	Frequência - fi
3,15	1
3,16	1
4,21	1
4,25	1
4,75	1
5,81	1
5,90	1
6,00	1

Média Harmônica: $H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{8}{\frac{1}{3{,}15} + \frac{1}{3{,}16} + \frac{1}{4{,}21} + \frac{1}{4{,}25} + \frac{1}{4{,}75} + \frac{1}{5{,}81} + \frac{1}{5{,}9} + \frac{1}{6}} = \mathbf{4,38}$

Desvio Simples – Desvio em Relação à Média - d

É a diferença entre cada elemento de um conjunto de valores e a média aritmética.

$$d = x_i - \bar{x}$$

Propriedade das Médias

1 – A soma algébrica de todos desvios é igual a zero

$$\sum_{i=1}^{k} d_i = 0$$

X	Média	Desvio
3,15	4,65375	≅ -1,50
3,16	≅ 4,65	≅ -1,49
4,21	4,65	-0,44
4,25	4,65	-0,40
4,75	4,65	0,10
5,81	4,65	1,16
5,90	4,65	1,25
6,00	4,65	1,35

Soma dos desvios ≅ 0

Propriedade das Médias

2 – Somando-se ou subtraindo-se uma constante c de todos valores de uma variável, a média do conjunto fica aumentada ou diminuída nessa constante

X	Média	X +2	Nova Média
3,15	4,65	5,15	6,65
3,16	4,65	5,16	6,65
4,21	4,65	6,21	6,65
4,25	4,65	6,25	6,65
4,75	4,65	6,75	6,65
5,81	4,65	7,81	6,65
5,90	4,65	7,90	6,65
6,00	4,65	8,00	6,65

Propriedade das Médias

3 – Multiplicando-se ou dividindo-se todos valores de uma variável por uma constante c, a média do conjunto fica multiplicada ou dividida por essa constante.

X	Média	X * 2	Nova Média
3,15	4,65	6,30	9,30
3,16	4,65	6,32	9,30
4,21	4,65	8,42	9,30
4,25	4,65	8,5	9,30
4,75	4,65	9,5	9,30
5,81	4,65	11,62	9,30
5,90	4,65	11,80	9,30
6,00	4,65	12,00	9,30

2- Moda

É o valor que ocorre com maior frequência na série de dados.

É utilizada quando:

a. Desejamos obter a medida rápida e aproximada de posição;

b. Quando a medida de posição deve ser o valor mais típico da distribuição;

c. Como uma medida simples de inferência: pode ser uma medida que representam todos os dados;

Dados não agrupados

Basta identificar o valor que mais se repete.

Dados agrupados

Basta fixar o valor de maior frequência.

X	Frequência
10	2
20	3
30	**4**
40	3
∑ TOTAL	12

$M_o = 30$

$\{7,8,9,10,10,10,11,12,13,15\}$ $M_o = 30$

X	Frequência - fi
3,15	1
3,16	1
4,21	1
4,25	1
4,75	1
5,81	1
5,90	1
6,00	1

Série amodal : não possui moda

Série: {2,3,4,4,4,5,6,7,7,7,8,9}

X	Frequência - fi
2	1
3	1
4	**3**
5	1
6	1
7	**3**
8	1
9	1

Série bimodal: possui duas modas: 4 e 7

A classe que apresenta a maior frequência é denominada **classe modal**.

Classe	X	Frequência - fi
1	3 ⊢ 4	2
2	**4 ⊢ 5**	**3**
3	5 ⊢ 6	2
4	6 ⊢ 7	1

Por definição o valor da moda para a série é o ponto médio entre os limites da classe, denominada de **moda bruta**:

$$M_o = \frac{l+L}{2} \qquad M_o = \frac{4+5}{2} = 4{,}5$$

2- Mediana

É o valor que se encontra no centro de uma série de números.

É utilizada quando:

a. Desejamos obter o ponto que divide a distribuição em partes iguais;

b. Há valores extremos que afetam de maneira acentuada a média;

c. Como uma medida simples de inferência: pode ser uma medida que representam todos os dados;

Dados não agrupados

Dada uma série de valores:

$$\{5,13,10,2,18,15,6,16,9\}.$$

O primeiro passo é ordenar os valores:

$$\{2,5,6,9,10,13,15,16,18\}.$$

Em seguida tomar o valor central que apresenta o mesmo número de elementos à esquerda e à direita:

$$M_D = 10$$

Caso a série conter um número par de elementos, por definição a mediana será qualquer dos números compreendidos entre os dois valores centrais da série. Convencionou-se utilizar o **ponto médio**.

{2,5,6,9,10,12,13,15,16,18}

$$M_D = 11$$

Caso a série conter um número ímpar de elementos a mediana M_D é:

$$M_D = \frac{n+1}{2}$$

Caso a série conter um número par de elementos a mediana M_D é:

Média dos termos $\frac{n}{2}$ e $\frac{n}{2}+1$

Dados agrupados

Calcula-se a mediana de maneira semelhante aso dados não agrupados, porém implicando a determinação prévia das frequências acumuladas

$$M_D = \frac{\sum f_i}{2}$$

X	Frequência - fi	Frequência Acumulada
0	2	2
1	6	8
2	10	18
3	12	30
4	4	34
Total	34	

$$M_D = \frac{\sum f_i}{2} = \frac{34}{2} = 17$$

A menor frequência acumulada que supera esse valor é 18, que corresponde ao valor 2 da variável.

No caso dos intervalos de classe passa a existir uma **classe mediana**.

Tal classe será aquela imediatamente superior a $\frac{\sum f_i}{2}$.

Feito isso, realizamos a interpolação admitindo-se que os valores distribuam-se uniformemente em todo intervalo de classe.

Passo a passo para interpolação :

1° - Determinamos a frequência acumulada

2° - Calculamos $\frac{\sum f_i}{2}$

3° - Marcamos a classe correspondente à frequência acumulada imediatamente superior à classe mediana.

4° - Empregamos a fórmula para mediana:

$$M_d = l^* + \frac{\left[\frac{\sum f_i}{2} - F(ant)\right] h^*}{f^*}$$

l^*: limite inferior da classe mediana;
$F(ant)$: frequência acumulada da classe anterior à classe mediana;
f^*: frequência simples da classe mediana;
h^*:amplitude do intervalo da classe mediana;

Classe	X	Frequência Simples - fi	Frequência Acumulada - Fi
1	150 ⊢ 154	4	4
2	154 ⊢ 158	9	13
3	158 ⊢ 162	11	24
4	162 ⊢ 166	8	32
5	166 ⊢ 170	5	37
6	170 ⊢ 174	3	40
∑ TOTAL		40	

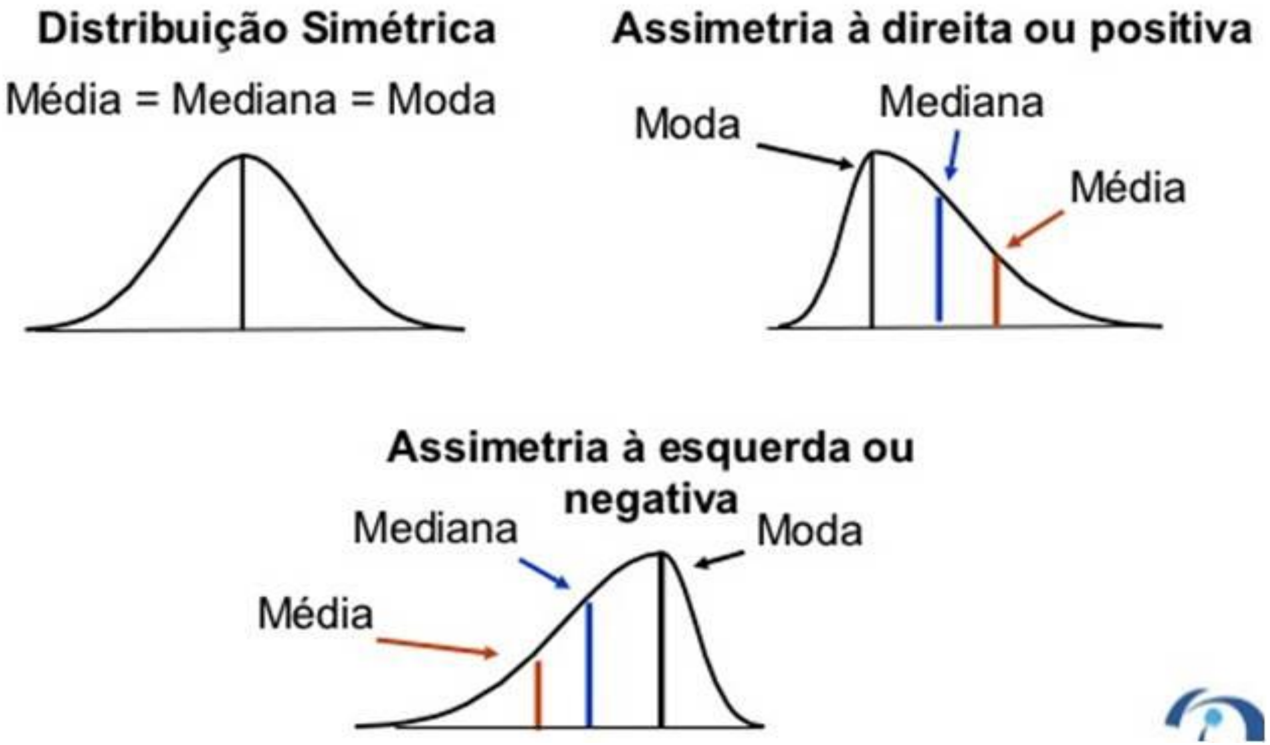
1° - Determinamos a frequência acumulada

2° - Calculamos $\frac{\sum f_i}{2} = \frac{40}{2} = 20$. Procuramos o valor que ocupa o 20° lugar.

3° - Marcamos a classe correspondente à frequência acumulada imediatamente superior à classe mediana. Ele está na classe 3, supondo frequências uniformemente distribuídas.

4° - Empregamos a fórmula para mediana: $M_d = l^* + \frac{\left[\frac{\sum f_i}{2} - F(ant)\right]h^*}{f^*} = 158 + \frac{[20-13]4}{11} = \mathbf{160,5}$

Posição relativa da média, mediana e moda



$\bar{X} = M_o = M_d$ **curva simétrica**.
Caso na distribuição normal.

$\bar{X} > M_d > M_o$ **curva assimétrica positiva**.

$\bar{X} < M_d < M_o$ **curva assimétrica negativa.**

Separatrizes

Além das medidas de tendência central há as separatrizes, que considerando também a mediana, separam as séries em partes iguais.

Quartis: os valores de uma série que a divide em quatro partes iguais;

Primeiro Quartil (Q_1)**:** valor situado na série de forma que uma quarta parte 25% dos dados é menor que ele e as três partes restantes 75% são maiores.

Segundo Quartil (Q_2)**:** coincide com a mediana: valor que divide a série em duas partes iguais.

Terceiro Quartil (Q_3)**:** valor situado de tal modo que as três quartas partes (75%) dos termos são menores que ele e uma quarta parte (25%) é maior.

$$Q_k = l^* + \frac{\left[\frac{k\sum f_i}{4} - F(ant)\right] h^*}{f^*}$$

k**:**Número de ordem do quartil

Percentis: os 99 valores que separam uma série em 100 partes iguais;

$$P_k = l^* + \frac{\left[\frac{k \sum f_i}{100} - F(ant)\right] h^*}{f^*}$$

k:Número de ordem do percentil

Decis: os 9 valores que separam uma série em 10 partes iguais;

$$D_k = l^* + \frac{\left[\frac{k \sum f_i}{10} - F(ant)\right] h^*}{f^*}$$

k:Número de ordem do decil

Competição

Get ready to compete!

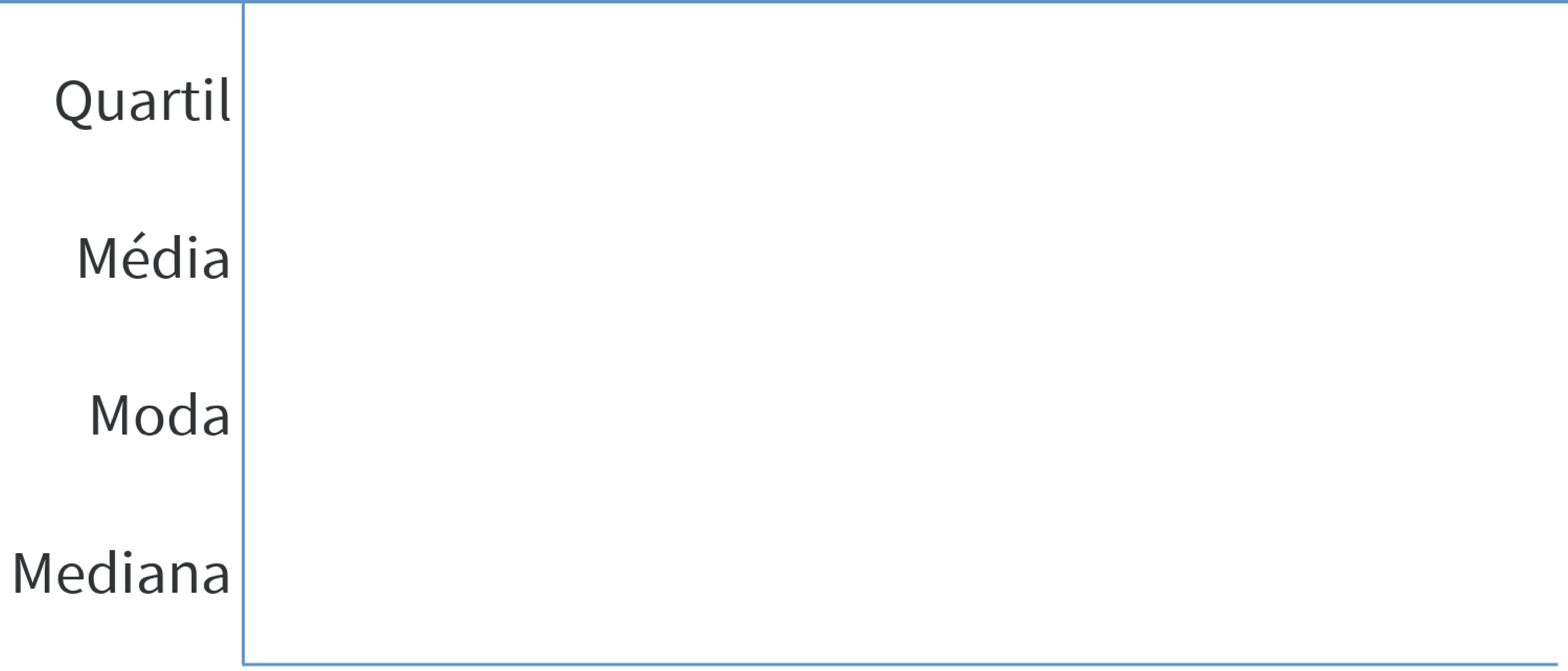
Não é uma medida de tendência central:

Não é uma medida de tendência central:

- Quartil
- Média
- Moda
- Mediana

Total Results: 0

Não é uma medida de tendência central:



Leaderboard

As médias podem ser:

As médias podem ser:

- Geométricas
- Aritméticas
- Agregadas
- Harmônicas

Total Results: 0

As médias não podem ser:

Geométricas

Aritméticas

Agregadas

Harmônicas

Leaderboard

Desvio em relação a média é:

Desvio em relação a média é:

- Alteração significativa do valor da série em relação à média
- Soma de todos valores da série divido pela quantidade de valores
- O valor da média dividido por cada valor observado
- Diferença entre o valor da série e a média

Total Results: 0

Desvio em relação a média é:

Alteração significativa do valor da série em relação à média

Soma de todos valores da série divido pela quantidade de valores

O valor da média dividido por cada valor observado

Diferença entre o valor da série e a média

Leaderboard

A soma de todos os desvios resulta em:

A soma de todos os desvios resulta em:

- 100% de cobertura amostral
- Distribuição de frequência
- Zero
- Dados normalizados

Total Results: 0

A soma de todos os desvios resulta em:

- 100% de cobertura amostral
- Distribuição de frequência
- Zero
- Dados normalizados

Leaderboard

Se todos valores de uma série de dados forem dividos por 3, pode remos observar que:

Se todos valores de uma série de dados forem dividos por 3, pode remos observar que:

- Sua média não se altera.
- A média também será divida por 3
- A média triplicará.
- A média diminuirá, mas numa proporção menor.

Total Results: 0

Leaderboard

Moda é:

Moda é:

O estilo dos dados.

Valor que divido a série de dados em duas partes iguais.

Valor que ocorre com maior frequência.

Forma como os dados são apresentados.

Total Results: 0

Moda é:

O estilo dos dados.

Valor que divido a série de dados em duas partes iguais.

Valor que ocorre com maior frequência.

Forma como os dados são apresentados.

Leaderboard

Mediana é:

Mediana é:

O valor que divide a série em dois tamanhos iguais.

O valor que se encontra entre dois valores de dados quaisquer.

O valor que representa o ponto médio entre dois valores de dados .

A metade de uma série de dados.

Total Results: 0

Mediana é:

O valor que divide a série em dois tamanhos iguais.

O valor que se encontra entre dois valores de dados quaisquer.

O valor que representa o ponto médio entre dois valores de dados .

A metade de uma série de dados.

Leaderboard

Quantos são necessários serem encontrados para dividir a série em quartis?

Quantos são necessários serem encontrados para dividir a série em quartis?

4

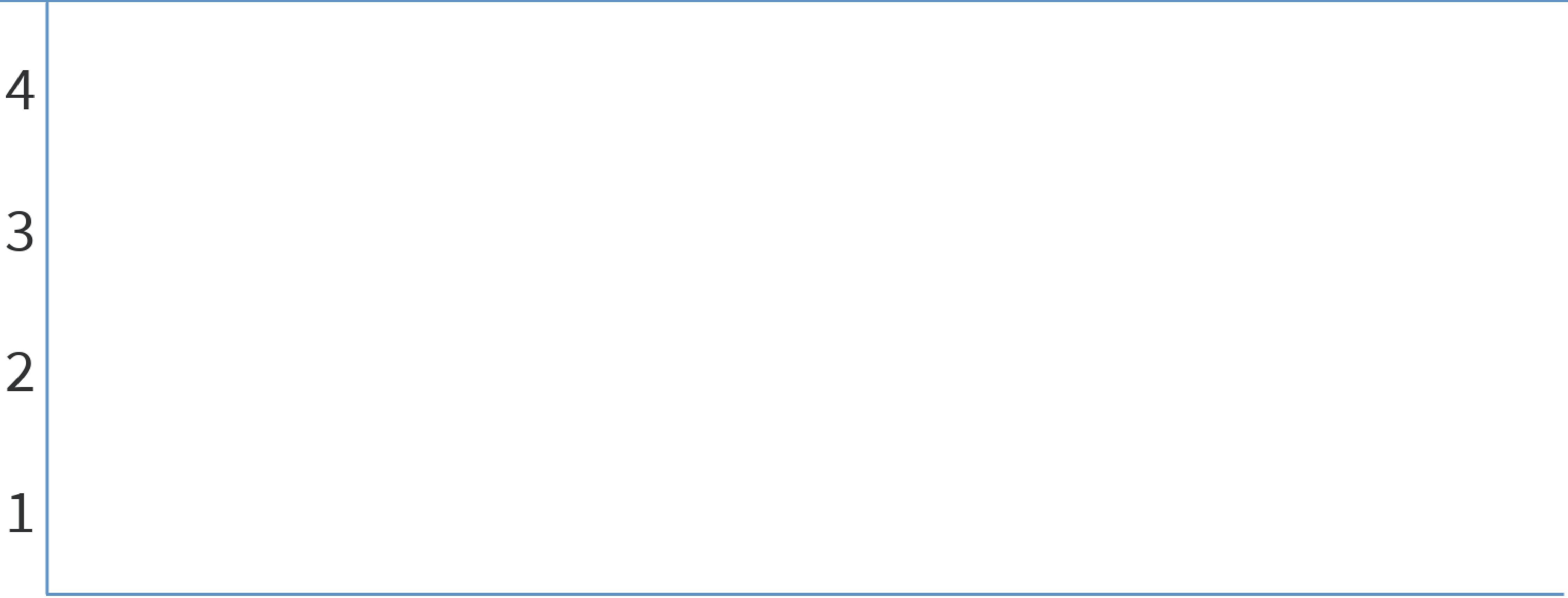
3

2

1

Total Results: 0

Quantos são necessários serem encontrados para dividir a série em quartis?



Leaderboard

Uma distribuição assimétrica positiva apresenta:

Uma distribuição assimétrica positiva apresenta:

$Me = Mo = Md$

$Me > Md > Mo$

$Me < Md < Mo$

Total Results: 0

Uma distribuição assimétrica positiva apresenta:

$Me = Mo = Md$

$Me > Md > Mo$

$Me < Md < Mo$

Leaderboard