

Projeto de Preços de Imóveis

Análise de fatores determinantes na precificação imobiliária utilizando Regressão Linear

Introdução

Contextualização:

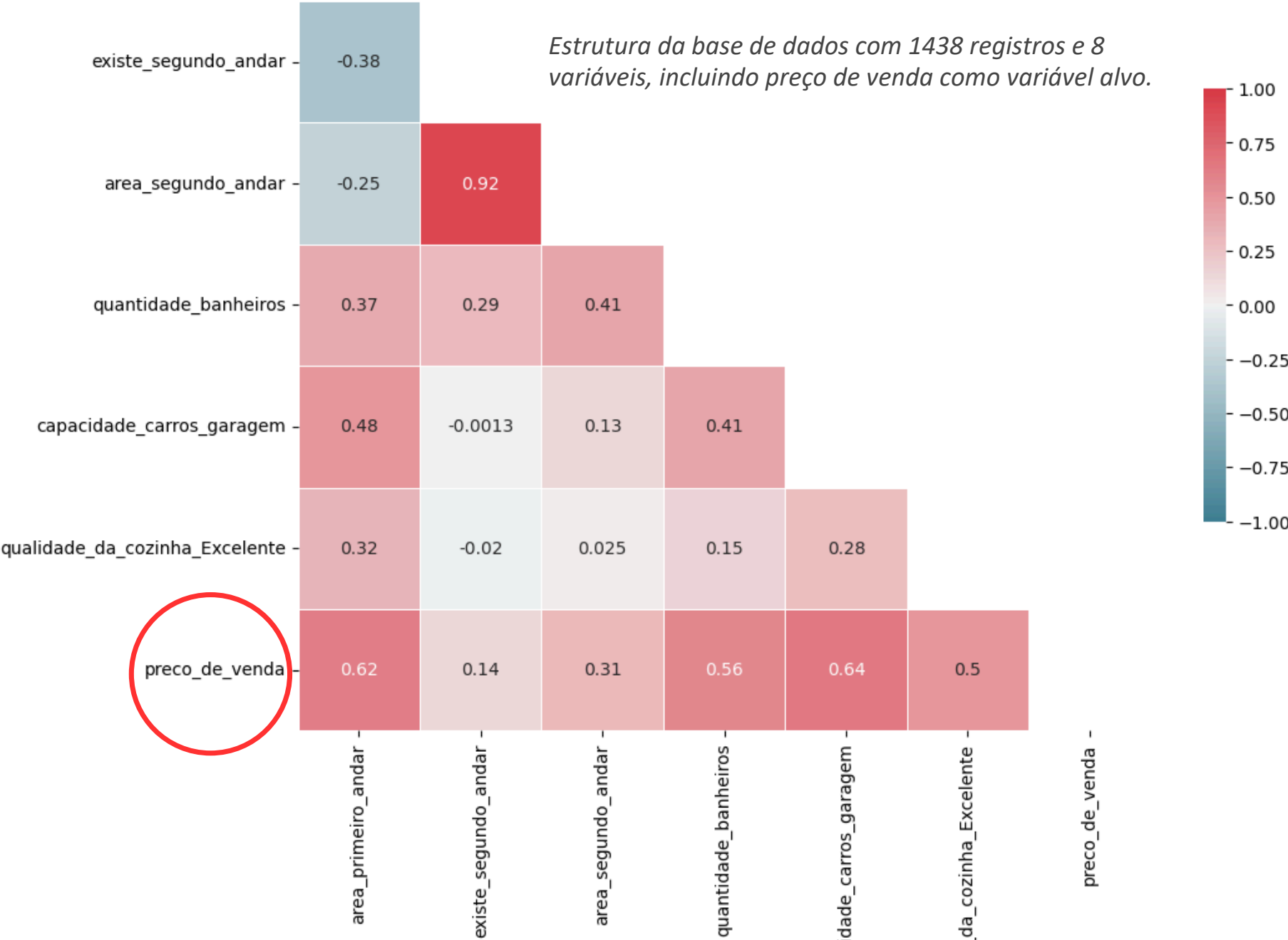
- O mercado imobiliário é altamente dinâmico e complexo
- A precificação adequada é essencial para compradores e vendedores
- Decisões baseadas em dados podem gerar vantagens competitivas

Objetivos do Projeto:

- Estimar preços de imóveis com base em suas características
- Identificar aspectos que contribuem para precificação
- Entender qual aspecto é mais relevante na determinação do preço
- Demonstrar como precificar um imóvel novo

Base de Dados

Estrutura da base de dados com 1438 registos e 8 variáveis, incluindo preço de venda como variável alvo.



Detalhamento da Base de Dados

Origem dos Dados:

- Base inspirada em 'House Prices' do Kaggle
- 1438 registros de imóveis

Variáveis Disponíveis:

- area_primeiro_andar: Área em metros quadrados
- existe_segundo_andar: Variável binária (0 ou 1)
- area_segundo_andar: Área em metros quadrados
- quantidade_banheiros: Número de banheiros
- capacidade_carros_garagem: Capacidade da garagem
- qualidade_da_cozinha_Excelente: Variável binária (0 ou 1)
- preco_de_venda: Variável alvo (em reais)

Correlação com o Preço de Venda

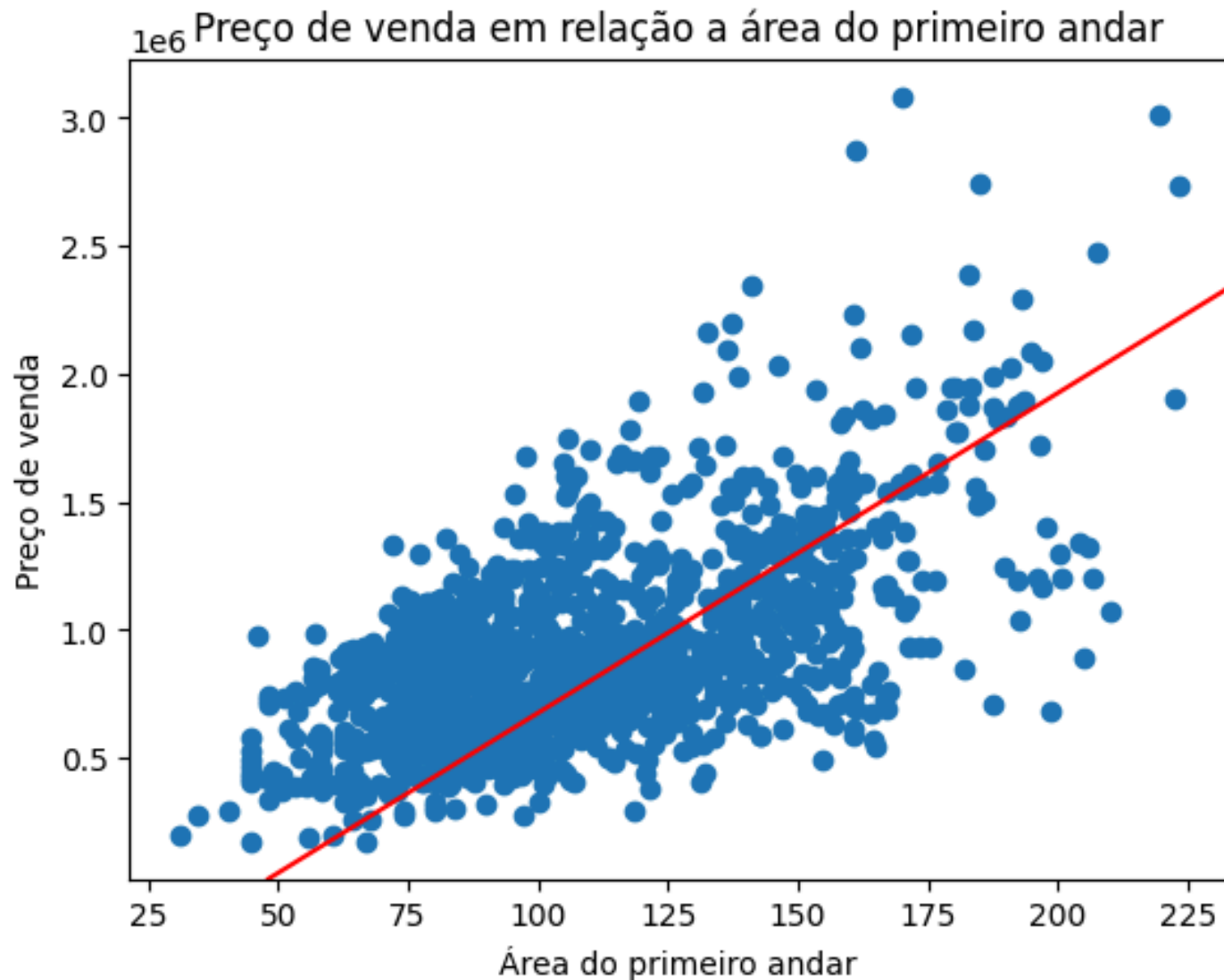
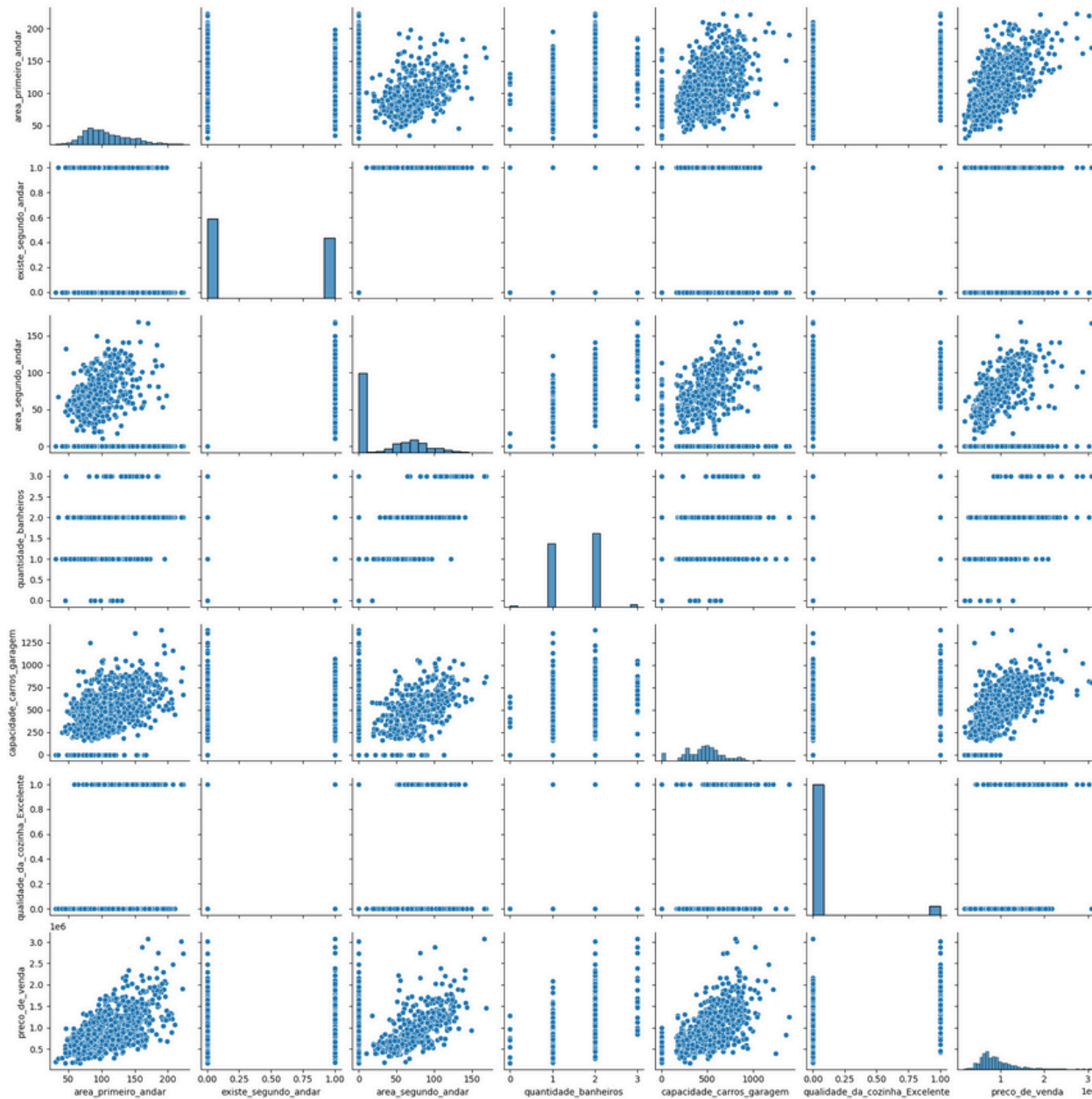


Tabela de correlações mostrando a relação entre a variável da área do primeiro andar e o preço de venda.

Matriz de Correlação (Heatmap)



Visualização das correlações entre todas as variáveis. Cores mais intensas indicam correlações mais fortes.

Interpretação das Correlações

Interpretação das Correlações:

- Correlação positiva: quando uma variável aumenta, o preço também tende a aumentar
- Quanto mais próximo de 1, mais forte é a correlação

Principais Descobertas:

- Capacidade da garagem (0.64) - correlação forte
- Área do primeiro andar (0.62) - correlação forte
- Quantidade de banheiros (0.56) - correlação moderada
- Qualidade da cozinha (0.50) - correlação moderada
- Área do segundo andar (0.31) - correlação fraca
- Existência de segundo andar (0.14) - correlação muito fraca

Modelagem com Regressão Linear

Conceito:

- A regressão linear busca estabelecer uma relação linear entre variáveis independentes e uma variável dependente

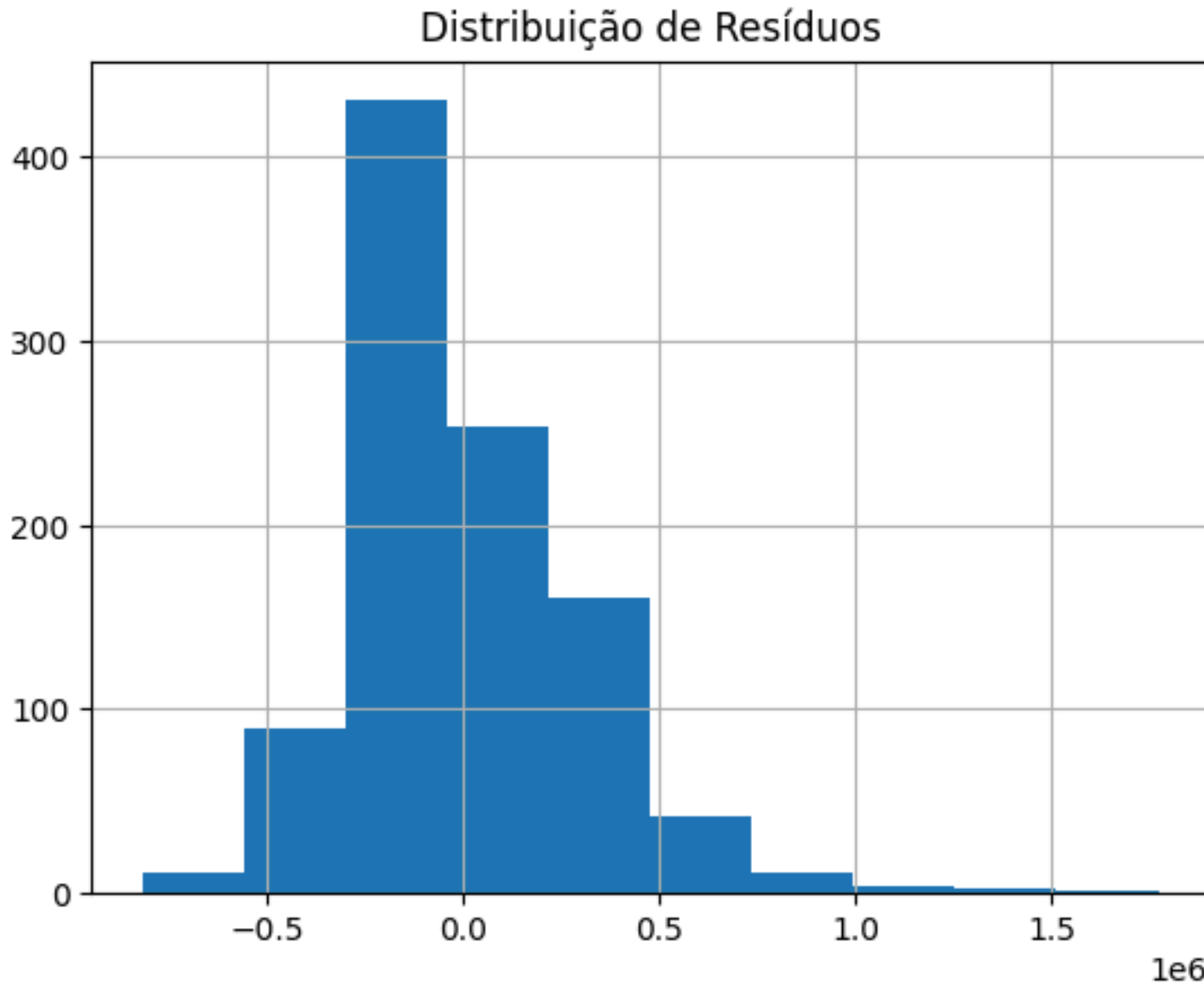
Processo:

- Divisão dos dados em conjuntos de treino (70%) e teste (30%)
- Construção do modelo utilizando as variáveis selecionadas
- Treinamento do modelo com os dados de treino

Equação do Modelo:

- $\text{Preço} = \beta_0 + \beta_1 \times (\text{area_primeiro_andar}) + \beta_2 \times (\text{existe_segundo_andar}) + \dots + \beta_n \times (\text{variável_n})$
- Os coeficientes (β) indicam o impacto de cada variável no preço final

Análise dos Resíduos



Distribuição dos resíduos do modelo. Uma distribuição próxima à normal indica um bom ajuste.

Avaliação do Modelo

Métricas de Desempenho:

- R^2 (coeficiente de determinação): Indica quanto da variação do preço é explicada pelo modelo
- Erro médio absoluto: Média dos erros absolutos entre preços reais e previstos
- Erro quadrático médio: Média dos quadrados dos erros

Análise dos Resíduos:

- Distribuição dos resíduos próxima à normal
- Resíduos sem padrões sistemáticos indicam um bom ajuste do modelo
- Verificação de pressupostos da regressão linear

Resultados e Insights

Fatores mais Influentes no Preço:

- Capacidade da garagem é o fator mais relevante
- Área do primeiro andar tem forte impacto no preço
- Quantidade de banheiros apresenta influência significativa
- Qualidade da cozinha também é um fator importante

Insights Principais:

- Investir em garagem pode valorizar significativamente o imóvel
- A área do primeiro andar tem mais impacto que a do segundo
- Banheiros adicionais aumentam consideravelmente o valor do imóvel
- Uma cozinha de excelente qualidade pode justificar um preço premium

Aplicação Prática

Como Precificar um Novo Imóvel:

- Coletar as características do imóvel (área, banheiros, garagem, etc.)
- Aplicar os valores na equação do modelo
- Obter a estimativa de preço

Exemplo:

Para um imóvel com as seguintes características:

- Área do primeiro andar: 100m²
- Sem segundo andar
- 2 banheiros
- Garagem para 2 carros
- Cozinha padrão (não excelente)

O modelo estima um preço de aproximadamente R\$ 686.448,23.

Limitações:

- O modelo não considera fatores como localização e idade do imóvel
- Funciona melhor para imóveis com características similares aos da base de dados

Conclusões

A capacidade da garagem é o fator mais determinante no preço

A área do primeiro andar tem forte impacto na valorização

Banheiros adicionais aumentam significativamente o valor

O modelo de regressão linear explica grande parte da variação

É possível estimar preços de novos imóveis com base em suas

características

Referências

Fonte dos Dados:

- Base inspirada em 'House Prices' do Kaggle

Metodologias Utilizadas:

- Análise exploratória de dados
- Correlação de Pearson
- Regressão Linear

Ferramentas e Bibliotecas:

- Python
- Pandas para manipulação de dados
- Matplotlib e Seaborn para visualizações
- Statsmodels para modelagem estatística