**MAPA – Material de Avaliação Prática da Aprendizagem**

|  |  |  |
| --- | --- | --- |
| **Acadêmico: Lucas Antonio Rocha Santos** | **R.A. 22134899-5** | |
| **Curso: Ciências de Dados e Análise de Comportamento** | | |
| **Disciplina: Visualização de Dados** | | |
| **Valor da atividade: 3,00** | | **Prazo: 22/09/2023** |

<https://colab.research.google.com/drive/1B4jYxAANOzZmH09T33Qfj656Ww7foIrS?usp=sharing>

**# 1. Importação dos dados: importe o arquivo CSV**

**# que está disponível em https://archive.ics.uci.edu/ml/machine-learning-databases/autos/imports-85.data.**

**#----------------------------------------------------------------------**

# Importando as bibliotecas pandas e matplotlib

import pandas as pd

import matplotlib.pyplot as plt

# Arquivo CSV

dados = "https://archive.ics.uci.edu/ml/machine-learning-databases/autos/imports-85.data"

# Colunas do conjunto de dados

colunas = ["symboling", "normalized-losses", "make", "fuel-type", "aspiration", "num-of-doors", "body-style",

           "drive-wheels", "engine-location", "wheel-base", "length", "width", "height", "curb-weight",

           "engine-type", "num-of-cylinders", "engine-size", "fuel-system", "bore", "stroke", "compression-ratio",

           "horsepower", "peak-rpm", "city-mpg", "highway-mpg", "price"]

# Importando o arquivo CSV para um DataFrame

df = pd.read\_csv(dados, names=colunas, na\_values="?")

# Exibindo as primeiras linhas do DataFrame

df.head()

**#2. Descrição estatística: forneça informações como a contagem de registros,**

**# média, desvio padrão, valores mínimo e máximo e os quartis para cada coluna numérica.**

**#----------------------------------------------------------------------**

# Estatísticas descritivas

descritivas = df.describe()

# Exibindo as estatísticas descritivas

descritivas

**#3. Transformação dos dados: substituia os valores faltantes por 0 e converta a coluna 'price'**

**# para valores numéricos em vez de strings.**

**#----------------------------------------------------------------------**

# Substituindo valores faltantes por 0 em todo o DataFrame

df.fillna(0, inplace=True)

# Convertendo a coluna 'price' para valores numéricos

df['price'] = pd.to\_numeric(df['price'], errors='coerce')

print(df['price'])

**#4. Visualização de dados: mostre a variação de preço (‘price’) em relação ao tamanho do motor ('engine-size')**

**# por intermédio de um gráfico de dispersão.**

**#----------------------------------------------------------------------**

# Criando um gráfico de dispersão

plt.figure(figsize=(10, 6))

plt.scatter(df['engine-size'], df['price'], alpha=0.5)

plt.title('Variação de Preço em Relação ao Tamanho do Motor')

plt.xlabel('Tamanho do Motor')

plt.ylabel('Preço')

plt.grid(True)

plt.show()