

## MÉDIA ARITMÉTICA

É a soma de todos os valores em um conjunto de dados dividida pelo número total de valores. É a média mais comumente usada e é representada pela fórmula:

### FÓRMULA

$$\overline{X} = \frac{x_1 + x_2 + \dots + x_N}{N}$$

### EXEMPLO:

**SALÁRIO 1:** 1.200

**SALÁRIO 2:** 2.500

**SALÁRIO 3:** 1.800

$$\textbf{MÉDIA} = \frac{1.200 + 2.500 + 1.800}{3}$$

$$\textbf{MÉDIA} = 1.833$$

## MÉDIA PONDERADA

É semelhante à média aritmética, mas atribui pesos diferentes aos valores, dependendo da sua importância relativa. A fórmula para a média ponderada é:

## FÓRMULA

$$\overline{X} = \frac{w_1 x_1 + w_2 x_2 + \dots w_N x_N}{w_1 + w_2 + w_n}$$

## EXEMPLO:

**NOTA 1: 5 PESO 1: 5**

**NOTA 2: 7 PESO 2: 2**

**NOTA 3: 8 PESO 3: 3**

$$\text{MÉDIA} = \frac{5*5 + 7*2 + 8*3}{5 + 2 + 3}$$

**MÉDIA PONDERADA = 6**

## MEDIANA

A mediana é uma medida estatística que representa o valor central de um conjunto de dados ordenados. Em outras palavras, é o valor que divide o conjunto ao meio, com metade dos valores sendo maiores que a mediana e a outra metade sendo menores.

### EXEMPLO:

Dada a série de valores:

**5,7,12,15,18**

O **valor central é 12** então ele será a mediana.

Mas e para conjuntos com números pares?

Dada a série de valores:

**5,7,12,15,18,20**

Realizamos a média dos dois valores centrais:

$$\text{MÉDIA} = \frac{12+15}{2} = 13.5$$

## MODA

Ela representa o valor que ocorre com mais frequência em um conjunto de dados. Em outras palavras, é o valor que possui a maior frequência.

## EXEMPLO:

Dada a série de valores:

**1,1,3,3,3,4,5,5,6,6**

Identificamos o valor que mais aparece, nesse caso a maior frequência é o 3, essa seria nossa moda.

## Conjunto amodal, Bimodal e Multimodal:

Caso não exista nenhum valor que tenha uma frequência maior, chamamos o conjunto de amodal.

Caso tenhamos 2 valores que se repetem mais então entendemos como bimodal.

## DESVIO PADRÃO

O desvio padrão é uma medida de dispersão que indica o quanto os valores de um conjunto de dados estão dispersos em relação à média. Em outras palavras, ele quantifica a variabilidade ou a dispersão dos valores em torno da média.

## FÓRMULA

$$D_p = \sqrt{\frac{\sum (X_1 - \bar{X})^2}{(N - 1)}}$$

## DESVIO PADRÃO – EXEMPLO PRÁTICO

Dada a série de idades:

**25,30,35,40,45**

1) Calculamos a média:

$$\text{MÉDIA} = \frac{20+30+35+40+45}{5} = 35$$

2) Aplicando na fórmula do desvio padrão:

$$\text{DESVIO PADRÃO} = \sqrt{\frac{(25-35)^2 + (30-35)^2 + (35-35)^2 + (40-35)^2 + (45-35)^2}{5-1}} = 7.91$$

$$D_p = \sqrt{\frac{\sum (X_i - \bar{X})^2}{(N - 1)}}$$

## VARIÂNCIA

A variância é uma medida estatística que indica a dispersão ou variabilidade dos valores de um conjunto de dados em relação à sua média. Em outras palavras, a variância mede o quão distantes os valores individuais estão da média.

## FÓRMULA

$$\text{VAR} = \left( \sqrt{\frac{\sum (X_1 - \bar{X})^2}{(N - 1)}} \right)^2$$

## VARIÂNCIA X DESVIO PADRÃO

- O desvio mede a dispersão dos valores em torno da média e é uma medida mais intuitiva da variabilidade dos dados, pois está na mesma unidade de medida dos dados originais
- A variância é expressa em unidades ao quadrado. Portanto, se os dados estiverem em uma unidade específica (por exemplo, metros, dólares, etc.), a variância será medida em unidades ao quadrado dessa unidade original.
- Aplicabilidade: Geralmente a variância é mais utilizada em modelos de machine learning, como por exemplo em modelos de regressão. Já o desvio por ser mais intuitivo é muito utilizado em análises.

**VARIÂNCIA =  
DESVIO AO  
QUADRADO**



## PROBABILIDADE

Probabilidade é uma medida quantitativa da incerteza associada a um evento ou resultado. Ela descreve a chance ou a probabilidade de que um evento específico ocorra em um experimento ou situação. Na teoria das probabilidades, a probabilidade de um evento é geralmente expressa como um número entre 0 e 1.

## FÓRMULA

### **Probabilidade**

Número de Resultados Favoráveis

Número Total de Resultados Possíveis

## EXEMPLO:

### Definição do Evento:

- **Evento:** Um cliente realiza uma compra após visitar o site.
- **Não evento:** Um cliente não realiza uma compra após visitar o site.

**Cenário:** Digamos que, em um determinado período de tempo, 500 clientes tenham visitado o site e 50 deles fizeram uma compra.

**Para calcular a probabilidade de um cliente fazer uma compra após visitar o site, você divide o número de compras pelo número total de visitas**

$$\frac{50}{500} = 0,1 \text{ ou } 10\% \text{ de probabilidade}$$