

Previsão estatística com ciência de dados (desplugado e colab)

Prof. Diego Vergaças

Objetivo

Demonstra como um valor futuro a partir de dados históricos simples, utilizando raciocínio estatístico e conceitos da modelagem de dados.

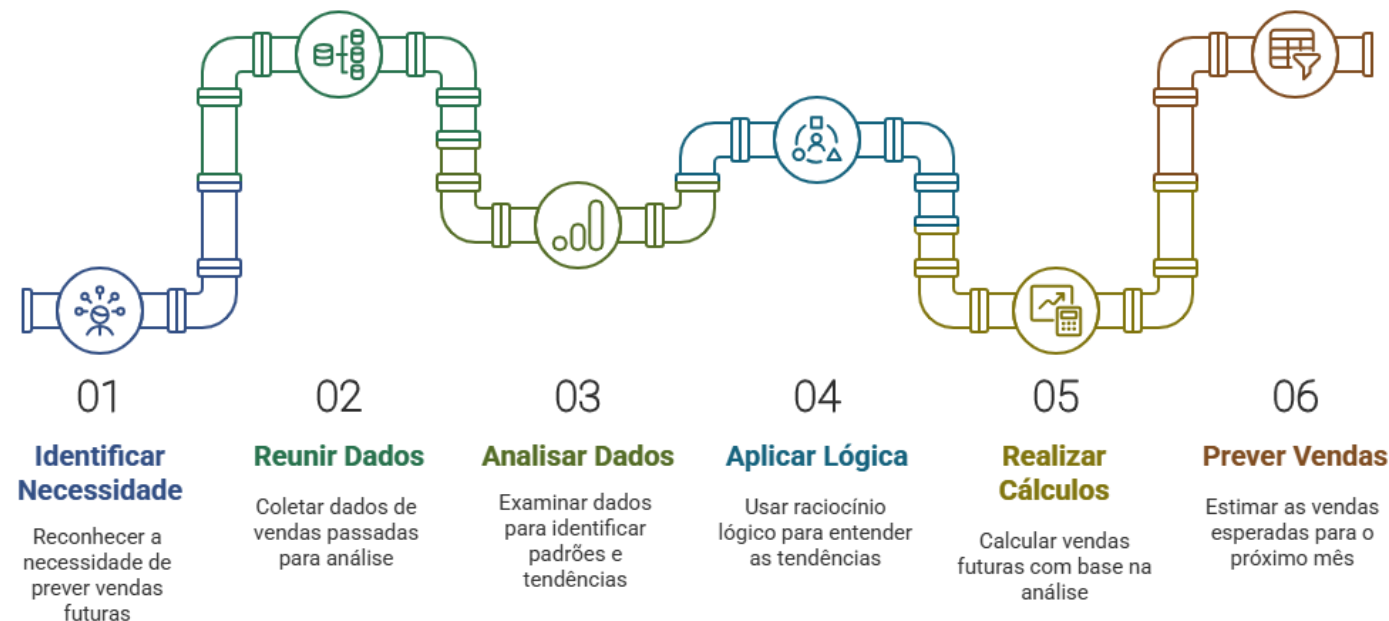
Previsão de Valores Futuros



Aplicação – Previsão de vendas

Imagine que você trabalha em uma loja e precisa prever quantos produtos vai vender no mês que vem. Hoje vamos fazer uma simulação simples de previsão estatística — sem computador, só com lógica, matemática e lápis na mão.

Simulação de Previsão Estatística



Dados

Mês	Vendas (unidades)
Janeiro	100
Fevereiro	120
Março	110
Abril	150
Maio	140
Junho	?

Quantas unidades vocês acham que vamos vender em junho? Como podemos prever isso de forma fundamentada?

Cálculos para linha de tendência

Utilizando a equação da reta:

$$y = ax + b$$

Onde,

Coeficiente angular

$$a = \frac{n \cdot \sum(xy) - \sum x \cdot \sum y}{n \cdot \sum(x^2) - (\sum x)^2}$$

Coeficiente angular

$$b = \frac{\sum y - a \cdot \sum x}{n}$$

x	y	x^2	$x \cdot y$
0	100	0	0
1	120	1	120
2	110	4	220
3	150	9	450
4	140	16	560

Fazendo as somatórias

$$\sum x = 0 + 1 + 2 + 3 + 4 = 10$$

$$\sum y = 100 + 120 + 110 + 150 + 140 = 620$$

$$\sum x^2 = 0 + 1 + 4 + 9 + 16 = 30$$

$$\sum x \cdot y = 0 + 120 + 220 + 450 + 560 = 1350$$

$$n = 5$$

x	y	x^2	$x \cdot y$
0	100	0	0
1	120	1	120
2	110	4	220
3	150	9	450
4	140	16	560

Coeficiente angular a :

$$a = \frac{n \cdot \sum(xy) - \sum x \cdot \sum y}{n \cdot \sum(x^2) - (\sum x)^2}$$

$$a = \frac{5 \cdot 1350 - 10 \cdot 620}{5 \cdot 30 - 10^2}$$

$$a = \frac{6750 - 6200}{150 - 100}$$

$$a = \frac{550}{50}$$

$$a = 11$$

$$\sum x = 10$$

$$\sum y = 620$$

$$\sum x^2 = 30$$

$$\sum x \cdot y = 1350$$

$$n = 5$$

Coeficiente angular b :

$$b = \frac{\sum y - a \cdot \sum x}{n}$$

$$b = \frac{620 - 11 \cdot 10}{5}$$

$$b = \frac{620 - 110}{5}$$

$$b = \frac{510}{5}$$

$$b = 102$$

$$\sum x = 10$$

$$\sum y = 620$$

$$\sum x^2 = 30$$

$$\sum x \cdot y = 1350$$

$$n = 5$$

Aplicando a equação da reta:

Equação da tendência é:

$$y = ax + b$$

$$y = 11 \cdot x + 102$$

$$b = 102$$

$$a = 11$$

$$x = 5$$

Nós vamos usar a equação da reta para prever o valor do próximo mês, junho.

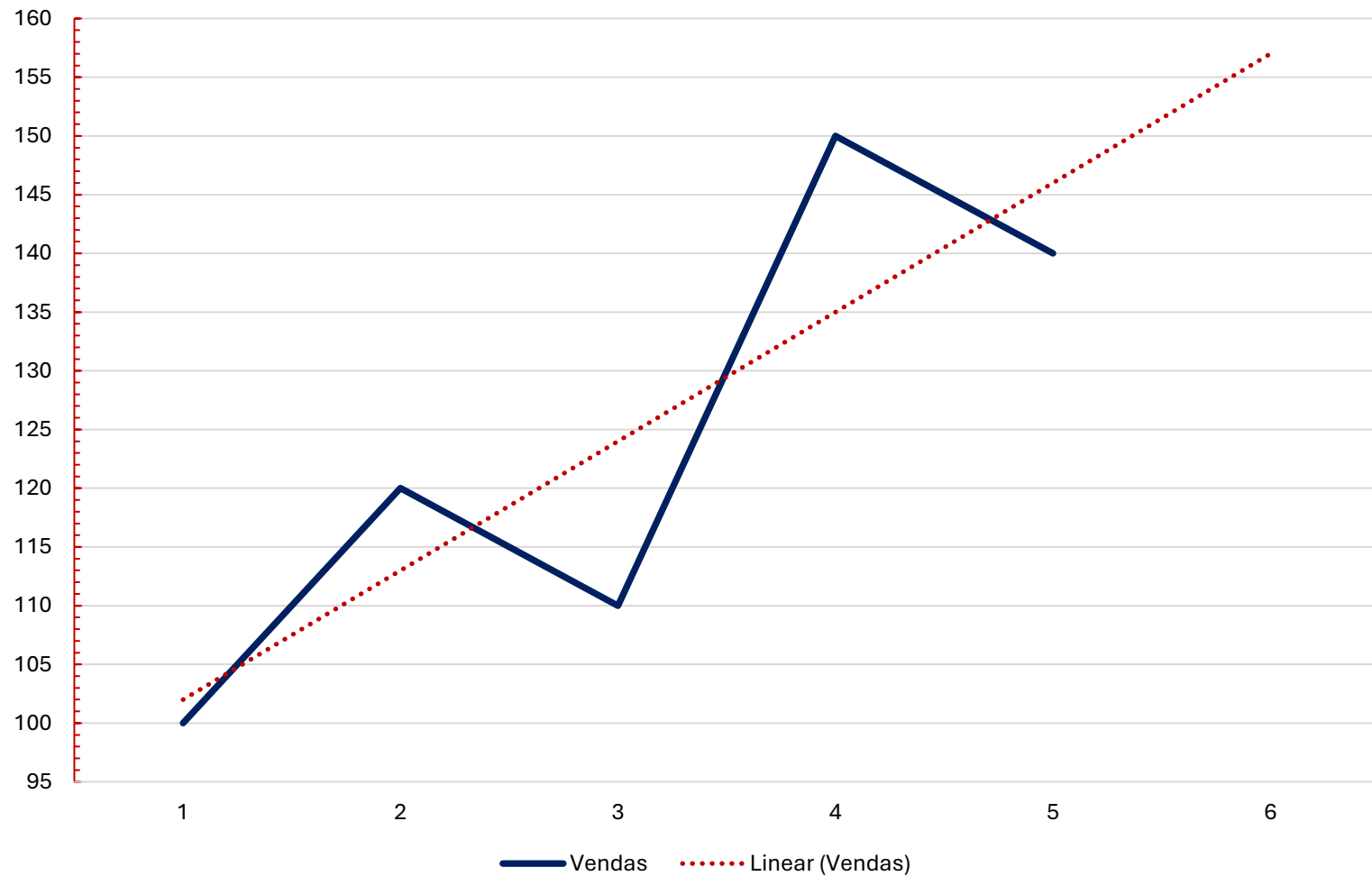
Para isso, **associamos a cada mês um valor de x** , representando o número de meses desde o início da nossa observação.

Como temos os dados de **5 meses completos** — janeiro, fevereiro, março, abril e maio — o próximo mês, junho, será o mês número 5 da nossa contagem.

$$y = 11 \cdot 5 + 102$$

$$y = 55 + 102$$

$$y = 157$$



Mês	Vendas (unidades)
Janeiro	100
Fevereiro	120
Março	110
Abril	150
Maio	140
Junho	157

Código em python – colab google

```
[1] import pandas as pd
```

```
[2] # Dados fictícios de vendas por mês
dados = {
    'Mês': ['Jan', 'Fev', 'Mar', 'Abr', 'Mai'],
    'Vendas': [100, 120, 110, 150, 140]
}
```

```
[3] df = pd.DataFrame(dados)
df
```

	Mês	Vendas
0	Jan	100
1	Fev	120
2	Mar	110
3	Abr	150
4	Mai	140

```
[4] import matplotlib.pyplot as plt
```

```
df['Índice'] = range(1, len(df) + 1) # Representação numérica dos meses
```

```
plt.plot(df['Índice'], df['Vendas'], marker='o')
```

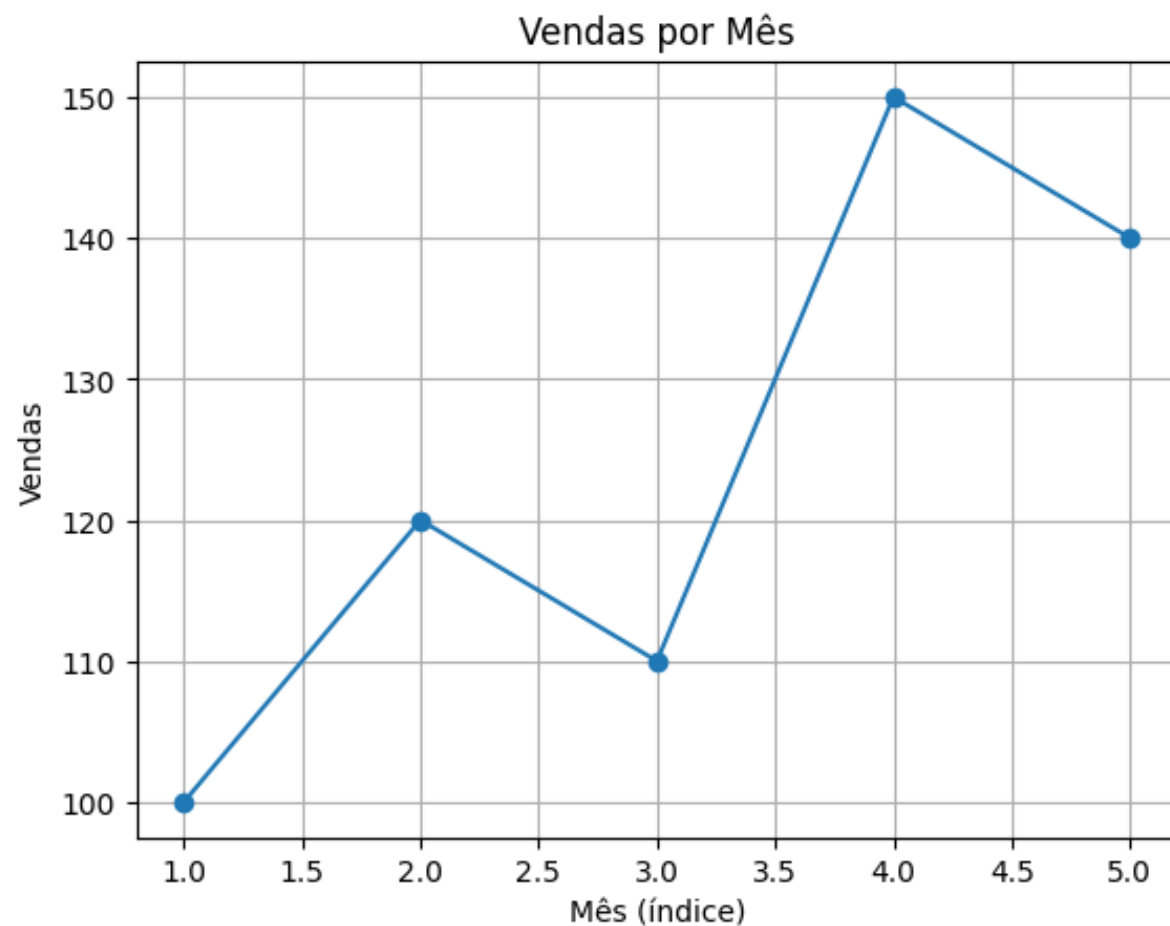
```
plt.title('Vendas por Mês')
```

```
plt.xlabel('Mês (índice)')
```

```
plt.ylabel('Vendas')
```

```
plt.grid(True)
```

```
plt.show()
```



```
✓ [8] import numpy as np
0s

✓ [9] # Variáveis independentes e dependentes
0s X = df['Índice']
y = df['Vendas']

✓ [▶] # Cálculo dos coeficientes da reta (y = a*x + b)
0s # Tendência linear --> np.polyfit()
a, b = np.polyfit(X, y, 1)

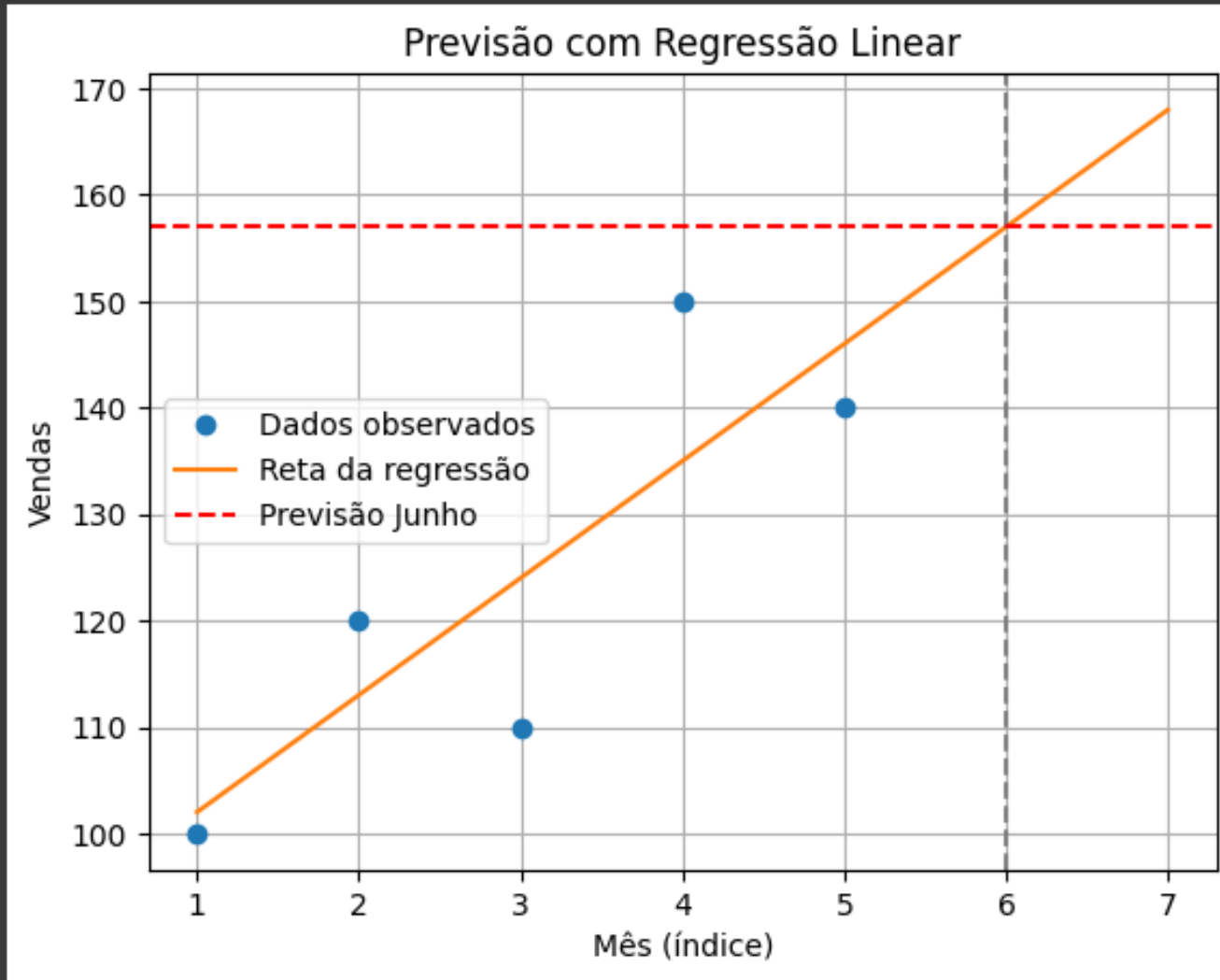
print(f"Equação da reta: y = {a:.2f}x + {b:.2f}")

⇒ Equação da reta: y = 11.00x + 91.00
```

```
✓ [▶] # Previsão para o mês 6 (junho)
0s mes_6 = 6
previsao = a * mes_6 + b
print(f"Previsão de vendas para junho: {previsao:.2f}")

⇒ Previsão de vendas para junho: 157.00
```

```
[ ] # Plot original + previsão
plt.plot(X, y, 'o', label='Dados observados')
plt.plot(range(1, 8), [a * x + b for x in range(1, 8)], '-', label='Reta da regressão')
plt.axvline(x=6, color='gray', linestyle='--')
plt.axhline(y=previsao, color='red', linestyle='--', label='Previsão Junho')
plt.legend()
plt.xlabel('Mês (índice)')
plt.ylabel('Vendas')
plt.title('Previsão com Regressão Linear')
plt.grid(True)
plt.show()
```



"Ciência de Dados não é prever o futuro com magia, mas com lógica, estatística e curiosidade."

Hoje vimos como dados simples podem revelar padrões e ajudar na tomada de decisões.

Com ferramentas como **Pandas**, **NumPy** e **modelos estatísticos**, transformamos informação em conhecimento, e conhecimento em ação.