

# Algoritmos Supervisionados – Regressão Linear Simples.

Os algoritmos clássicos para prever valores



# Algoritmos Supervisionados – Regressão Linear Simples

O objetivo deste módulo é apresentar conceitualmente os **principais algoritmos de regressão** para que possamos desenvolver projetos de machine learning que fazem previsões de valores. E faremos um projeto explorando o primeiro destes algoritmos, que é o de **regressão linear simples**, onde faremos o **processo completo** desde o EDA até a entrega do modelo através de uma API para consumo por outras aplicações.

---

<b>Algoritmos Supervisionados – Regressão Linear Simples</b>	01.	O que é a análise de regressão?
	02.	Um passeio pelos tipos de regressão
	03.	Projeto - Regressão Linear Simples

# O que é Análise de Regressão?

A **análise de regressão** é uma abordagem **estatística** que busca **investigar e quantificar as relações** entre variáveis. Ela é usada para entender como uma variável dependente (target) está relacionada a uma ou mais variáveis independentes (fatores que acreditamos influenciar a variável dependente).

Essa técnica permite construir um **modelo matemático**, geralmente na forma de uma equação linear, que representa essa relação. Ao ajustar o modelo aos dados observados, podemos **estimar** os parâmetros e entender como as **mudanças nas variáveis independentes** afetam a variável dependente.

Isso é fundamental para **fazer previsões**, tomar decisões embasadas em evidências e **compreender melhor os padrões em dados do mundo real**.

---

# Um passeio pelos tipos de Regressão

## Linear Simples

Envolve uma variável independente e uma variável dependente, sendo representada por uma linha reta.

**Exemplo:** Prever pontuação num exame com base no tempo de estudo.

## Linear Múltipla

Utiliza várias variáveis independentes para prever uma variável dependente, resultando em um modelo linear mais complexo.

**Exemplo:** Prever desempenho de alunos com base em múltiplas variáveis.

## Logística

Usada geralmente para prever probabilidades em problemas de classificação (binária, multinomial, ordinal, aninhada)

**Exemplo:** Prever se cliente fará compra online (sim/não)

## Polinomial

Permite modelar relações não-lineares entre variáveis. A regressão polinomial usa uma curva (um polinômio) para se ajustar aos dados.

**Exemplo:** Previsão de vendas de produtos com base no clima/temperatura.

# Um passeio pelos tipos de Regressão

## Poisson

Adequada para modelar dados de contagem, como o número de eventos em um determinado período. É comum em estudos epidemiológicos e de contagem. **Exemplo:** Número de Reclamações de Sinistros

## Ridge, Lasso e ElasticNet

Técnicas de regularização para lidar com multicolinearidade e overfitting. São usadas em análises onde há muitas variáveis independentes. **Exemplo:** Estimar preços de casas com base em muitas variáveis.

## Gamma

Usada quando os dados não seguem uma distribuição normal e exibem uma assimetria positiva (à direita). Além disso, o target é contínuo e positivo. **Exemplo:** Previsão de tempo de permanência de pacientes em um hospital.

## Beta

Aplicada quando a variável dependente está restrita ao intervalo [0, 1], comum em análises de proporções e taxas. **Exemplo:** Taxa de conversão com base numa campanha de Mkt.

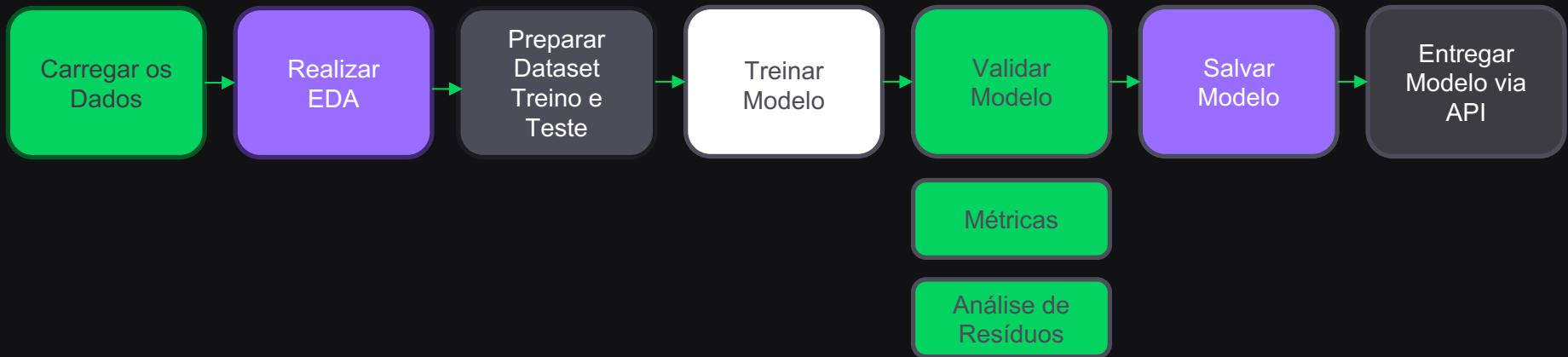
# Regressão Linear Simples – Projeto

Uma plataforma online de educação, que consegue acompanhar a quantidade de horas que seus alunos passam estudando numa determinada trilha, deseja avaliar se esta quantidade de horas influencia na pontuação do teste final.

Para isso, iremos treinar um **algoritmo de regressão linear**, de forma que seja possível **prever a pontuação do teste final**, dada a **quantidade de horas de estudo**.

---

# Regressão Linear Simples



# Regressão Linear Simples

$$y = mx + b$$

single value of dependent variable

slope

single value of independent variable

y-intercept

$$Y = \beta_0 + \beta_1 X + \epsilon$$

all observed values for dependent variable

y-intercept aka "bias"

slope aka "coefficient"

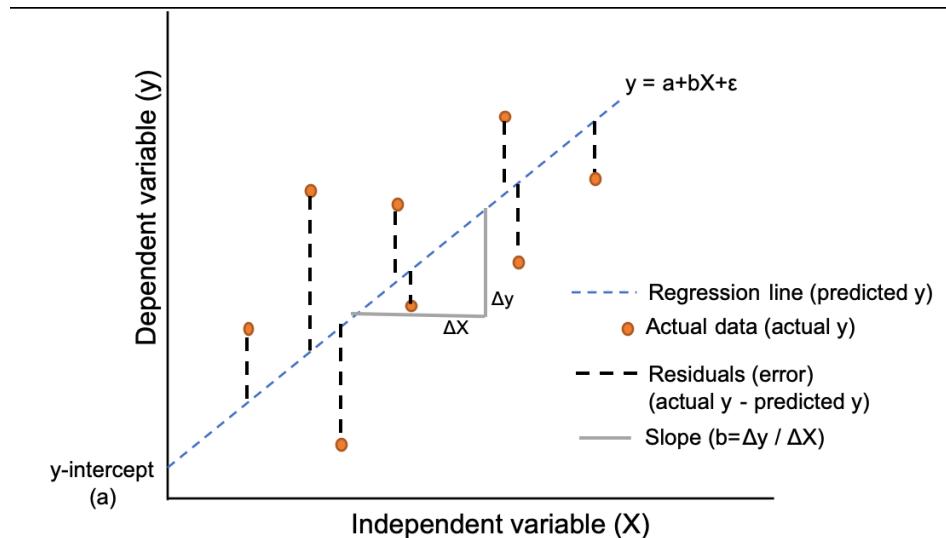
all observed values of independent variable

error\*

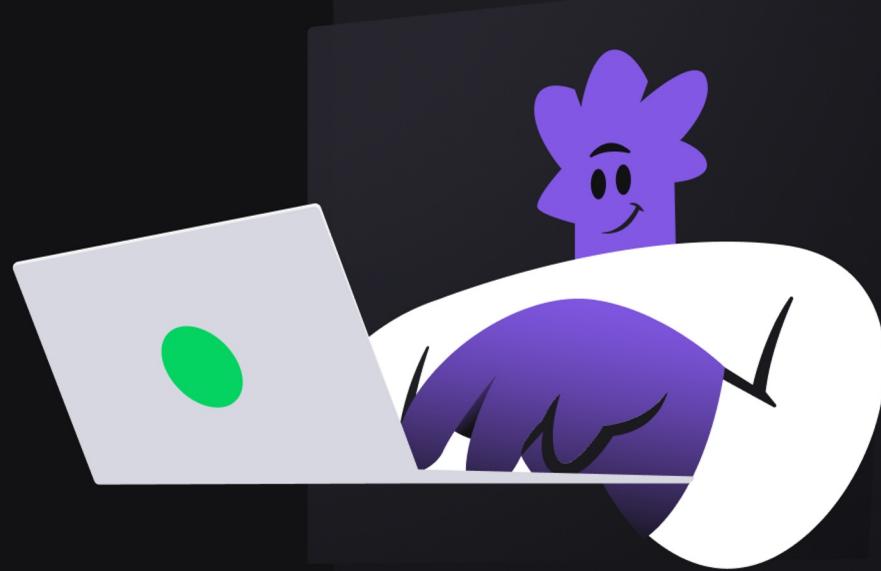
\* additional term

$\alpha$

Ordinary Least Squares – Método dos Mínimos Quadrados



# Let's Go, Let's Go



Code time ...



# Boosting People.

[rocketseat.com.br](https://rocketseat.com.br)

---