

```

# -*- coding: utf-8 -*-
"""Business Analytics Project 3.ipynb

Automatically generated by Colaboratory.

Original file is located at
    https://colab.research.google.com/drive/16t_9QFehuDRix8n3LxCUGfvJWd9oSuuM
"""

#Sai Kiran Mangali
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

file_path = '/content/Electric_Vehicle_Population_Data.csv'
data = pd.read_csv(file_path)

#print(data.describe())

plt.figure(figsize=(8, 6))
sns.countplot(x='Electric Vehicle Type', data=data)
plt.title('Distribution of Electric Vehicle Types')
plt.show()

#sns.pairplot(data=data, vars=['Model Year', 'Electric Range'])
#plt.title('Relationship between Model Year and Electric Range\n')
#plt.show()

from sklearn.model_selection import train_test_split
from sklearn.impute import SimpleImputer
from sklearn.experimental import enable_hist_gradient_boosting
from sklearn.ensemble import HistGradientBoostingClassifier
from sklearn.metrics import accuracy_score, classification_report

selected_features = ['Model Year', 'Electric Range', 'Legislative District', 'DOL
Vehicle ID', '2020 Census Tract'] # Replace 'Other_Column' with actual column names
X = data[selected_features]
y = data['Clean Alternative Fuel Vehicle (CAFV) Eligibility']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
imputer = SimpleImputer(strategy='mean')
X_train_imputed = imputer.fit_transform(X_train)
X_test_imputed = imputer.transform(X_test)
model = HistGradientBoostingClassifier()
model.fit(X_train_imputed, y_train)
predictions = model.predict(X_test_imputed)

accuracy = accuracy_score(y_test, predictions)
report = classification_report(y_test, predictions)
print(f"Accuracy: {accuracy}")
print(f"Classification Report:\n {report}")

```