**# Import Libraries**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.model\_selection import train\_test\_split

from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import mean\_absolute\_error

**# Import CSV File**

housing\_data=pd.read\_csv('Housing.csv')

housing\_data

**# Preprocessing in wrangler**

Convert string to Boolean

hd\_data=pd.read\_csv('hd.csv')

hd\_data

**# Separate features and target**

X = hd\_data.drop('price', axis=1)

y = hd\_data['price']

**# Split the Data**

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

 Initialize the model

rf = RandomForestRegressor(

n\_estimators=100,

max\_depth=None,

random\_state=42

)

**# Fit the model**

rf.fit(X\_train, y\_train)

Predict on test data

y\_pred = rf.predict(X\_test)

**# Evaluation**

mae = mean\_absolute\_error(y\_test, y\_pred)

print("Mean Absolute Error (MAE):", mae)

mpae = np.mean(np.abs((y\_test - y\_pred) / y\_test)) \* 100

print("Mean Percentage Absolute Error (MPAE):", mpae)

**# Visualisation**

**# Correlation heatmap**

corr = hd\_data.corr()

plt.figure(figsize=(12, 8))

sns.heatmap(corr[['price']].sort\_values(by='price', ascending=False), annot=True, cmap='coolwarm')

plt.title("Features Correlated with price")

plt.show()

**# Box Plot:**

**# Preferred Area vs. Price**

sns.boxplot(x='prefarea', y='price', data=hd\_data)

plt.title('Price by Preferred Area')

plt.show()

**# Stories vs. Price**

sns.boxplot(x='stories', y='price', data=hd\_data)

plt.title('Price by stories')

plt.show()

**# Bedrooms vs. Price**

sns.boxplot(x='bedrooms', y='price', data=hd\_data)

plt.title('Price by bedrooms')

plt.show()

**# Bathrooms vs. Price**

sns.boxplot(x='bathrooms', y='price', data=hd\_data)

plt.title('Price by bathrooms')

plt.show()