

# Modelos de regressão

Professor: Gabriel  
Oliveira Assunção

# Dados



## Features

- Dados para explicar o comportamento do target
- Conhecido também como variáveis explicativas.



## Target

- Dado alvo, usualmente o que queremos prever.
- Conhecido também como variável resposta.

# Modelo de Regressão

## Inputs

### Features

Conjunto de dado conhecido



### Equação

$$Y = a + bX$$



## Outputs

### Target

Predição do target,  
dado desconhecido



# Modelo de Regressão

## Inputs

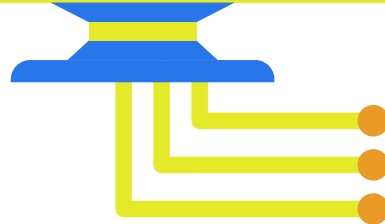
### Features

Conjunto de dado conhecido



### Equação

$$Y = a + bX + cZ + \dots$$



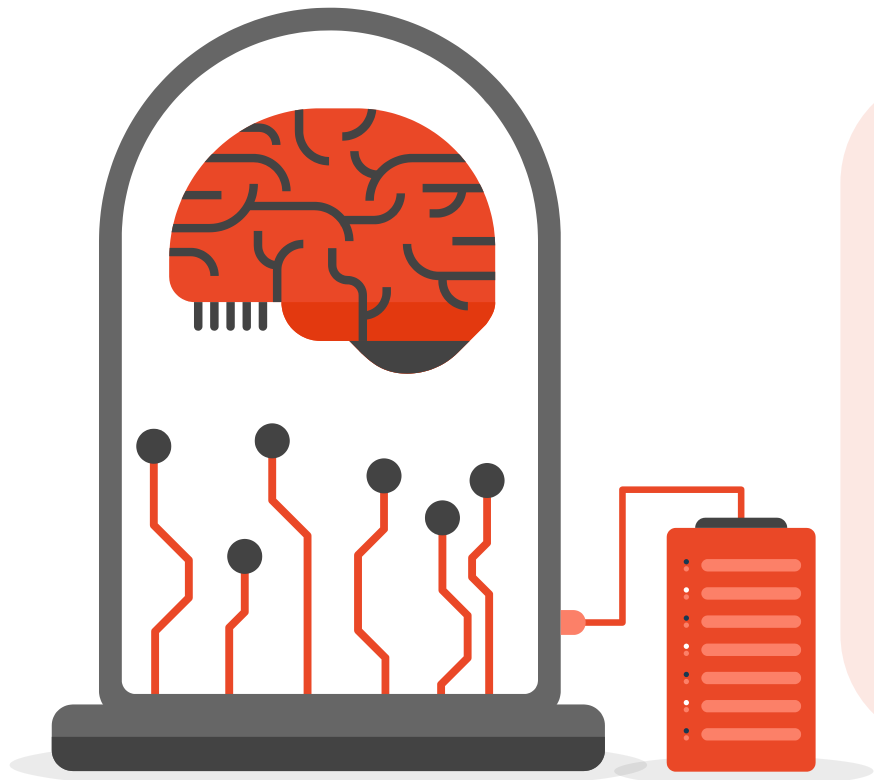
## Outputs

### Target

Predição do target,  
dado desconhecido



# Função de perda



O motor para a construção do modelo é conhecido como função de perda.

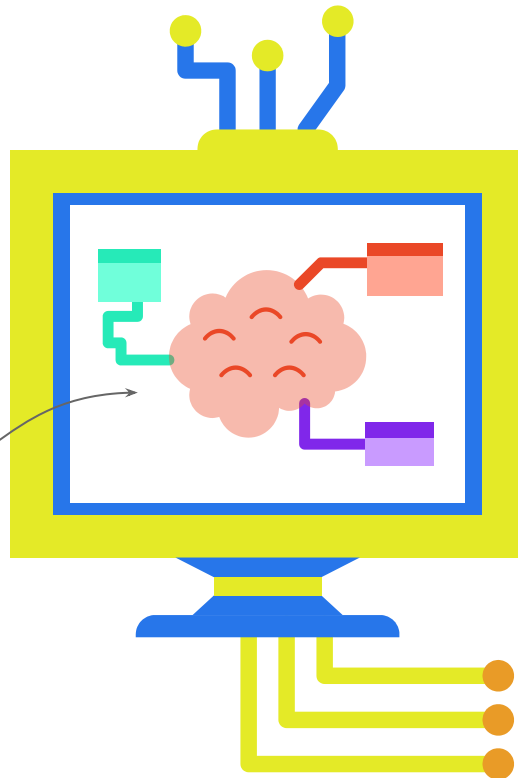
**Função de perda:** é uma função matemática que buscamos otimizar da forma que nosso modelo gere boas previsões.

# Modelo de regressão

## Inputs

Dados sobre o imóvel

- Número de quartos
- Tamanho
- Número de banheiros

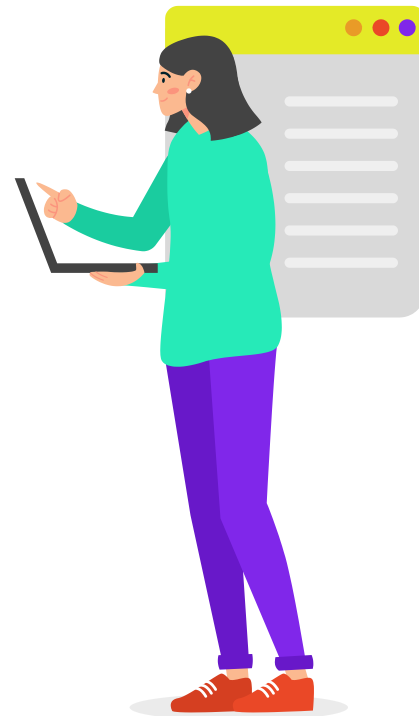
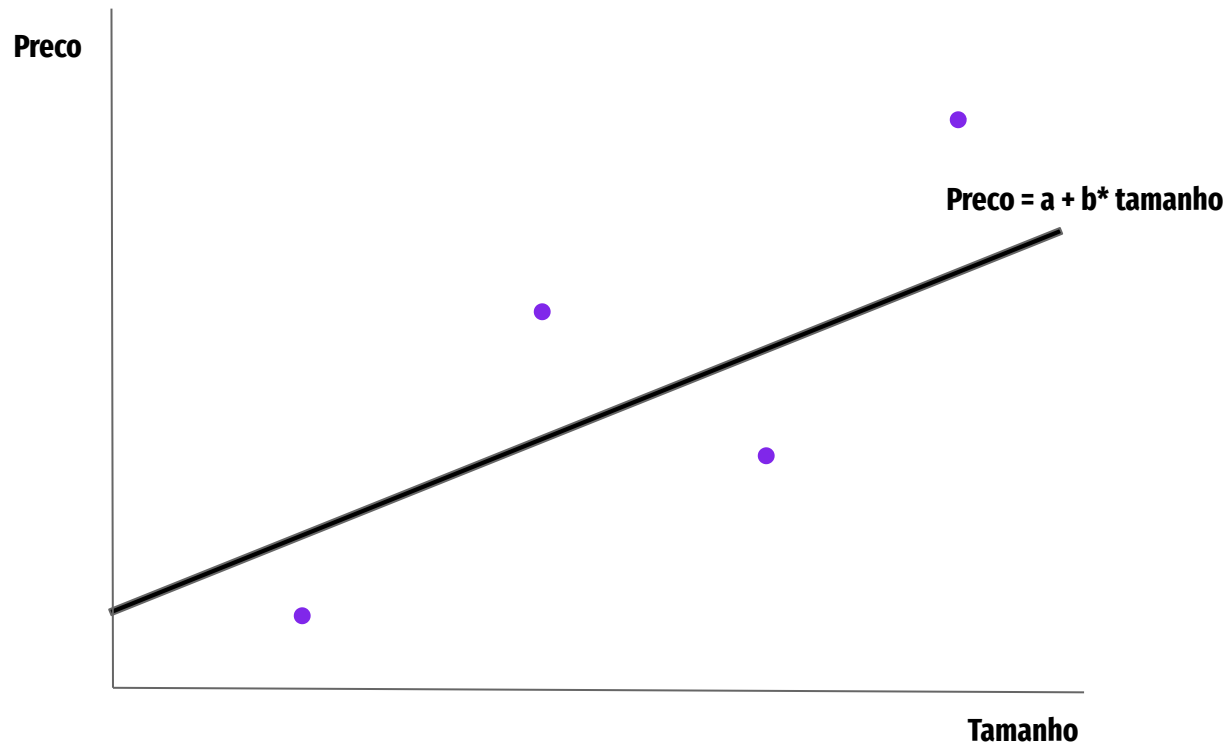


## Outputs

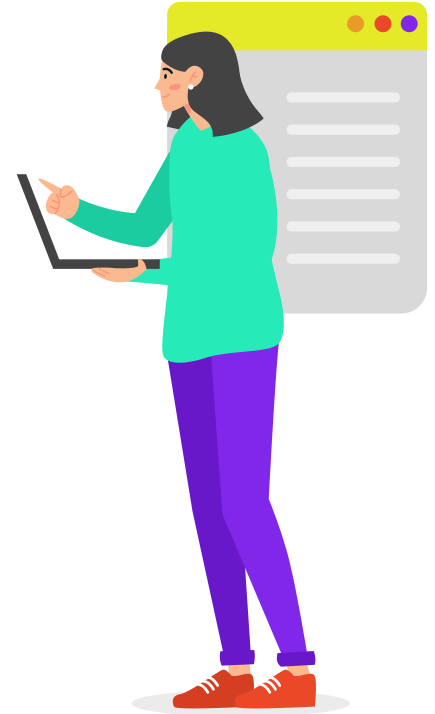
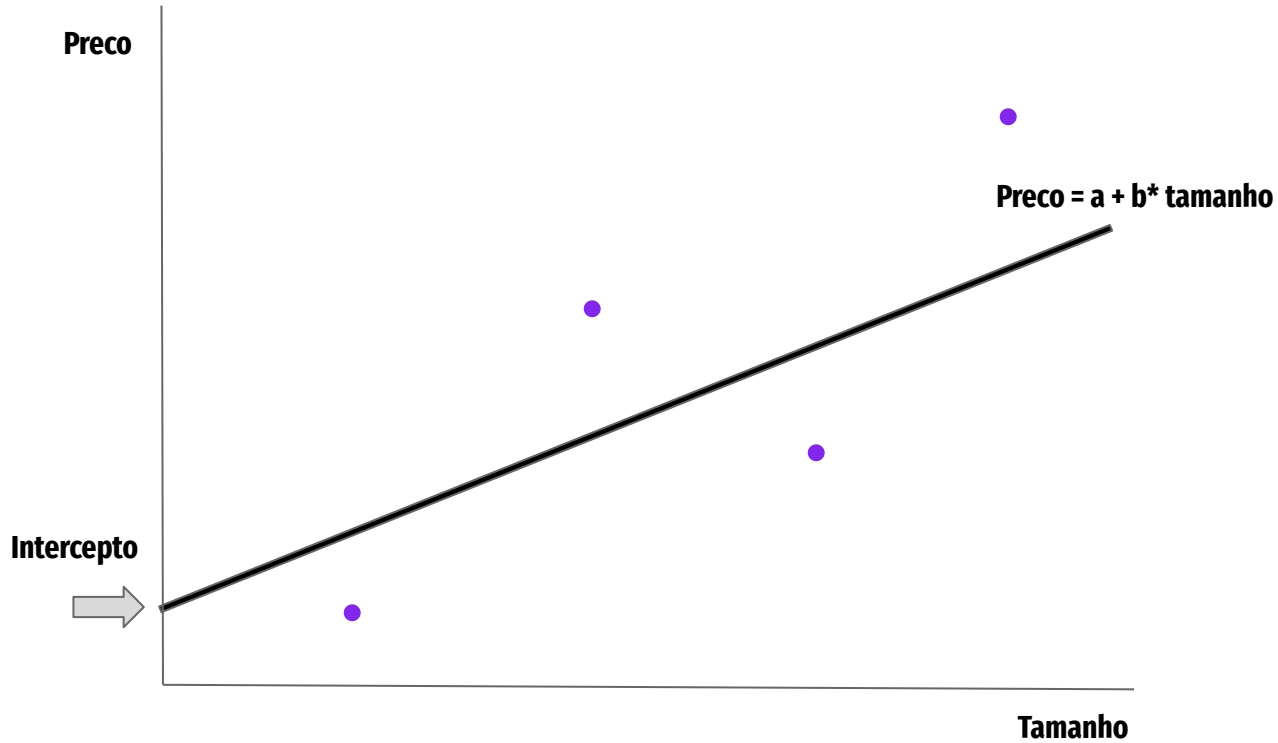
Preço do imóvel

- Constante/  
Intercepto
- Coeficientes

# Análise da reta

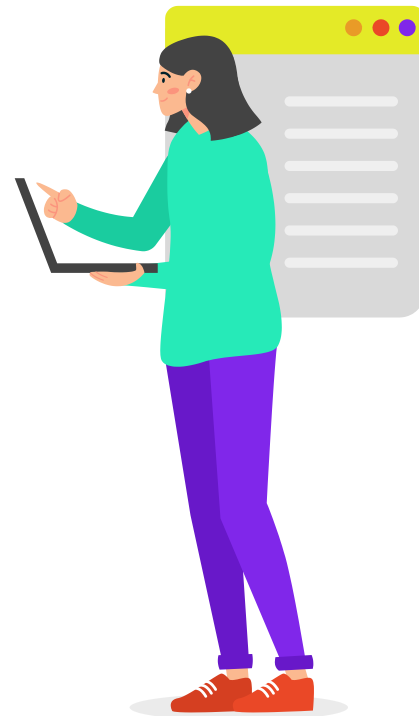
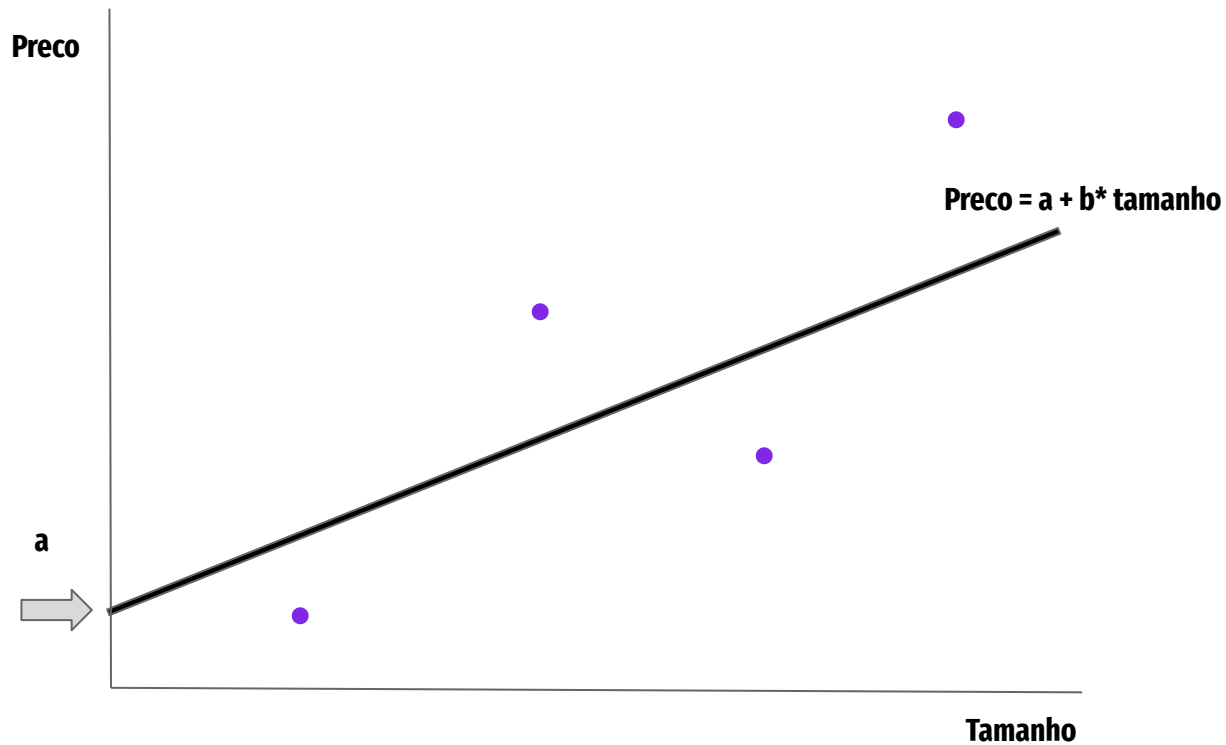


# Análise da reta

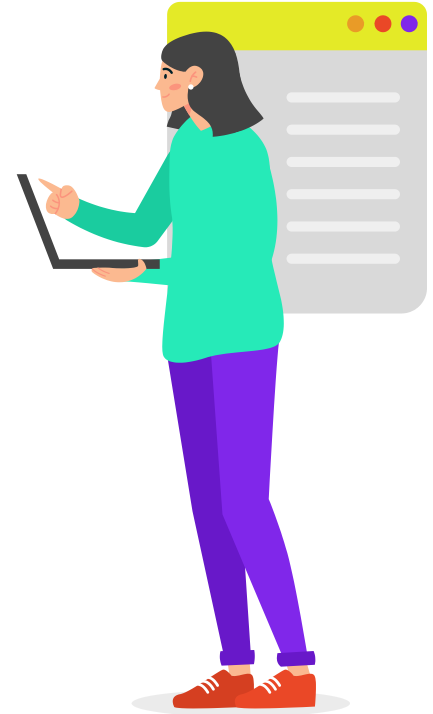
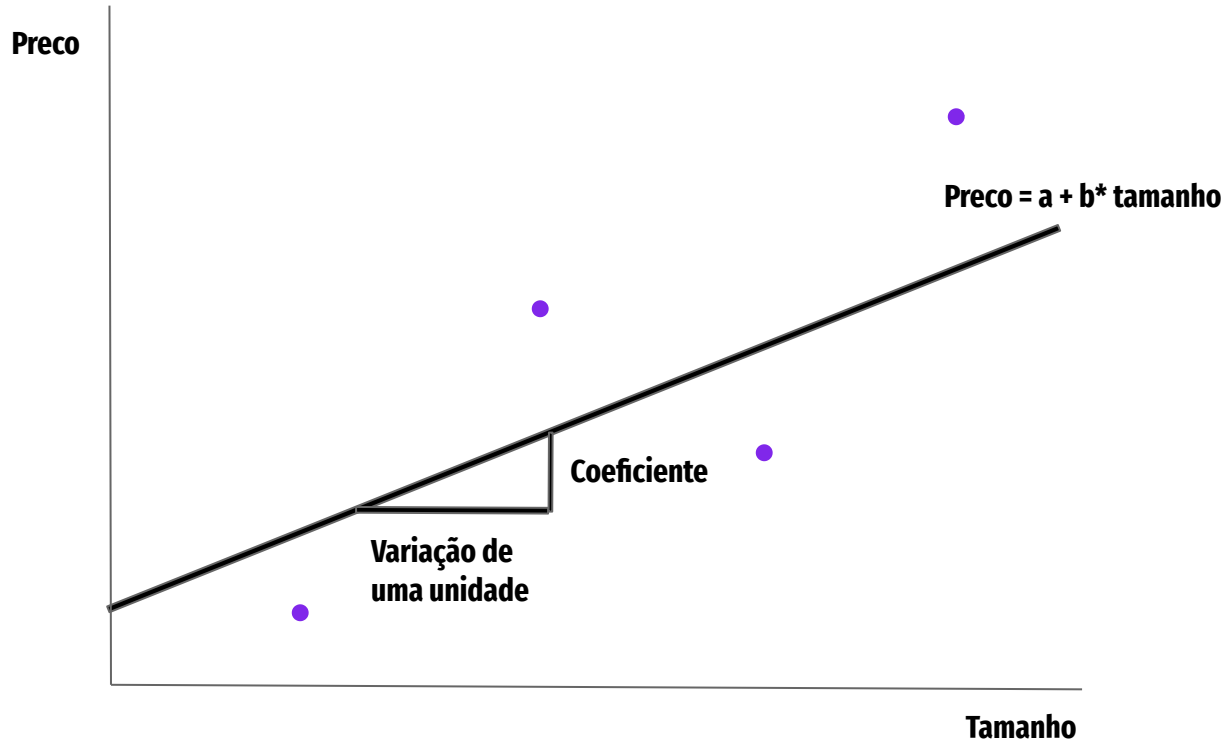




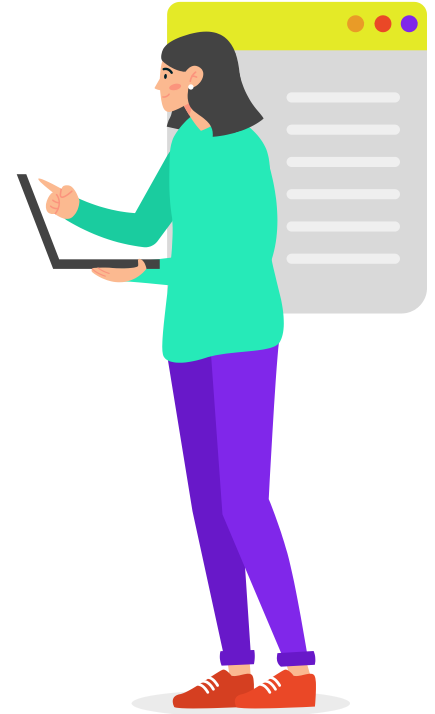
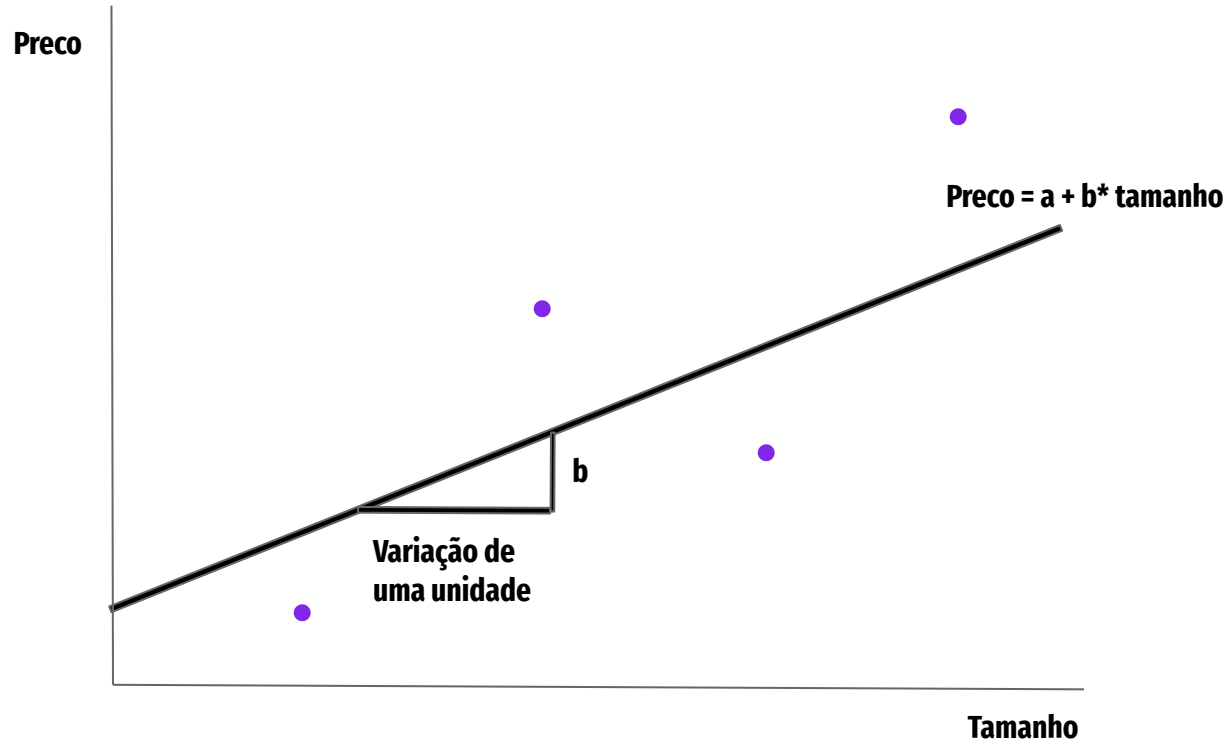
# Análise da reta



# Análise da reta



# Análise da reta



# Um pouco de teoria

## Teoria

$$Y = a + bX + \text{resíduo}$$



## Multicolineariedade

Os dados possuem correlação forte

## Independência

As observações não possuem dependência

## Heterocedasticidade

Variância constante

## Normalidade

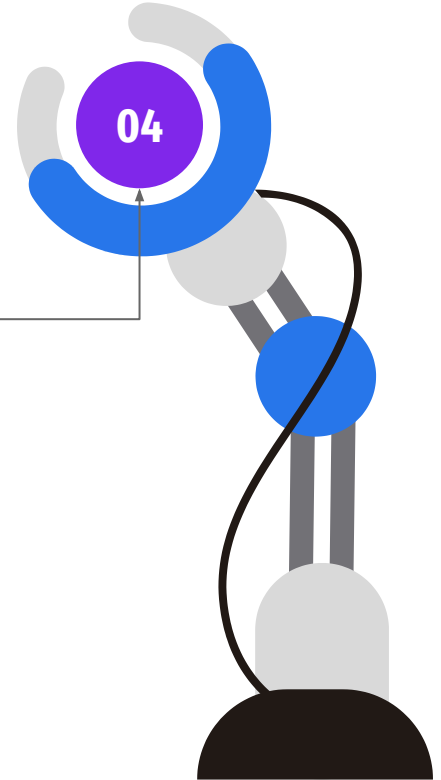
Os resíduos e o target são normais

01

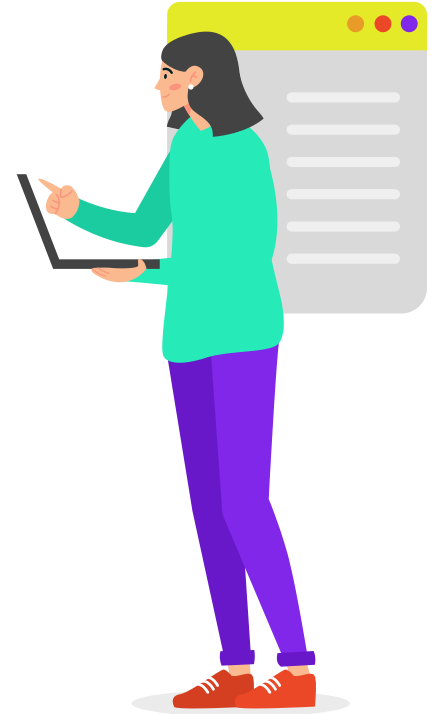
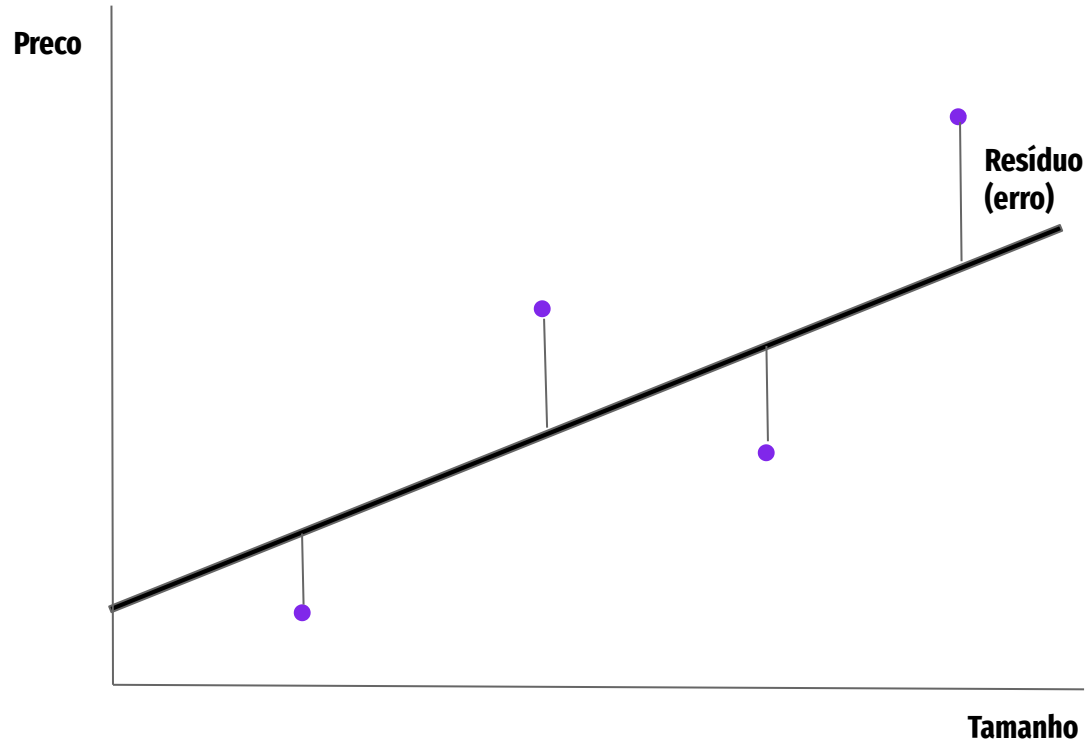
02

03

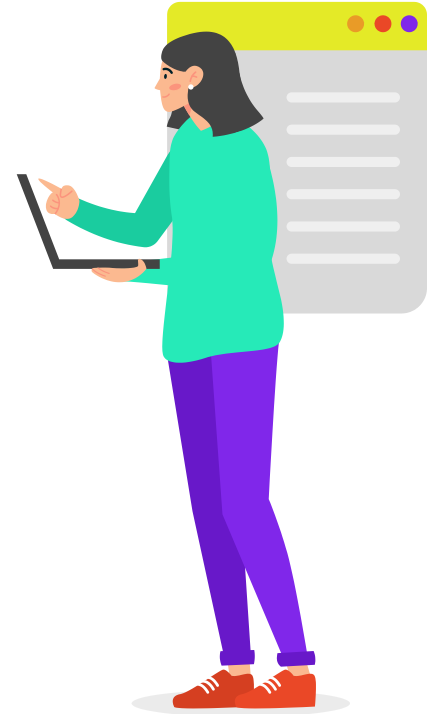
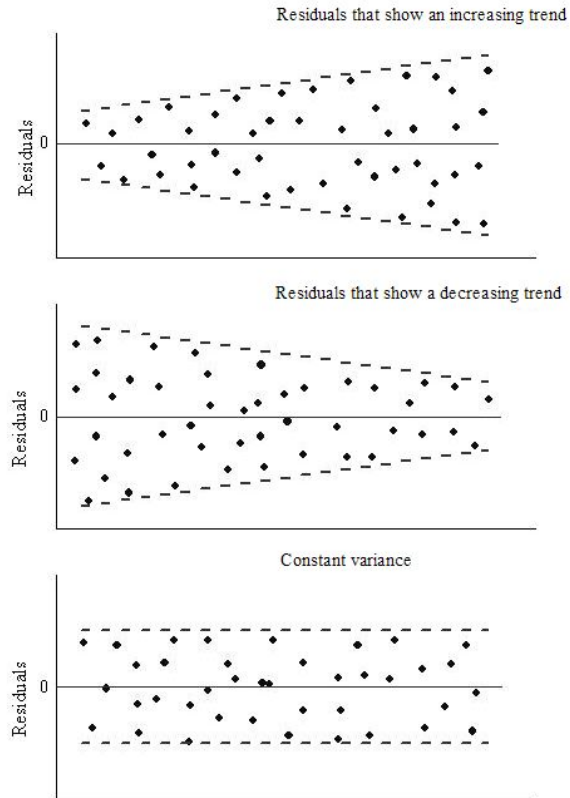
04



# Análise de resíduo (erro)

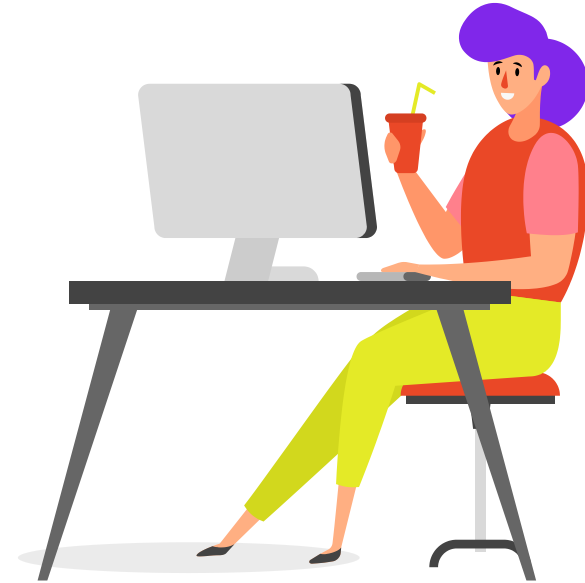


# Análise de resíduo (erro)

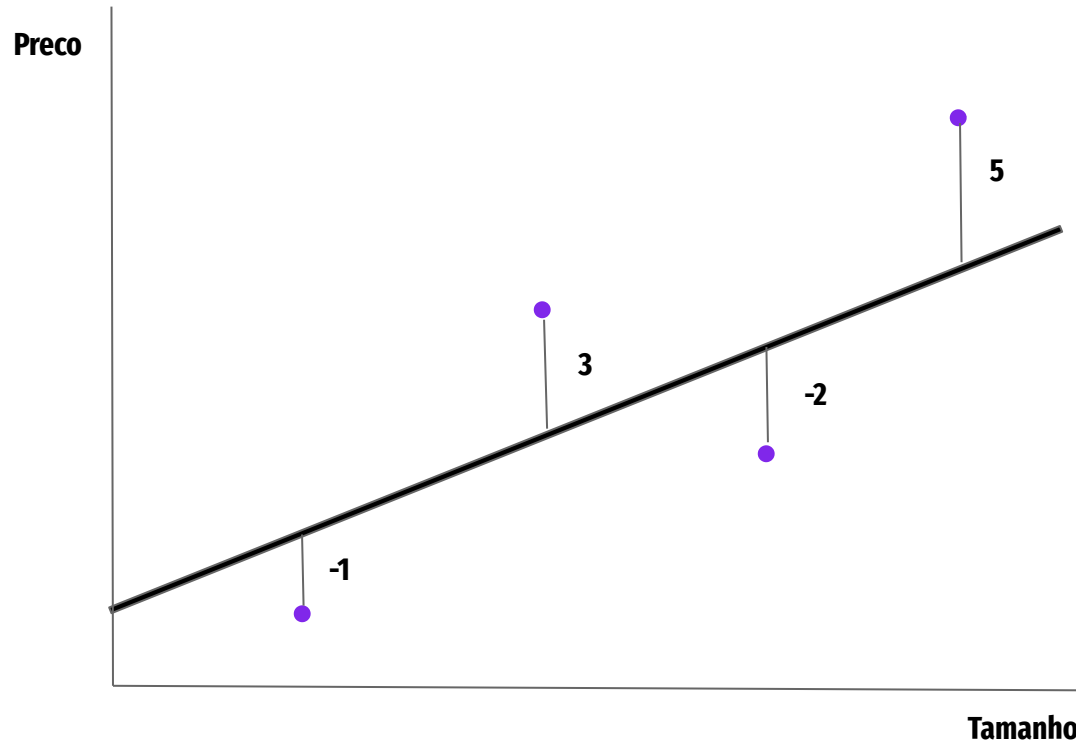


# Métricas de regressão

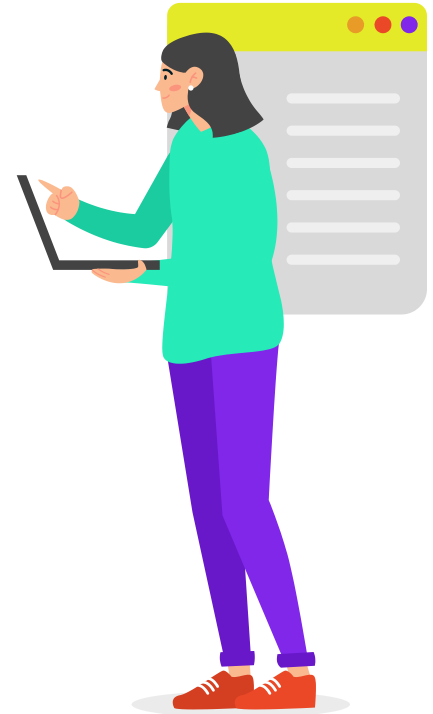
Métricas	
Erro quadrado médio (EQM)	A média da diferença entre o predito e o real ao quadrado
Raiz do erro quadrado médio (REQM)	Raiz quadrada do EQM. Prove a magnitude dos erros.
Erro absoluto médio (EAM)	A média da diferença absoluta entre o predito e o real
R <sup>2</sup>	Medida variabilidade que é explicada pelo modelo



# EQM

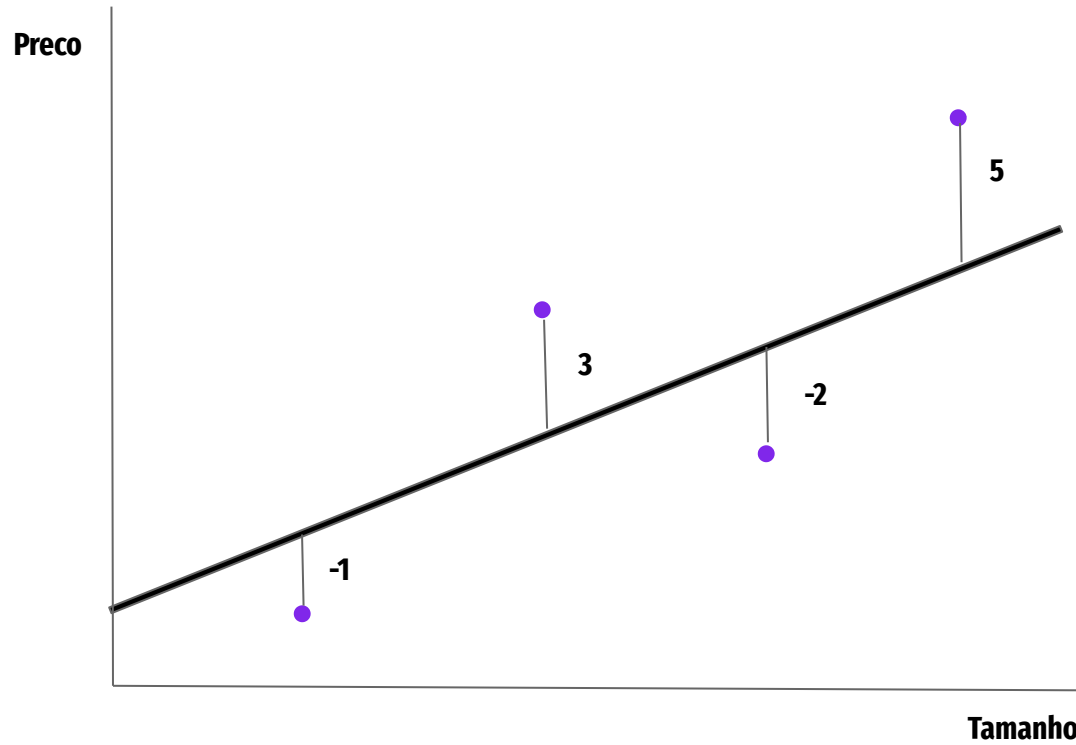


EQM = 9,75

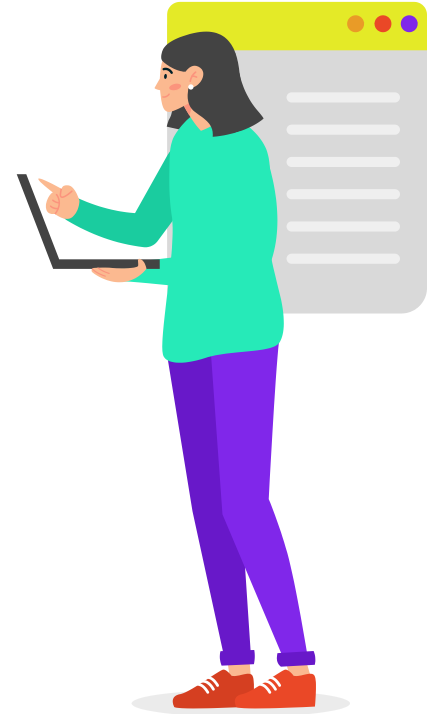




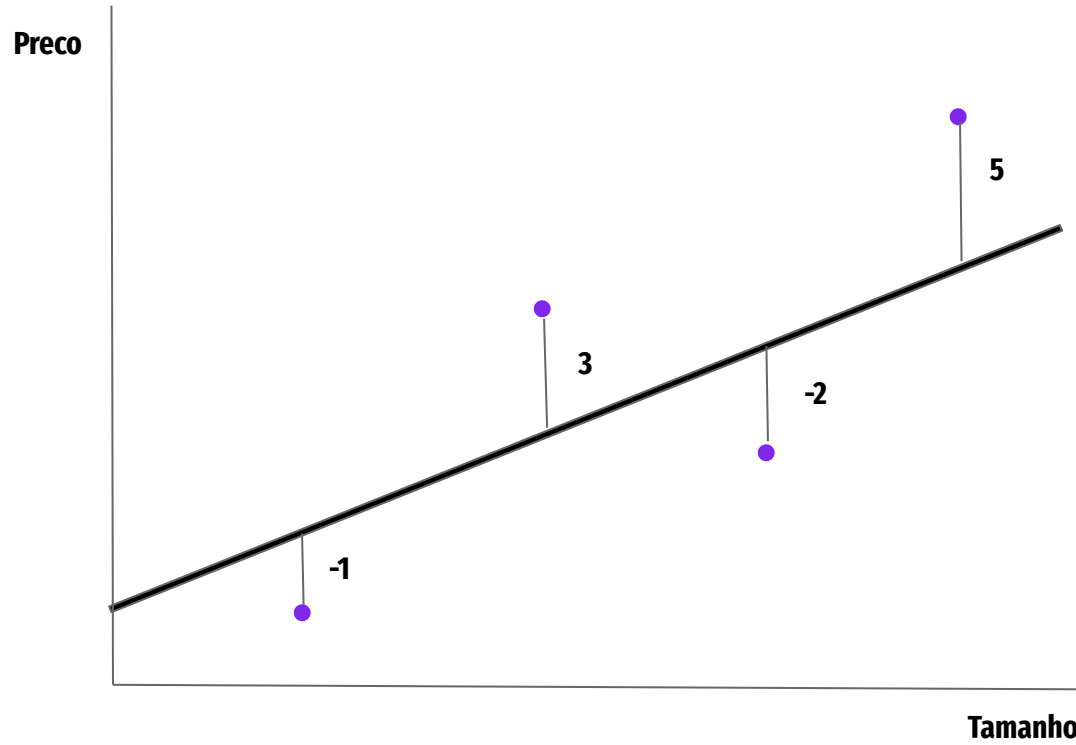
# REQM



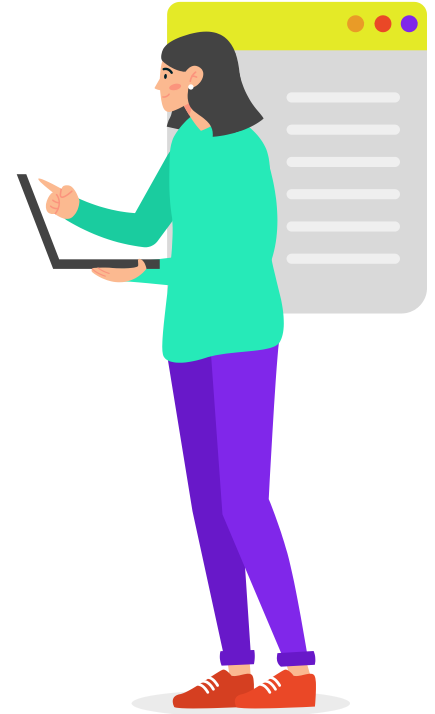
REQM = 3,12



# EAM



EAM = 3,25



# Outras métricas



**Métricas  
estatísticas**

**01**

## **Testes estatístico**

Testes para qualidade do modelo e coeficientes

**02**

## **Intervalo de confiança**

Estimativa intervalar para a variável alvo



**Hora da prática!!!!**