

Análise e Predição do Valor de Aluguéis de Imóveis no Brasil

INTELIGÊNCIA ARTIFICIAL

Discentes:

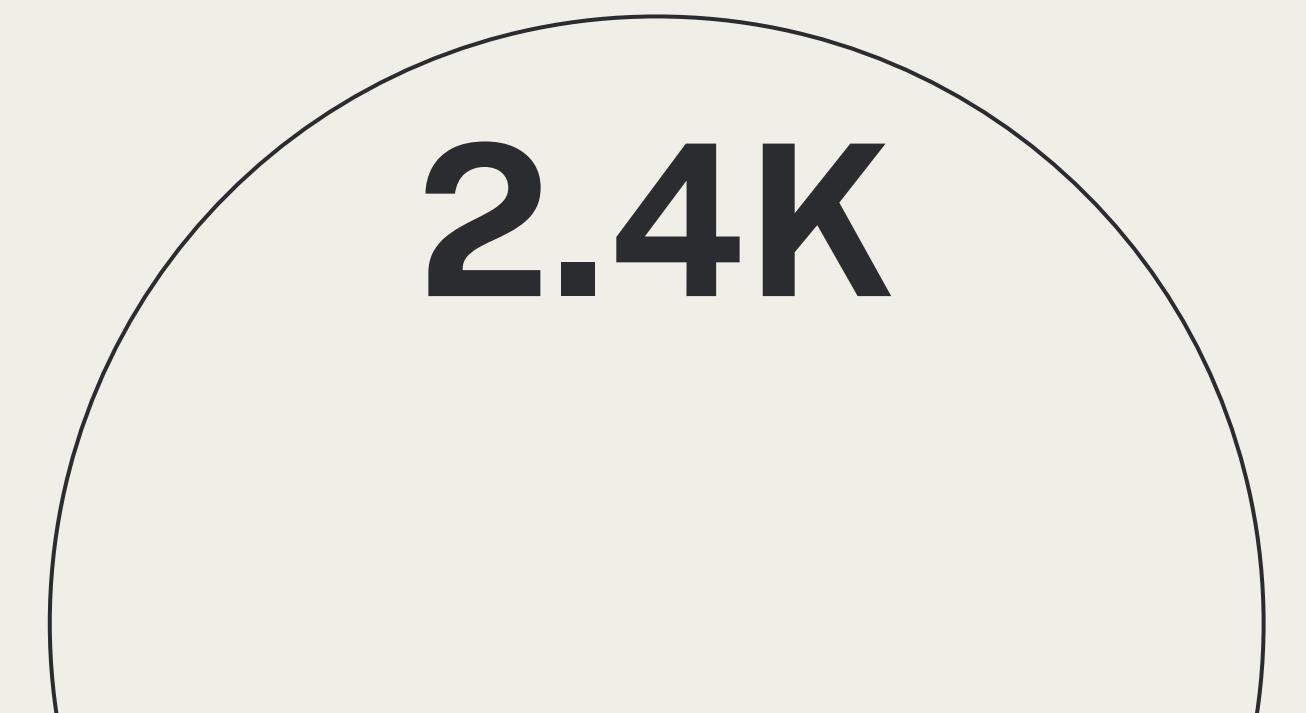
Diego Metring Teixeira
Murilo Ribeiro da Costa



PROBLEMA, DOMÍNIO E BASE DE DADOS



35.7K



2.4K

PROBLEMA

Entender e prever o valor de aluguéis em um mercado complexo.

DOMÍNIO

Mercado Imobiliário de Aluguéis no Brasil

BASE DE DADOS

Brazilian Real Estate to Rent (Kaggle)

Foco da análise: Imóveis na faixa de R\$300 a R\$1300, visando uma média estimada do custo dos imóveis na época da pesquisa

Volume inicial: 35.773 entradas, 78 colunas

Volume final: 2399 entradas, 13 colunas



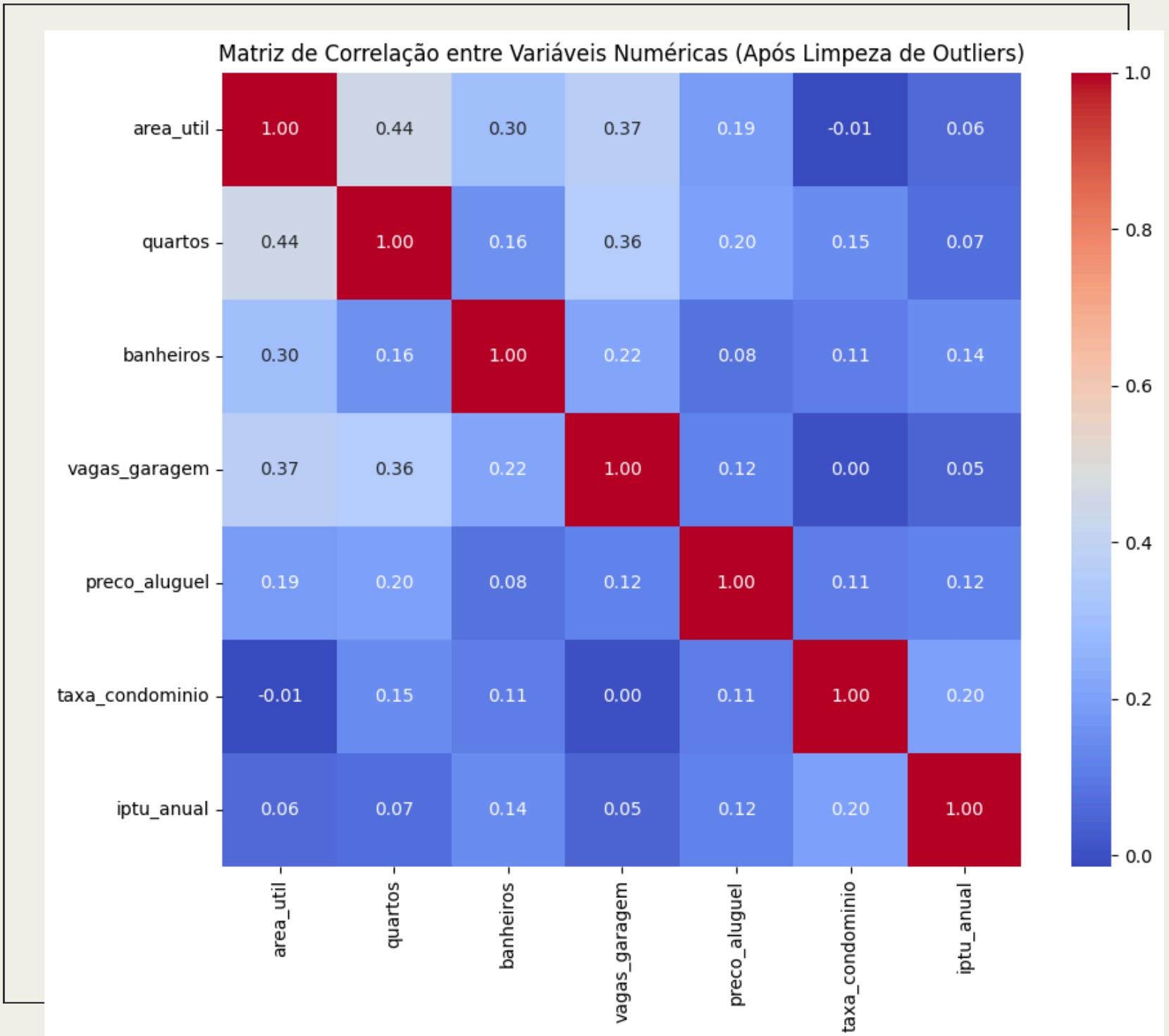
**CENTRO
UNIVERSITÁRIO**

PRÉ-PROCESSAMENTO E LIMPEZA DE DADOS

<i>Seleção de Colunas Relevantes</i>	<i>Renomeação e Conversão de Tipos</i>	<i>Preenchimento de Valores Ausentes</i>	<i>Remoção de Outliers e Filtragem</i>	<i>AED e Análise de Correlação</i>
Como principais atributos do dataset, foram selecionadas as colunas cidade, estado, área útil, quartos, banheiros, suítes, vagas na garagem, se é mobiliado, se tem piscina, se tem academia, valor do aluguel, taxa de condomínio e IPTU.	Colunas renomeadas para facilitar o uso, conversão de dados textuais para numéricos onde aplicável, valores ausentes preenchidos.	Linhas com valores zero ou nulos nas colunas preco_aluguel e area_util removidas.	Outliers identificados e removidos com base na coluna area_util utilizando o percentil 99. Foram incluídos somente os imóveis com faixa de preço de aluguel entre R\$300 e R\$1300.	Foi gerado um resumo estatístico das colunas numéricas após a última etapa para entender a distribuição e as principais medidas de tendência central e dispersão. As relações lineares entre os atributos foram identificadas por uma matriz de correlação.

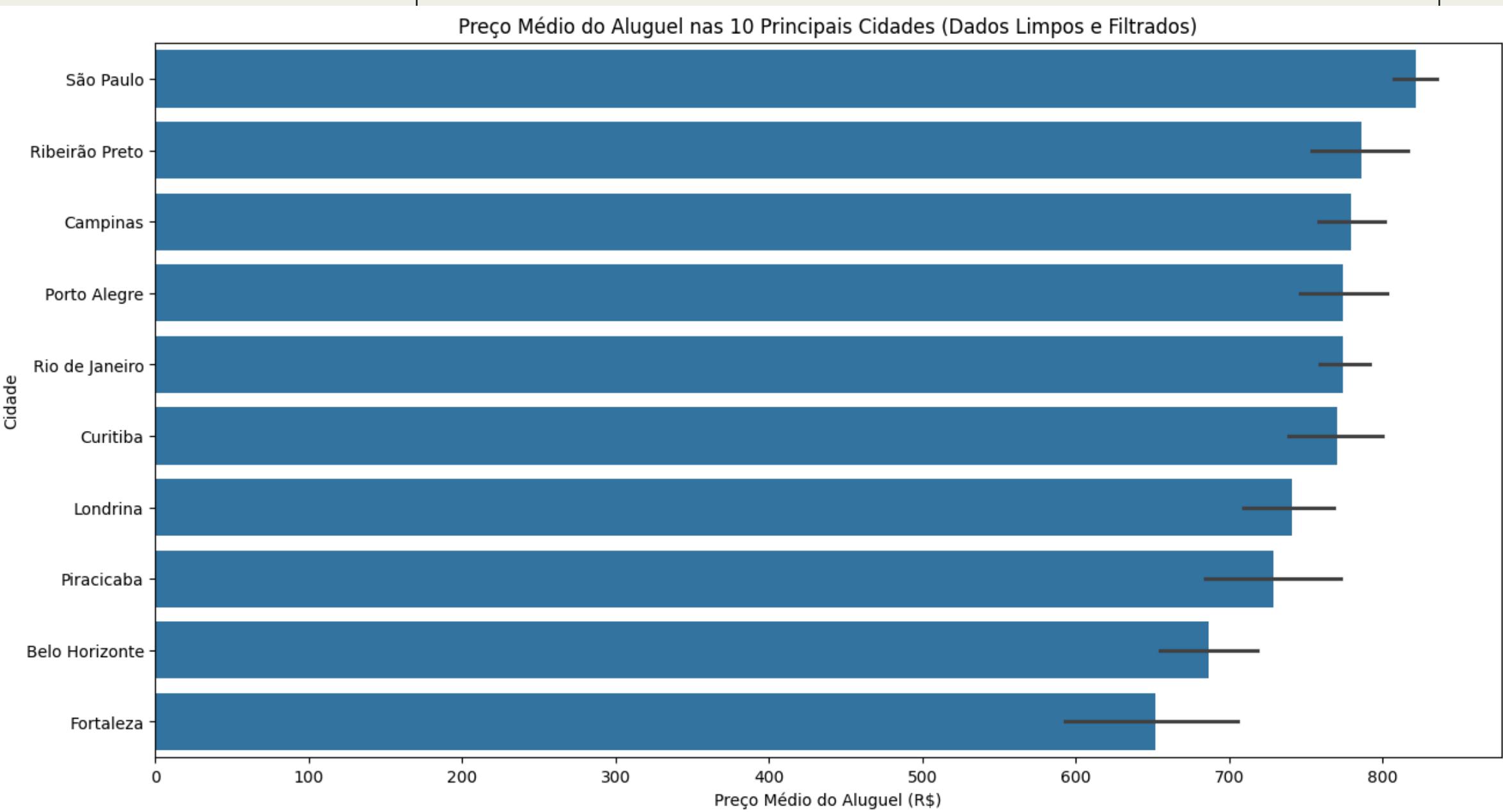
RESULTADOS DA AED

- O preço do aluguel apresenta maior correlação positiva com os atributos:
 - **Número de quartos (0.20)**
 - **Área útil (0.19)**
 - **Vagas na garagem e IPTU Anual (0.12)**



RESULTADOS DA AED

- A cidade também pode ser classificada como um dos principais fatores influentes no valor do aluguel, como pode ser visto na imagem



PREPARAÇÃO PARA MODELAGEM

- **Variável Alvo (y):** preco_aluguel.
- **Features (X):** Demais colunas após limpeza.
- **Engenharia de Atributos:**
 - One-Hot Encoding para variáveis categóricas (cidade, estado, mobiliado, piscina, academia).
 - Preenchimento de NaNs restantes em X com 0.
- **Divisão Treino/Teste:** 80% Treino, 20% Teste
(random_state=42)

MACHINE LEARNING (ALGORITMO 1)

- **Algoritmo:** Random Forest Regressor.
- **Funcionamento:** Combina várias árvores de decisão para prever valores contínuos.
- **Treinamento:** Modelo inicial (model.fit(X_train, y_train)).
- **Otimização de Hiperparâmetros:** RandomizedSearchCV para buscar melhores n_estimators, max_depth, etc.
- **Métricas de Avaliação (Teste):**
 - **MAE (Erro Médio Absoluto):** R\$80.05 (inicial), R\$79.85 (otimizado).
 - **R² (Explicabilidade):** 0.42 (inicial), 0.43 (otimizado).

MACHINE LEARNING (ALGORITMO 2)

- **Algoritmo:** Gradient Boosting Regressor.
- **Funcionamento:** Constrói modelos sequencialmente, corrigindo erros dos modelos anteriores.
- **Treinamento:** Modelo `gbr_model.fit(X_train, y_train)`.
- **Métricas de Avaliação (Teste):**
 - **MAE:** R\$ 107.11
 - **R²:** 0.36

COMPARATIVO DE PERFORMANCE

Random Forest Inicial

- $R^2 = 0.42$
- **MAE** = R\$80,05
- **MSE** = 13411.78

Random Forest Otimizado

- $R^2 = 0.43$
- **MAE** = R\$79.85
- **MSE** = 13227.05

Gradient Boosting Regressor

- $R^2 = 0.36$
- **MAE** = R\$107,11
- **MSE** = 17471.63

INSIGHTS

Modelo Vencedor

Random Forest
(Otimizado) com MAE de
~R\$80 e R^2 de 0.43.

Fatores Não Capturados

Estado de conservação, vista,
andar, barulho, proximidade a
serviços, etc.

Limitação

O modelo explica 43% da
variação de preço; 57% ainda
não é capturada



CENTRO
UNIVERSITÁRIO

REFERÊNCIAS

maverickjpa. (Ano do Upload ou Acesso, se disponível, caso contrário, omitir). Brazilian Real Estate to Rent. Kaggle. Disponível em: <https://www.kaggle.com/datasets/maverickjpa/brazilian-real-estate-to-rent>. Acesso em: 25 de novembro de 2025.

Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830. Documentação oficial de *RandomForestRegressor* e *GradientBoostingRegressor*. Disponível em: <https://scikit-learn.org/stable/>. Acesso em: 25 de novembro de 2025.

Obrigado!



CENTRO
UNIVERSITÁRIO