ANÁLISE DE PREÇOS DE VEÍCULOS USADOS

A BASE DE DADOS

• Dataset sintético com 1.000.000 de registros de veículos usados.

20 colunas de informação

Contém 25 das marcas de carros mais comuns.

Automotive Price Prediction [

Synthetic dataset of 1M+ vehicles for high-accuracy price prediction.

Data Card Code (4) Discussion (0) Suggestions (0)

About Dataset

Overview

This comprehensive dataset contains 1,000,000 entries for used vehicles, designed s models. The data was synthetically generated using a Python script that establishes re market price. It includes 25 of the most common car brands, covering a wide range of

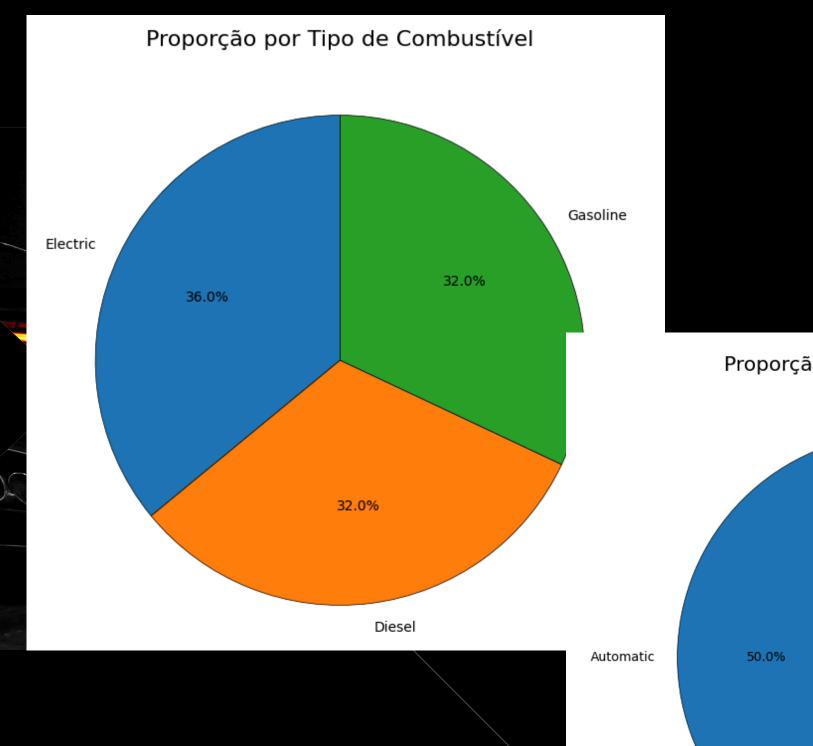
How the Data Was Generated

The dataset was created programmatically. The script's logic ensures reglistic day

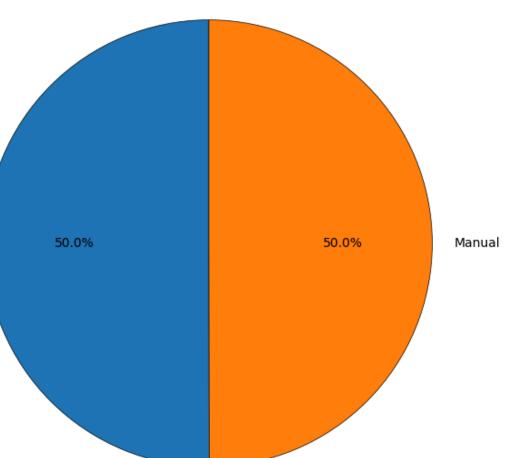
A BASE DE DADOS

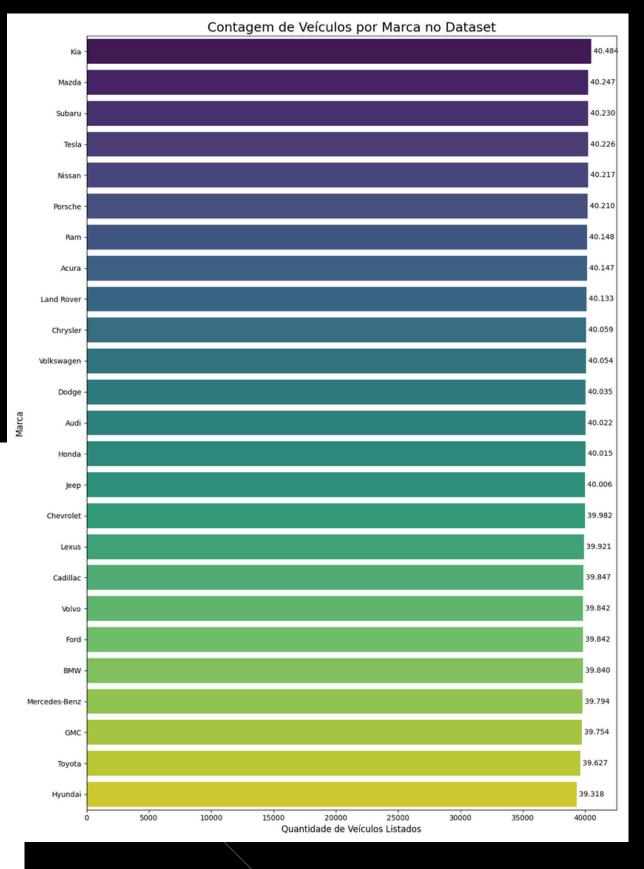
```
base = pd.read_csv(caminho_csv)
    base.columns = base.columns.str.strip().str.replace('\n', '')
    pd.set_option('display.max_columns', None)
    colunas_para_remover = ['year', 'mileage_per_year', 'trim', 'brand_popularity', 'exterior_color', 'interior_color']
    base.drop(columns=colunas_para_remover, inplace=True)
    print("Colunas removidas com sucesso! Colunas restantes no dataset:")
    print(base.columns)
 8
    print("\n--- 1. Verificação de Valores Nulos (NaN) ---")
    print(base.isnull().sum())
    print("\n--- 2. Verificação de Valores Vazios ('') em colunas de texto ---")
    for coluna in base.select_dtypes(include=['object']).columns:
        vazios = (base[coluna] == '').sum()
13
        if vazios > 0:
14
            print(f"Coluna '{coluna}' possui {vazios} valores vazios ('').")
15
    print("\n--- 3. Verificação de Valores Zerados (0) em colunas numéricas ---")
    valores_zerados = (base.select_dtypes(include=['number']) == 0).sum()
    print(valores_zerados[valores_zerados > 0])
    base['accident_history'].fillna('No Accidents', inplace=True)
    print("\n--- Verificação de 'accident_history' após o tratamento ---")
    print(base['accident_history'].value_counts())
    print("\n--- Re-verificação de Valores Nulos ---")
    print(base.isnull().sum())
    print(base.describe())
```

PORCENTAGENS







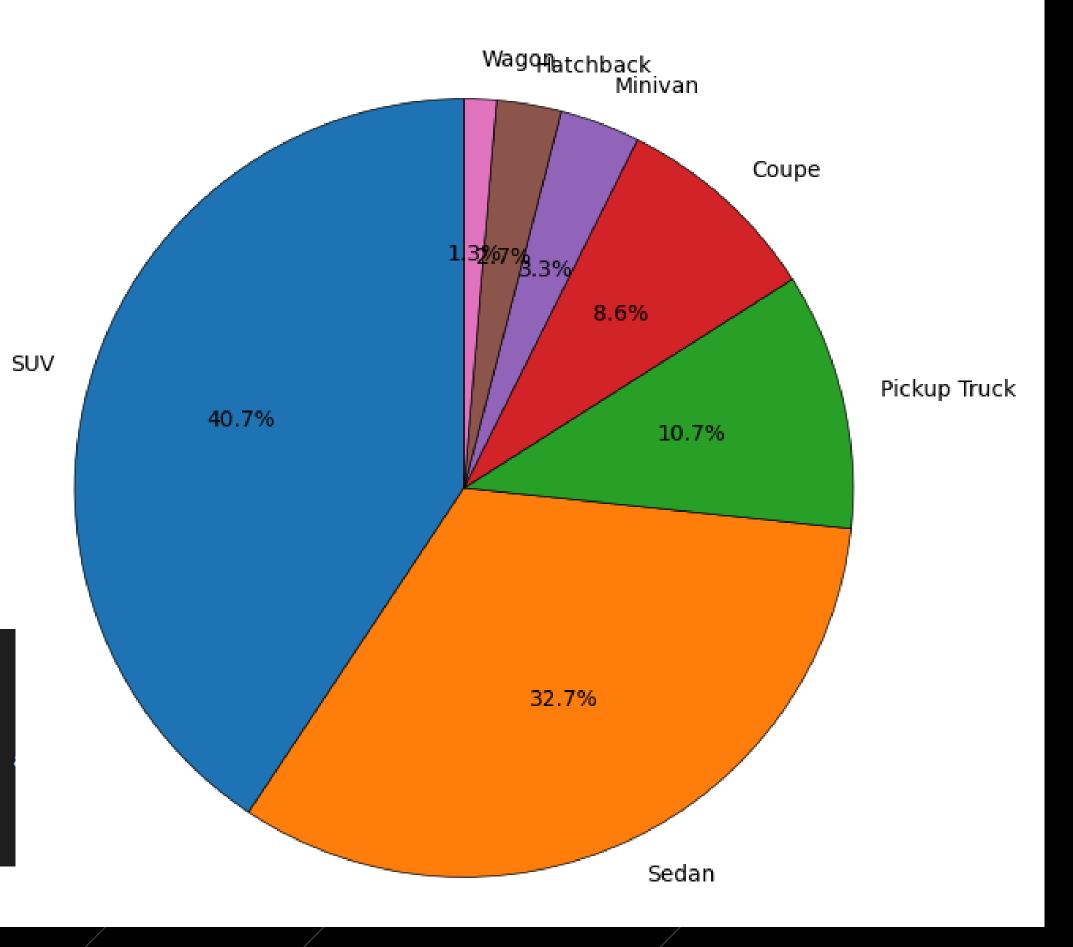


PORCENTAGENS



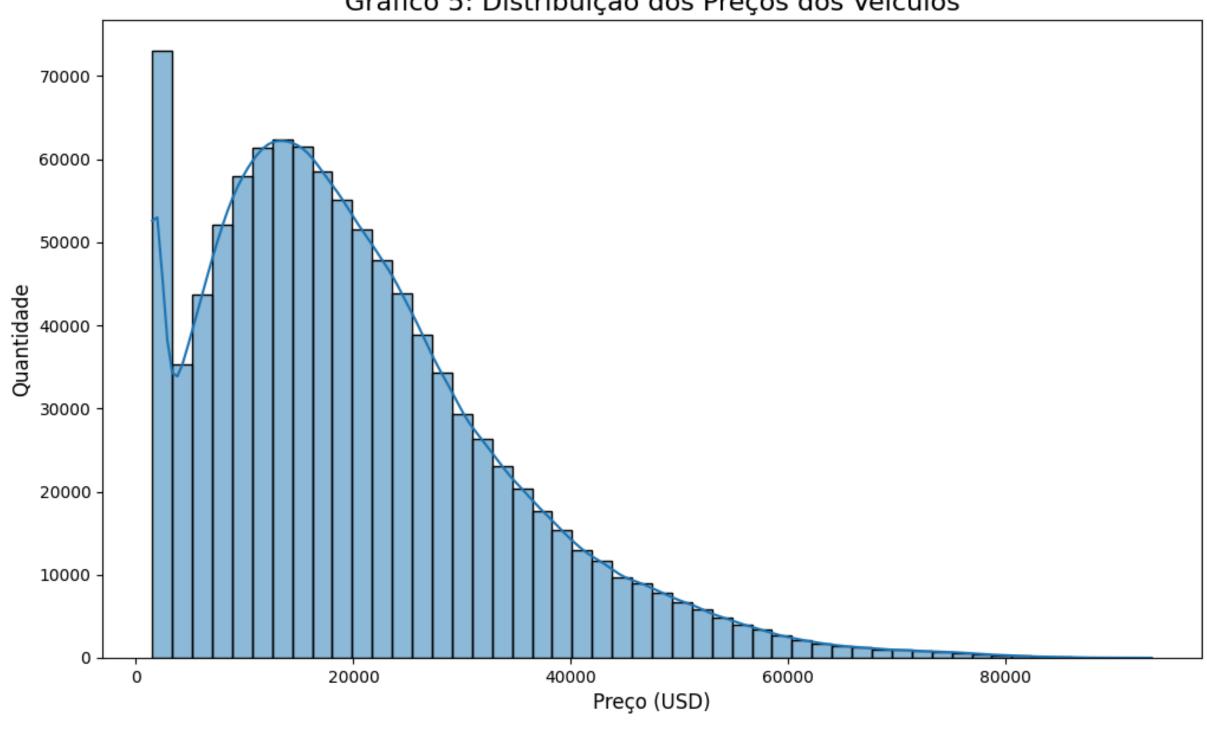
- 1 plt.figure(figsize=(10, 8))
- contagem_carroceria = base['body_type'].value_counts()
- 3 plt.pie(contagem_carroceria, labels=contagem_carroceria.index,
- 4 plt.title('Proporção por Tipo de Carroceria', fontsize=16)
- 5 plt.ylabel('')
- 6 plt.savefig(pasta_destino / '2_pizza_carroceria.png')

Proporção por Tipo de Carroceria



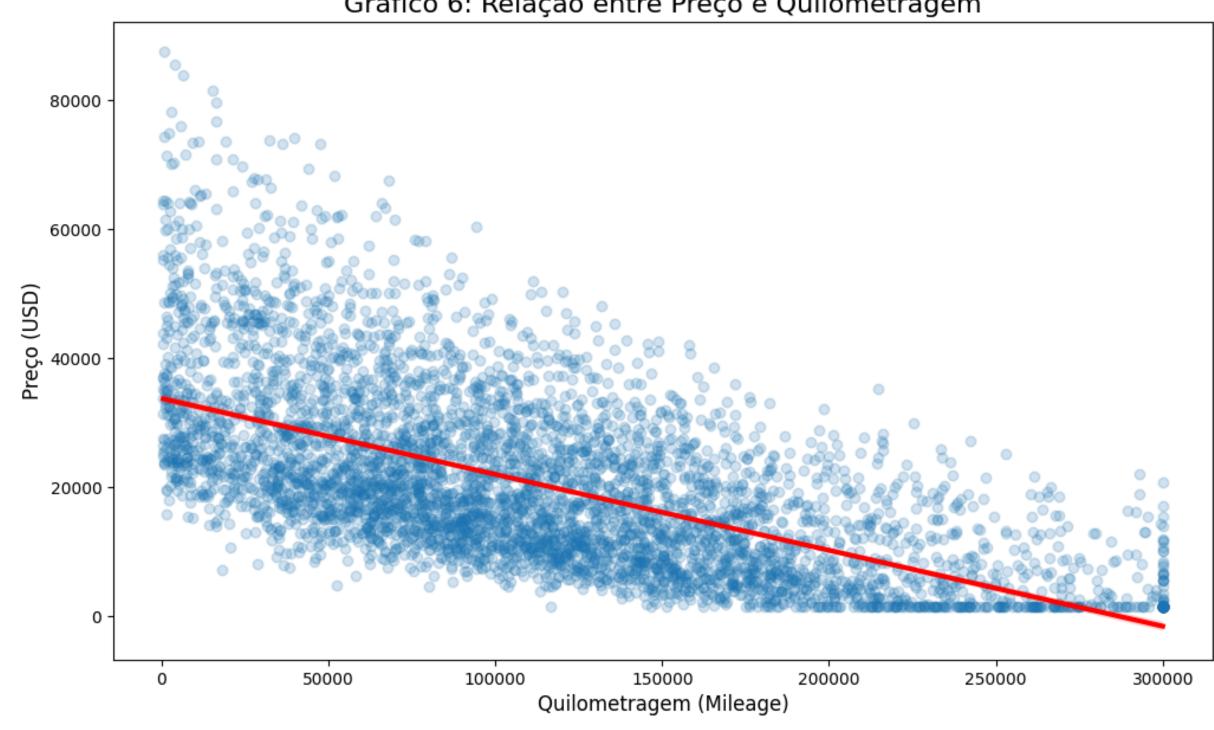
ANALISES DOS DADOS

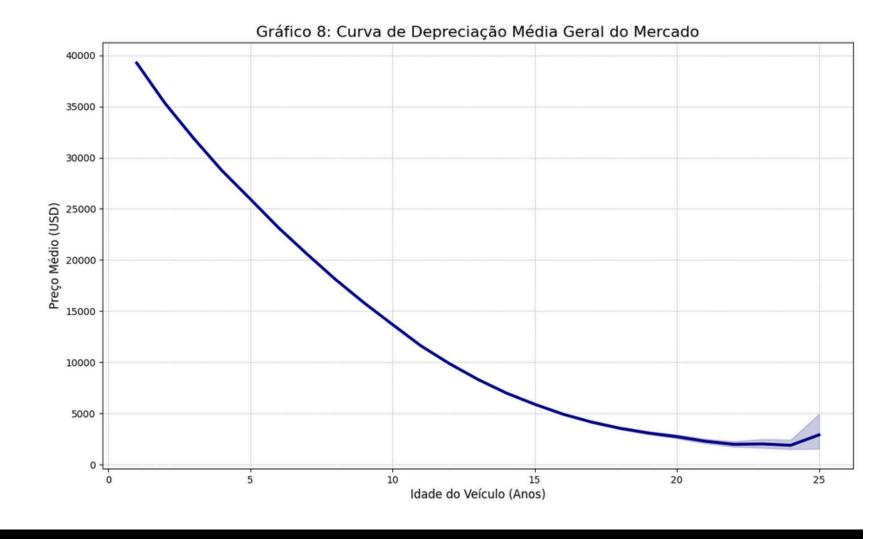
Gráfico 5: Distribuição dos Preços dos Veículos



ANALISES DOS DADOS

Gráfico 6: Relação entre Preço e Quilometragem





ANALISES DOS DADOS

