

Introdução à Estatística

O que é?

Ramo da matemática que trata da:

- Coleta
- Análise
- Interpretação
- Apresentação

de massas de dados numéricos.

Coleta

População e Amostra

População: é um grupo completo de elementos que possuem uma mesma característica.

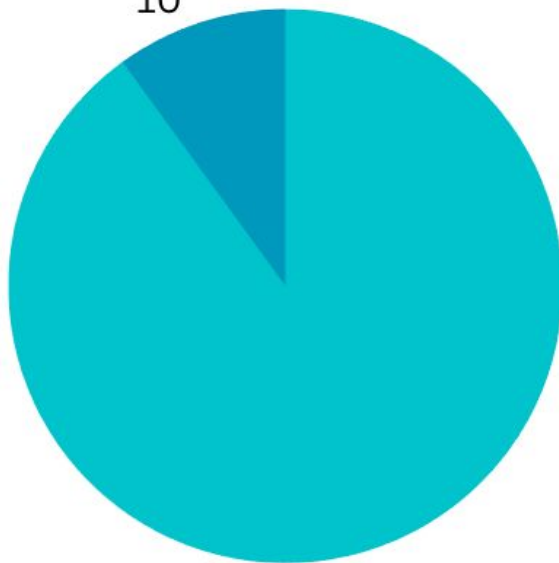
Amostra: é um sub-conjunto da população, que pode ser estudado para gerar inferências sobre uma característica da população.

Exemplos

População - 100 pessoas

Amostra

10



Exemplos

- Consumo de produtos
- Intenção de votos
- Renda média
- Análise de clientes
- Pesquisa de satisfação

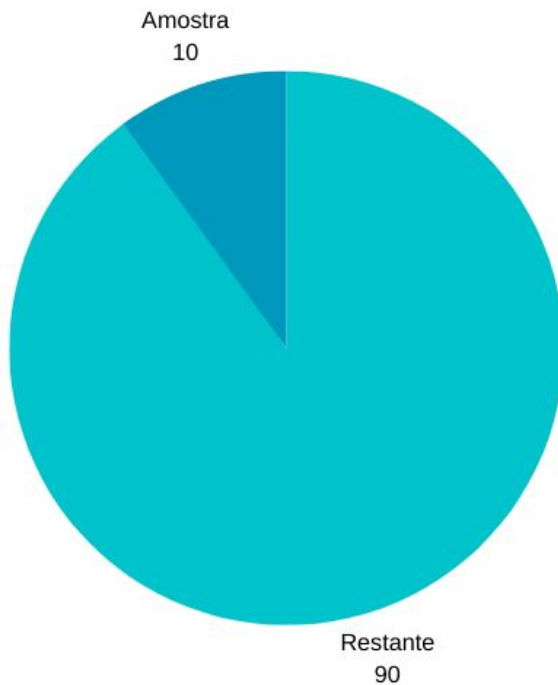
Dados absolutos e relativos

Dados absolutos: são dados os dados coletados da fonte sem qualquer tipo de transformação.

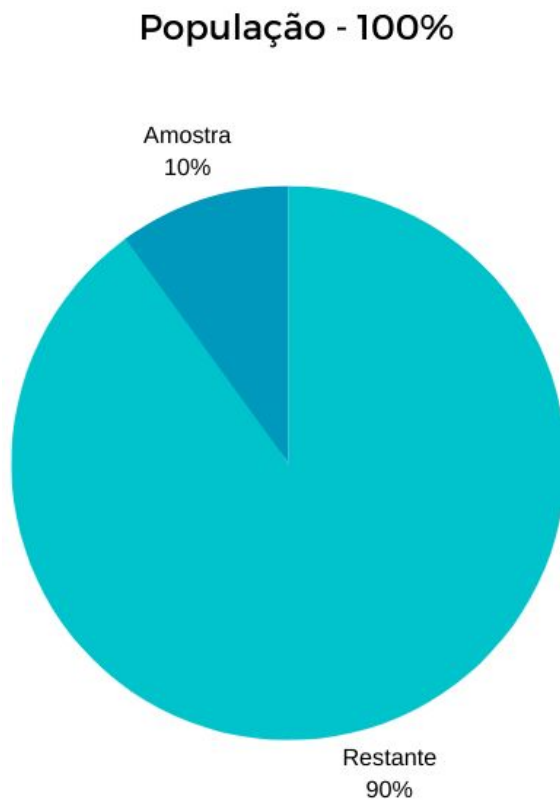
Dados relativos: são valores obtidos a partir dos dados absolutos, geralmente por divisões, utilizados para comparações (índices, coeficientes, taxas, porcentagens, etc.)

Exemplos

População - 100 pessoas



Exemplos



Análise

Análise

- Distribuição de frequência
- Medidas de posição
- Medidas de dispersão
- Distribuição normal

Distribuição de Frequência

Pode ser representada por uma tabela de duas colunas, onde a primeira representa um grupo ou categoria dos dados, e a segunda representa o número de vezes que aquela categoria foi observada. Também é comum ver distribuição de frequências absolutas e relativas.

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|---------|---------------|
| João | Azul |
| Ana | Vermelho |
| Nasser | Vermelho |
| Paulo | Azul |
| Matheus | Vermelho |
| Clara | Azul |
| Alex | Vermelho |

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|---------|---------------|
| João | Azul |
| Ana | Vermelho |
| Nasser | Vermelho |
| Paulo | Azul |
| Matheus | Vermelho |
| Clara | Azul |
| Alex | Vermelho |

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|------|---------------|
| Azul | 3 |

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|---------|---------------|
| João | Azul |
| Ana | Vermelho |
| Nasser | Vermelho |
| Paulo | Azul |
| Matheus | Vermelho |
| Clara | Azul |
| Alex | Vermelho |

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|----------|---------------|
| Azul | 3 |
| Vermelho | 4 |

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|----------|---------------|
| Azul | 3 |
| Vermelho | 4 |

Distribuição de frequência absoluta ou relativa?

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|----------|---------------|
| Azul | 3 |
| Vermelho | 4 |

Distribuição de frequência absoluta ou relativa?

Absoluta

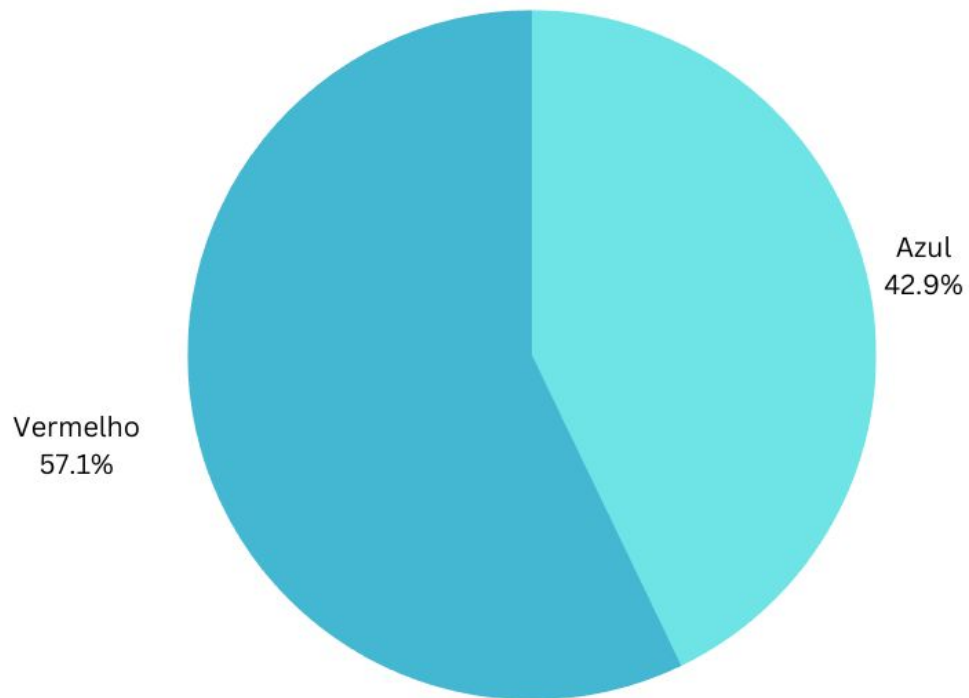
Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|----------|---------------|
| Azul | 3 |
| Vermelho | 4 |
| Total | 7 |

Distribuição de frequência - Exemplo

| Nome | Cor preferida |
|----------|----------------|
| Azul | $3/7 = 42,9\%$ |
| Vermelho | $4/7 = 57,1\%$ |
| Total | $7/7 = 100\%$ |

Distribuição de frequência - Exemplo



Medidas de posição

Utilizadas para resumir um conjunto de dados.

Medidas de posição

- Moda
- Média
- Mediana
- Quartil
- Percentil
- Min
- Max

Medidas de posição - Média

É a soma de todos os valores dividido pela quantidade.

Medidas de posição - Média

| Bimestre | Notas |
|----------|-------|
| 1 | 8 |
| 2 | 7 |
| 3 | 10 |
| 4 | 5 |

Medidas de posição - Média

| Bimestre | Notas |
|----------|-------|
| 1 | 8 |
| 2 | 7 |
| 3 | 10 |
| 4 | 5 |

Medidas de posição - Média

| |
|----|
| 8 |
| 7 |
| 10 |
| 5 |

$$8+7+10+5 = 30$$

Medidas de posição - Média

| |
|----|
| 8 |
| 7 |
| 10 |
| 5 |

4 elementos

$$8+7+10+5 = 30$$

Medidas de posição - Média

| |
|----|
| 8 |
| 7 |
| 10 |
| 5 |

4 elementos

$$8+7+10+5 = 30$$

$$\text{Média} = 30/4 = 7,5$$

Medidas de posição - Moda

É o valor mais frequente da série.

Medidas de posição - Moda

| Aluno | Médias |
|---------|--------|
| Matheus | 8 |
| Ana | 7 |
| Paulo | 10 |
| João | 5 |
| Nasser | 8 |

Medidas de posição - Moda

| Médias | Frequência |
|--------|------------|
| 8 | |
| 7 | |
| 10 | |
| 5 | |
| 8 | |

Medidas de posição - Moda

| Médias | Frequência |
|--------|------------|
| 8 | 2 |
| 7 | |
| 10 | |
| 5 | |

Medidas de posição - Moda

| Médias | Frequência |
|--------|------------|
| 8 | 2 |
| 7 | 1 |
| 10 | |
| 5 | |

Medidas de posição - Moda

| Médias | Frequência |
|--------|------------|
| 8 | 2 |
| 7 | 1 |
| 10 | 1 |
| 5 | 1 |



A moda é 8

Medidas de posição - Tipos de Moda

Dependendo do conjunto de dados, ela pode ser:

- **Sem moda:** quando nenhum valor se repete.
- **Unimodal:** quando existe apenas um valor repetido com maior frequência.
- **Bimodal:** quando existem dois valores com a mesma maior frequência.
- **Multimodal:** quando mais de dois valores se repetem com a mesma frequência.

Medidas de posição - Mediana

É o valor que separa a série ordenada no meio.

Medidas de posição - Mediana

| Aluno | Médias |
|---------|--------|
| Matheus | 8 |
| Ana | 7 |
| Paulo | 10 |
| João | 5 |
| Nasser | 8 |


Medidas de posição - Mediana

| Aluno | Médias |
|---------|--------|
| João | 5 |
| Ana | 7 |
| Nasser | 8 |
| Matheus | 8 |
| Paulo | 10 |

Ordenam-se os valores

Medidas de posição - Mediana

| Aluno | Médias |
|---------|--------|
| João | 5 |
| Ana | 7 |
| Nasser | 8 |
| Matheus | 8 |
| Paulo | 10 |



Mediana = 8

Medidas de posição - Mediana

| Aluno | Médias |
|---------|--------|
| João | 5 |
| Ana | 7 |
| Matheus | 8 |
| Paulo | 10 |

E se a quantidade de valores for par?

Medidas de posição - Mediana

| Aluno | Médias |
|---------|--------|
| João | 5 |
| Ana | 7 |
| Matheus | 8 |
| Paulo | 10 |

Nesses casos a mediana é a média entre os valores do meio

Medidas de posição - Mediana

| Aluno | Médias |
|---------|--------|
| João | 5 |
| Ana | 7 |
| Matheus | 8 |
| Paulo | 10 |

$$\text{Mediana} = (7+8)/2 = 7,5$$

Medidas de posição - Percentil

É o valor em que certa porcentagem da série ordenada fica abaixo.

Ex: Mediana é o valor que separa a série ordenada em ao meio, ou seja metade da série ordenada fica abaixo da mediana (percentil 50%).

Fórmula = $(\text{quantidade} + 1) * \text{percentil}$

*Arredondando para cima.

Medidas de posição - Percentil 50%

| | |
|---------|----|
| José | 5 |
| Maria | 9 |
| Ana | 4 |
| João | 5 |
| Julia | 10 |
| Paulo | 1 |
| Matheus | 3 |
| Nasser | 9 |
| Aline | 8 |

Medidas de posição - Percentil 50%

Primeiro ordena-se a série

| | |
|---------|----|
| José | 5 |
| Maria | 9 |
| Ana | 4 |
| João | 5 |
| Julia | 10 |
| Paulo | 1 |
| Matheus | 3 |
| Nasser | 9 |
| Aline | 8 |

Medidas de posição - Percentil 50%

Primeiro ordena-se a série

| | |
|---------|----|
| Paulo | 1 |
| Matheus | 3 |
| Ana | 4 |
| João | 5 |
| José | 5 |
| Aline | 8 |
| Maria | 9 |
| Nasser | 9 |
| Julia | 10 |

Medidas de posição - Percentil 50%

Segundo quantifica a série

| | |
|---------|----|
| Paulo | 1 |
| Matheus | 3 |
| Ana | 4 |
| João | 5 |
| José | 5 |
| Aline | 8 |
| Maria | 9 |
| Nasser | 9 |
| Julia | 10 |

9
Itens

Medidas de posição - Percentil 50%

Terceiro aplica a fórmula $(9+1)*0.5=5$

| | |
|---------|----|
| Paulo | 1 |
| Matheus | 3 |
| Ana | 4 |
| João | 5 |
| José | 5 |
| Aline | 8 |
| Maria | 9 |
| Nasser | 9 |
| Julia | 10 |

Medidas de posição - Percentil 50%

Quarto encontra o valor na posição

| | | |
|---------|----|---------------|
| Paulo | 1 | |
| Matheus | 3 | |
| Ana | 4 | |
| João | 5 | |
| José | 5 | ← Posição 5 |
| Aline | 8 | Percentil 50% |
| Maria | 9 | |
| Nasser | 9 | |
| Julia | 10 | |

Medidas de posição - Quartis

São marcas que dividem os valores ordenados em 4 partes (25%, 50% - mediana, e 75%).

- Q2 - 50% - Mediana - $(\text{quantidade}+1)*0.5$
- Q1 - 25% - Mediana da metade inferior - $(\text{quantidade}+1)*0.25$
- Q3 - 75% - Mediana da metade superior - $(\text{quantidade}+1)*0.75$

Medidas de posição - Quartis

| | |
|---------|----|
| José | 5 |
| Maria | 9 |
| Ana | 4 |
| João | 5 |
| Julia | 10 |
| Paulo | 1 |
| Matheus | 3 |
| Nasser | 9 |
| Aline | 8 |

Medidas de posição - Quartis

Primeiro ordena-se a série

| | |
|---------|----|
| José | 5 |
| Maria | 9 |
| Ana | 4 |
| João | 5 |
| Julia | 10 |
| Paulo | 1 |
| Matheus | 3 |
| Nasser | 9 |
| Aline | 8 |

Medidas de posição - Quartis

Primeiro ordena-se a série

| | |
|---------|----|
| Paulo | 1 |
| Matheus | 3 |
| Ana | 4 |
| João | 5 |
| José | 5 |
| Aline | 8 |
| Maria | 9 |
| Nasser | 9 |
| Julia | 10 |

Medidas de posição - Quartis

Segundo quantifica a série

| | |
|---------|----|
| Paulo | 1 |
| Matheus | 3 |
| Ana | 4 |
| João | 5 |
| José | 5 |
| Aline | 8 |
| Maria | 9 |
| Nasser | 9 |
| Julia | 10 |

9
Itens

Medidas de posição - Quartis

Terceiro definir Q2

$$(9+1)*0.5=5$$

| | |
|---------|----|
| Paulo | 1 |
| Matheus | 3 |
| Ana | 4 |
| João | 5 |
| José | 5 |
| Aline | 8 |
| Maria | 9 |
| Nasser | 9 |
| Julia | 10 |

Q2

Medidas de posição - Quartis

Quarto definir o Q1

$$(9+1)*0.25=2.5$$

| | | |
|---------|----|------|
| Paulo | 1 | |
| Matheus | 3 | |
| Ana | 4 | ← Q1 |
| João | 5 | |
| José | 5 | ← Q2 |
| Aline | 8 | |
| Maria | 9 | |
| Nasser | 9 | |
| Julia | 10 | |

Medidas de posição - Quartis

Quinto definir o Q3

$$(9+1)*0.75=7.5$$

| | | |
|---------|----|------|
| Paulo | 1 | |
| Matheus | 3 | |
| Ana | 4 | ← Q1 |
| João | 5 | |
| José | 5 | ← Q2 |
| Aline | 8 | |
| Maria | 9 | |
| Nasser | 9 | ← Q3 |
| Julia | 10 | |

Medidas de posição - Mínimo e Máximo

Mínimo: é o menor valor da série

Máximo: é o maior valor da série

Medidas de posição - Quartis

| | |
|---------|-------------|
| José | 5 |
| Maria | 9 |
| Ana | 4 |
| João | 5 |
| Julia | 10 ← Máximo |
| Paulo | 1 ← Mínimo |
| Matheus | 3 |
| Nasser | 9 |
| Aline | 8 |

Medidas de dispersão

- Amplitude
- Desvio Médio
- Variância
- Desvio Padrão

Medidas de dispersão

| Grupo A | Notas |
|---------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |

Medidas de dispersão - Amplitude

Amplitude é a diferença entre o maior(máximo) e o menor(mínimo) valor.

Medidas de dispersão - Amplitude

| Grupo A | Notas |
|---------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |

Medidas de dispersão - Amplitude

| Grupo A | Notas |
|---------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |

← Máximo

← Mínimo

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |

← Mínimo

← Máximo

Medidas de dispersão - Amplitude

| Grupo A | Notas |
|-----------|------------|
| José | 6 ← Máximo |
| João | 6 |
| Paulo | 5 ← Mínimo |
| Matheus | 5 |
| Média | 5.5 |
| Amplitude | $6-5=1$ |

| Grupo B | Notas |
|-----------|-------------|
| Maria | 9 |
| Ana | 1 ← Mínimo |
| Julia | 10 ← Máximo |
| Aline | 2 |
| Média | 5.5 |
| Amplitude | $10-1=9$ |

Medidas de dispersão - Amplitude

| Grupo A | Notas |
|-----------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Amplitude | 1 |

| Grupo B | Notas |
|-----------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |
| Amplitude | 9 |

Medidas de dispersão - Desvio médio

Desvio médio é a média do módulo (abs) dos desvios da média ou mediana.

Medidas de dispersão - Desvio médio

| Grupo A | Notas |
|---------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |

Medidas de dispersão - Desvio médio

| Grupo A | Notas | Desvio | Módulo |
|---------|-------|--------------|-----------------------|
| José | 6 | $6-5.5=0.5$ | 0.5 |
| João | 6 | $6-5.5=0.5$ | 0.5 |
| Paulo | 5 | $5-5.5=-0.5$ | 0.5 |
| Matheus | 5 | $5-5.5=-0.5$ | 0.5 |
| Média | 5.5 | | $(0.5+0.5+0.5+0.5)/4$ |

Medidas de dispersão - Desvio médio

| Grupo A | Notas | Desvio | Módulo |
|---------|-------|------------------|-------------|
| José | 6 | $6 - 5.5 = 0.5$ | 0.5 |
| João | 6 | $6 - 5.5 = 0.5$ | 0.5 |
| Paulo | 5 | $5 - 5.5 = -0.5$ | 0.5 |
| Matheus | 5 | $5 - 5.5 = -0.5$ | 0.5 |
| Média | 5.5 | | $2/4 = 0.5$ |

Medidas de dispersão - Desvio médio

| Grupo A | Notas |
|-----------------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Desvio Médio | 0.5 |

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |

Medidas de dispersão - Desvio médio

| Grupo B | Notas | Desvio | Módulo |
|---------|-------|------------------|-------------------------------|
| Maria | 9 | $9 - 5.5 = 3.5$ | 3.5 |
| Ana | 1 | $1 - 5.5 = -4.5$ | 4.5 |
| Julia | 10 | $10 - 5.5 = 4.5$ | 4.5 |
| Aline | 2 | $2 - 5.5 = -3.5$ | 3.5 |
| Média | 5.5 | | $(3.5 + 4.5 + 4.5 + 3.5) / 4$ |

Medidas de dispersão - Desvio médio

| Grupo B | Notas | Desvio | Módulo |
|---------|-------|------------------|------------|
| Maria | 9 | $9 - 5.5 = 3.5$ | 3.5 |
| Ana | 1 | $1 - 5.5 = -4.5$ | 4.5 |
| Julia | 10 | $10 - 5.5 = 4.5$ | 4.5 |
| Aline | 2 | $2 - 5.5 = -3.5$ | 3.5 |
| Média | 5.5 | | $16/4 = 4$ |

Medidas de dispersão - Desvio médio

| Grupo A | Notas |
|--------------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Desvio Médio | 0.5 |

| Grupo B | Notas |
|--------------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |
| Desvio Médio | 4 |

Medidas de dispersão - Variância

O desvio médio é baseado em uma operação não algébrica (módulo) dificultando o estudo de suas propriedades. Como alternativa utilizamos a variância.

Variância é a soma dos quadrados dos desvios.

Medidas de dispersão - Variância

| Grupo A | Notas |
|---------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |

Medidas de dispersão - Variância

| Grupo A | Notas | Desvio | Desvio ao Quadrado |
|---------|-------|--------------|---------------------------|
| José | 6 | $6-5.5=0.5$ | 0.25 |
| João | 6 | $6-5.5=0.5$ | 0.25 |
| Paulo | 5 | $5-5.5=-0.5$ | 0.25 |
| Matheus | 5 | $5-5.5=-0.5$ | 0.25 |
| Média | 5.5 | | $(0.25+0.25+0.25+0.25)/4$ |

Medidas de dispersão - Variância

| Grupo A | Notas | Desvio | Desvio ao Quadrado |
|---------|-------|------------------|--------------------|
| José | 6 | $6 - 5.5 = 0.5$ | 0.25 |
| João | 6 | $6 - 5.5 = 0.5$ | 0.25 |
| Paulo | 5 | $5 - 5.5 = -0.5$ | 0.25 |
| Matheus | 5 | $5 - 5.5 = -0.5$ | 0.25 |
| Média | 5.5 | | $1/4 = 0.25$ |

Medidas de dispersão - Variância

| Grupo A | Notas |
|-----------------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Desvio Médio | 0.5 |
| Variância | 0.25 |

| Grupo B | Notas |
|---------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |

Medidas de dispersão - Variância

| Grupo B | Notas | Desvio | Desvio ao Quadrado |
|---------|-------|------------------|---------------------------------------|
| Maria | 9 | $9 - 5.5 = 3.5$ | 12.25 |
| Ana | 1 | $1 - 5.5 = -4.5$ | 16.25 |
| Julia | 10 | $10 - 5.5 = 4.5$ | 16.25 |
| Aline | 2 | $2 - 5.5 = -3.5$ | 12.25 |
| Média | 5.5 | | $(12.25 + 16.25 + 16.25 + 12.25) / 4$ |

Medidas de dispersão - Variância

| Grupo B | Notas | Desvio | Desvio ao Quadrado |
|---------|-------|------------------|--------------------|
| Maria | 9 | $9 - 5.5 = 3.5$ | 12.25 |
| Ana | 1 | $1 - 5.5 = -4.5$ | 16.25 |
| Julia | 10 | $10 - 5.5 = 4.5$ | 16.25 |
| Aline | 2 | $2 - 5.5 = -3.5$ | 12.25 |
| Média | 5.5 | | $57/4 = 14.25$ |

Medidas de dispersão - Variância

| Grupo A | Notas |
|--------------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Desvio Médio | 0.5 |
| Variância | 0.25 |

| Grupo B | Notas |
|--------------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |
| Desvio Médio | 4 |
| Variância | 14.25 |

Medidas de dispersão - Desvio Padrão

Para manter a mesma unidade de medida dos dados originais, definimos o **desvio padrão** como a raiz da variância.

Medidas de dispersão - Desvio Padrão

| Grupo A | Notas |
|------------------|---------------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Desvio Médio | 0.5 |
| Variância | 0.25 |
| Desvio Padrão | $\sqrt{0.25}$ |

| Grupo B | Notas |
|------------------|----------------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |
| Desvio Médio | 4 |
| Variância | 14.25 |
| Desvio Padrão | $\sqrt{14.25}$ |

Medidas de dispersão - Desvio Padrão

| Grupo A | Notas |
|------------------|-------|
| José | 6 |
| João | 6 |
| Paulo | 5 |
| Matheus | 5 |
| Média | 5.5 |
| Desvio Médio | 0.5 |
| Variância | 0.25 |
| Desvio Padrão | 0.5 |

| Grupo B | Notas |
|------------------|-------|
| Maria | 9 |
| Ana | 1 |
| Julia | 10 |
| Aline | 2 |
| Média | 5.5 |
| Desvio Médio | 4 |
| Variância | 14.25 |
| Desvio Padrão | 3.77 |