

# Python

Fundamentos de **BIG DATA** e **DATA ANALYTICS**

# Matemática

## Estatística

- > Definição
- > Funções Interativas
  - **Média**
  - **Moda**
  - **Variância**
  - **Desvio Padrão**
- > Estatística Descritiva
- > Tabulação de Dados
- > Gráficos
  - **Linha**
  - **Pizza**
  - **Dispersão**

# Matemática Estatística | Definição

---

## **O que é Matemática Estatística ?**

É a área da matemática que estuda a coleta, registro, organização e análise dos dados de uma pesquisa.

# Matemática Estatística | Funções Interativas - Média

---

## O que é Média ?

É a medida central de tendência de uma série de dados.

Formula:

$$X = \frac{(X1 + X2 + X3 + \dots + Xn)}{n}$$

**n** => Número de elementos

# Matemática Estatística | Funções Interativas - Média

---

## Exemplo:

Idade dos funcionários do departamento de recursos humanos de uma empresa: {28, 30, 29, 31, 32, 33, 34}

Formula:

$$X = \frac{(28+30+29+31+32+33+34)}{7} \Rightarrow X = 31$$

# Matemática Estatística | Funções Interativas - Moda

---

## O que é Moda ?

É o valor que mais se repete em uma série de dados.

Exemplo :

*Lista = {35, 37, 36, 34, 38, 35, 37, 37, 33, 36, 38, 37, 35, 37, 34, 33, 37, 36, 35, 38, 36, 35, 36, 37, 38, 39, 37, 37, 36, 37, 33, 37, 35, 37, 39}*

*Quantas vezes o valor 33 aparece ? R: 3 vezes*

*Quantas vezes o valor 34 aparece ? R: 2 vezes*

*Quantas vezes o valor 35 aparece ? R: 6 vezes*

*Quantas vezes o valor 36 aparece ? R: 6 vezes*

*Quantas vezes o valor 37 aparece ? R: 12 vezes*

*Quantas vezes o valor 38 aparece ? R: 4 vezes*

*Quantas vezes o valor 39 aparece ? R: 2 vezes*

# Matemática Estatística | Funções Interativas - Variância

---

## O que é Variância ?

É a média da somatória dos desvios ao quadrado dos dados.

Formula:

$$V = \sqrt{\frac{\sum (X_i - X)^2}{n - 1}}$$

$V$  => Variância

$X_i$  => Valor que está sendo analisado

$X$  => Média aritmética dos elementos

$n$  => Número de elementos

# Matemática Estatística | Funções Interativas - Variância

## Exemplo :

*Lista* = {3, 7, 6, 5, 4}

- Média =  $\frac{3+7+6+5+4}{5} \Rightarrow 25/5 = 5$
- Desvio de cada item em relação a média:  
(média-valorItem) = Dv

$$5-3 = 2$$

$$5-7 = -2$$

$$5-6 = -1$$

$$5-5 = 0$$

$$5-4 = 1$$

- Variância :

$$V = \frac{2^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2}{5} \Rightarrow \frac{4+4+1+0+1}{5} \Rightarrow$$

$$\frac{10}{5} \Rightarrow$$

$$V = 2$$



## O que é Desvio Padrão ?

É a raiz quadrada da variância

Formula:

$$DP = \sqrt{V}$$

DP => Desvio Padrão      V => Variância

# Matemática Estatística | Funções Interativas – Desvio Padrão

---

**Exemplo:**

$$DP = \sqrt{2}$$

$$**DP = 1,4142**$$

## **O que é Estatística Descritiva ?**

É o ramo da estatística que visa sumarizar e descrever qualquer conjunto de dados.

# Matemática Estatística | **Tabulação de Dados**

---

## **O que é Tabulação de Dados ?**

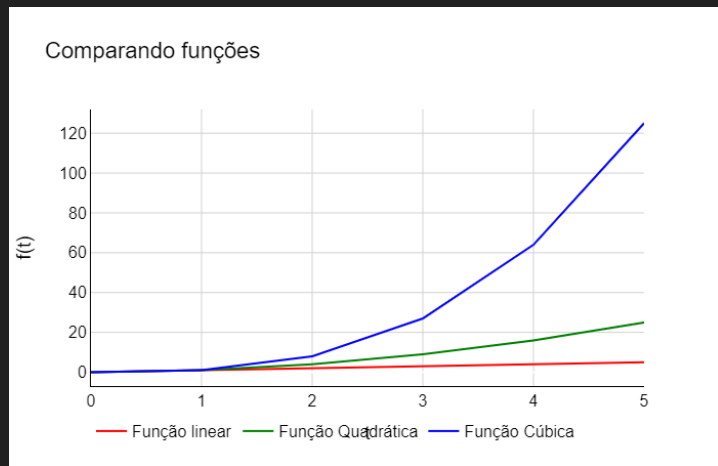
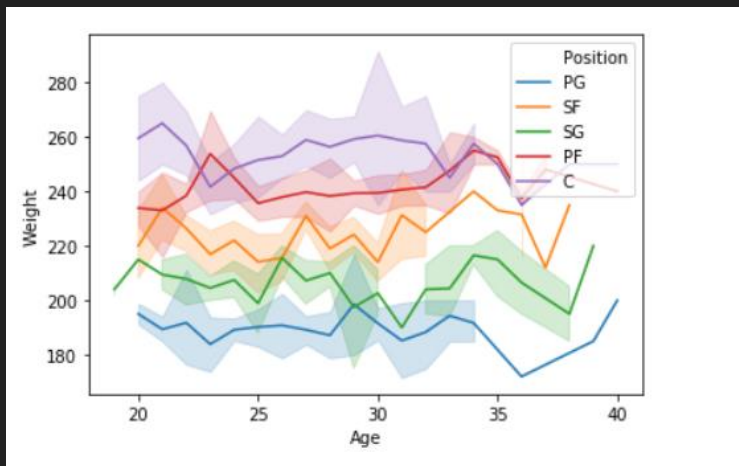
É a transformação de todas as informações coletadas em material que possa ser analisado.

## O que são Gráficos ?

É a tentativa de expressar visualmente valores numéricos, de maneira diferente, facilitando a compreensão.

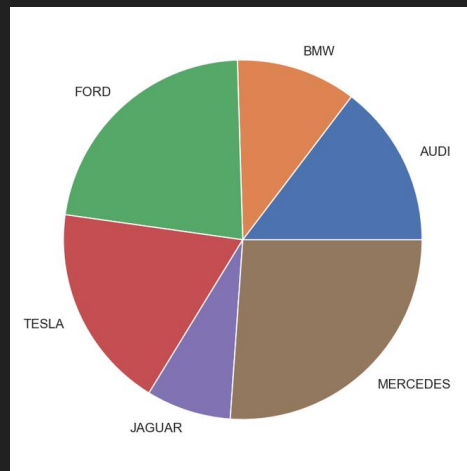
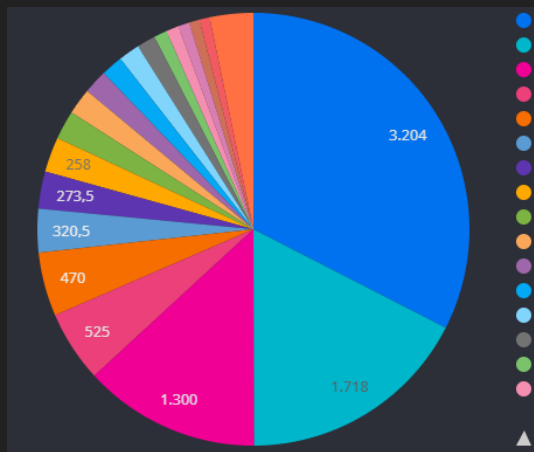
# Matemática Estatística | Gráficos - Linha

**O Gráfico de Linhas:** Mostram as informações como uma série de pontos de dados que estão conectadas por segmentos de linha reta.



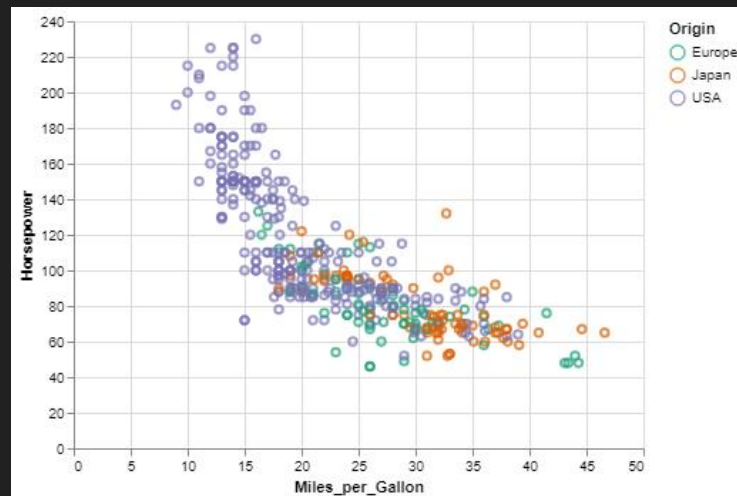
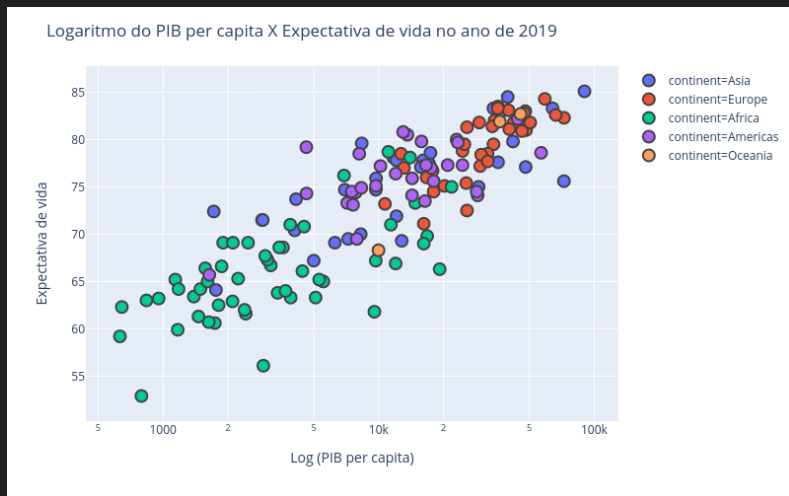
# Matemática Estatística | Gráficos - Pizza

**O Gráfico de Setores:** É uma visualização que representa um valor relativo de cada categoria estabelecida em relação a um todo



# Matemática Estatística | Gráficos - Dispersão

**O Gráfico de Dispersão:** É utilizado para pontuar dados em um eixo vertical e horizontal com a intenção de exibir quanto uma variável é afetada por outra





Obrigado !

---