

# Inteligência Artificial

2024/2



---

Profa. Dra. Juliana Félix

[jufelix16@gmail.com](mailto:jufelix16@gmail.com)



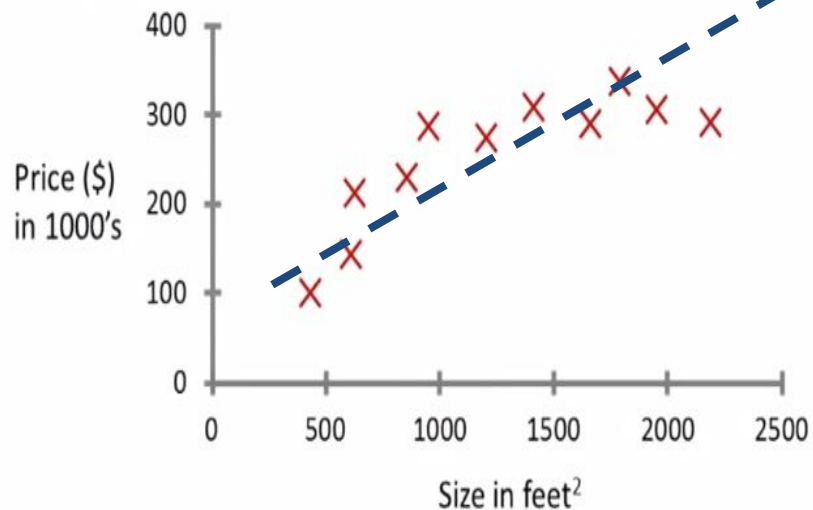
# Regressão Linear

# Regressão Linear

Modelo

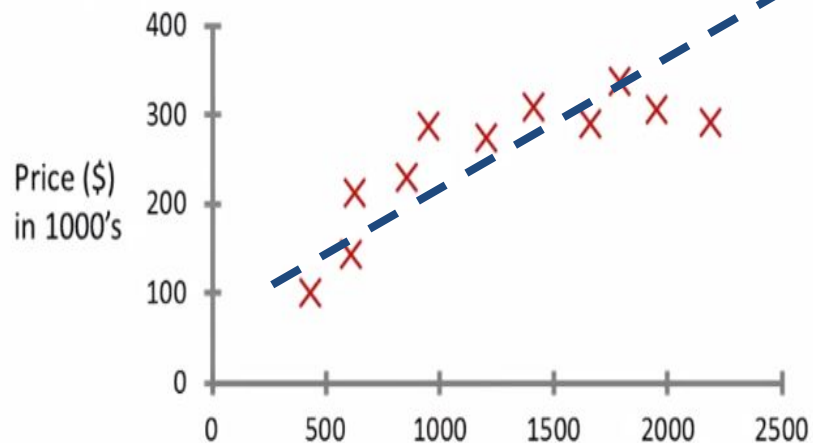
$$y = \theta_0 + \theta_1 x$$

Housing price prediction.



# Regressão Linear

Housing price prediction.



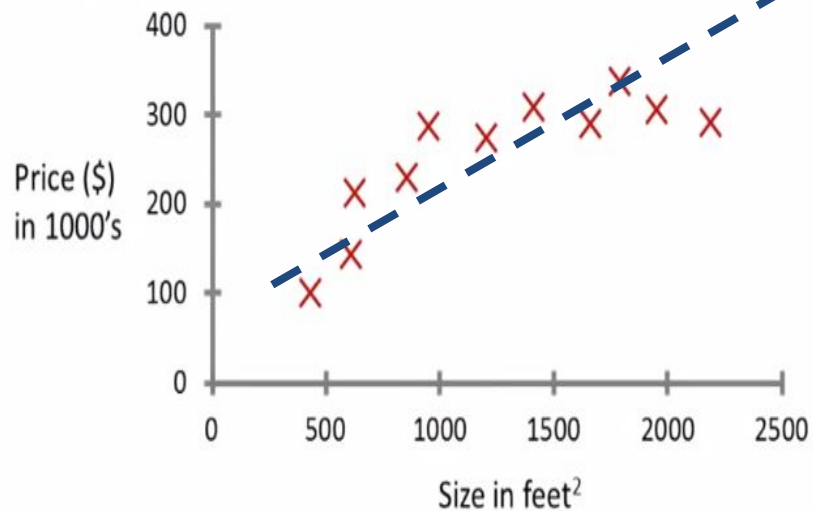
$$y = \theta_0 + \theta_1 x$$

Características  
(features)

Size in feet²

# Regressão Linear

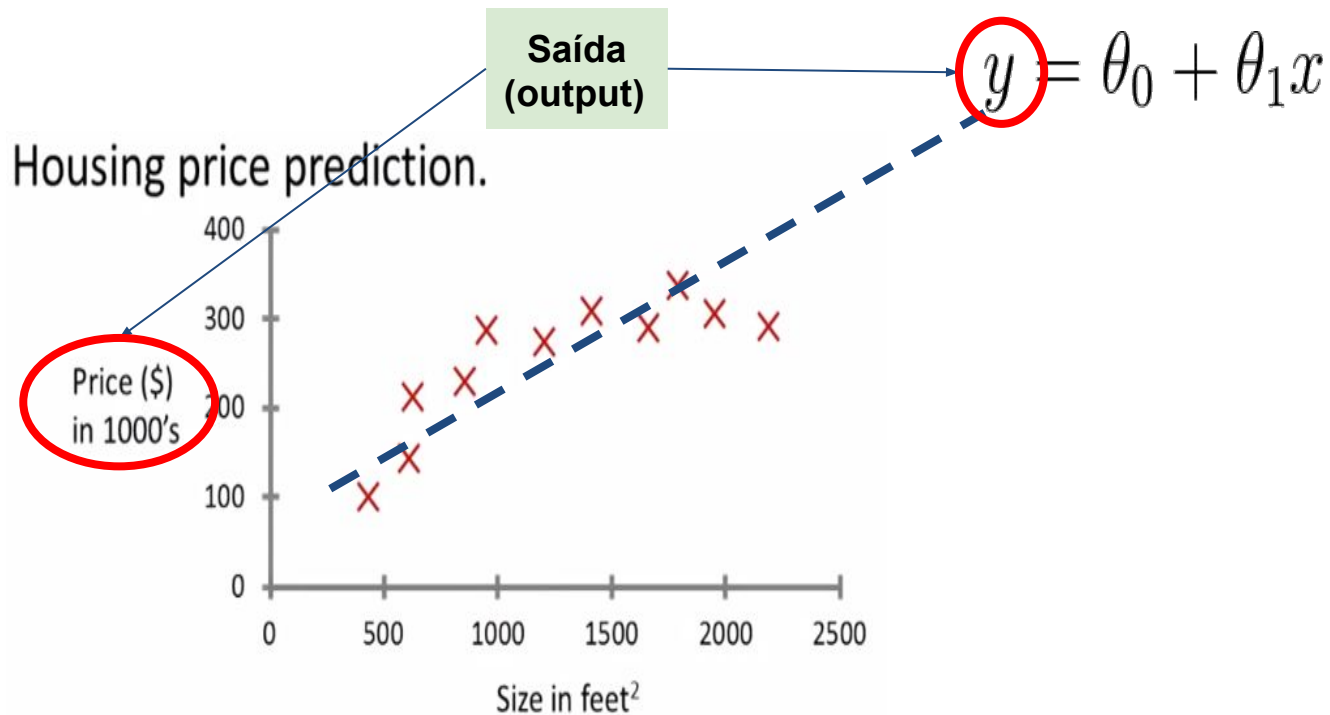
Housing price prediction.



$$y = \theta_0 + \theta_1 x$$

Parâmetros

# Regressão Linear





# Exemplo

# Conjunto de dados (*Dataset*)

Um **Dataset** (conjunto de dados) combina amostras com

- Valores ou **variáveis de entrada** (features, características, recursos) e
- Valores de **saída** (outcome, labels) utilizados no aprendizado supervisionado.



# Dataset

Exemplo:

Total amostras  
= 47

Preço de imóveis na 'Terra tão tão distante'	
Tamanho do imóvel (em m <sup>2</sup> ) - X	Preço (R\$) - Y
2104	399.900
1600	329.900
2400	369.000
...	...

Feature  
(característica)

Outcome  
(saída)

# Notação

Podemos pensar no problema anterior como um problema que tem:

- Um total de  $m$  amostras/*samples* ( $m = 47$ )
- Cada amostra tem 1 única *feature*/característica (tamanho do imóvel)
  - Costumamos representar uma variável de entrada por  $\mathbf{x}$
- Para cada amostra, temos uma única saída (preço do imóvel).
  - Costumamos representar uma variável de saída por  $\mathbf{y}$
- Cada amostra pode ser representada por um par, ou tupla  $(\mathbf{x}, \mathbf{y})$ 
  - Uma tupla  $(\mathbf{x}^i, \mathbf{y}^i)$  representa a  $i$ -ésima amostra do problema, com  $1 \leq i \leq m$

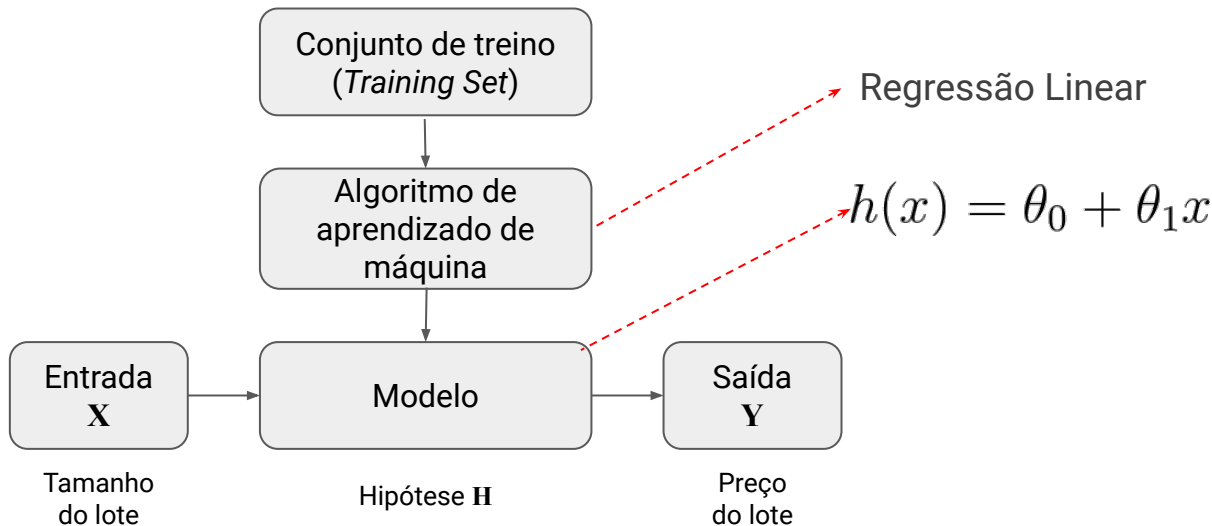
# Notação

- No conjunto abaixo, o par  $(x^1, y^1)$  refere-se aos dados da primeira linha (2104, 399900).

Preço de lotes na 'Terra tão tão distante'	
Tamanho do lote (em m <sup>2</sup> ) - X	Preço do lote (R\$) - Y
2104	399.900
1600	329.900
2400	369.000
...	...

# Processo básico de Machine Learning

A base de qualquer processo de machine learning consiste em mapear um dado de entrada **X** em um dado de saída **Y**.



# Desvio

Quando fazemos a predição de um valor, o **desvio** é a diferença entre o **valor esperado** (conhecido) e o **valor predito** pelo modelo construído.

$$\text{desvio}^i = Y^i - h(x^i)$$

$$\text{desvio}^i = Y^i - \hat{Y}^i$$

# MSE

O **Mean Square Error** (MSE - Erro Médio Quadrático) é a **média** do **quadrado** dos **erros** obtidos pelo modelo.

$$MSE = \frac{1}{m} \sum_{i=1}^m (Y^i - h(x^i))^2$$

# MSE

O **Mean Square Error** (*MSE* - Erro Médio Quadrático) é a **média** do **quadrado** dos **erros** obtidos pelo modelo.

$$MSE = \frac{1}{2m} \sum_{i=1}^m (Y^i - h(x^i))^2$$

Na prática, a média é dividida pela metade (1/2) como uma conveniência para o cálculo do **gradiente descendente**, método utilizado na regressão linear, que cancelará o termo 1/2.

Em outras palavras, dividir por  $m$  ou  $2m$  não traz diferenças significativas para o cálculo dos valores analisados.

# MSE

$$MSE = \frac{1}{2m} \sum_{i=1}^m (Y^i - h(x^i))^2$$

Considerando os seguintes valores preditos, podemos calcular o *MSE* do modelo.

Preço de lotes na 'Terra tão tão distante'				
x	y	h(x)	desvio	desvio <sup>2</sup>
2104	399.900	399.800		
1600	329.900	339.900		
2400	369.000	367.000		
			<b>Soma</b>	
			<b>MSE</b>	



# MSE

$$MSE = \frac{1}{2m} \sum_{i=1}^m (Y^i - h(x^i))^2$$

Considerando os seguintes valores preditos, podemos calcular o *MSE* do modelo.

Preço de lotes na 'Terra tão tão distante'				
x	y	h(x)	desvio	desvio <sup>2</sup>
2104	399.900	399.800	100	10.000
1600	329.900	339.900	-10.000	10.000.000
2400	369.000	367.000	2.000	4.000.000
Soma				14.010.000
MSE				2.335.000

# Exercício

Considerando os valores  $x$  e  $y$  fornecidos, tente encontrar, manualmente, uma reta que melhor se ajuste aos dados abaixo:

```
x = np.array([480, 510, 520, 850, 960, 1200, 1400, 1650, 1700, 1920, 2350])  
y = np.array([98, 110, 200, 210, 280, 265, 300, 287, 325, 300, 290])
```

Faça isso utilizando:

- a) Apenas  $\theta_0$
- b) Apenas  $\theta_1$
- c) Atribuindo valores para  $\theta_0$  e  $\theta_1$

Para cada reta, calcule o respectivo  $MSE$ . Plote as retas e o  $MSE$  encontrado em todos os casos.

# Aula de hoje...

- Falamos sobre o funcionamento da regressão linear
  - Entendemos o que é um **modelo** de RL
  - Como funciona o treinamento de um modelo de RL
  - Como o cálculo do erro é realizado

## Na próxima aula:

- Veremos como minimizar o erro (função custo)
- Apresentaremos o algoritmo do gradiente descendente para minimização da função custo.

# Atividade 2

- A atividade pode ser feita em duplas.
- Sua solução deve ser entregue em um notebook python (.ipynb).
- Reinicie o ambiente de execução e rode todas as células antes de gerar o arquivo a ser entregue (garantindo a exibição de todos os gráficos, por ex.)
- A entrega deve ser realizada até 23h59 de segunda-feira, 16/09/2024.
- O ambiente de entrega ainda será definido (estou aguardando acesso aos sistemas da universidade) e a turma será avisada.

# Leitura recomendada

Calculadora gráfica: [Desmos | Calculadora Gráfica](#)

Regressão linear: [Explicação sobre o modelo de regressão linear](#)