import pandas as pd  
import numpy as np  
  
from sklearn.model\_selection import train\_test\_split  
from sklearn.linear\_model import LinearRegression  
from sklearn.metrics import r2\_score, mean\_squared\_error

df = pd.read\_csv("../data/Housing.csv")  
  
X = df[['area']]  
y = df['price']

x\_train, x\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.3, random\_state=42)

lm = LinearRegression()  
  
lm.fit(x\_train, y\_train)

LinearRegression()

y\_pred = lm.predict(x\_test)  
  
mse = mean\_squared\_error(y\_test, y\_pred)  
rmse = np.sqrt(mse)  
  
print(f"Mean Squared Error: {mse:.2f}")  
print(f"Root Mean Squared Error: {rmse:.2f}")  
print(f"R-squared: {r2\_score(y\_test, y\_pred):.2f}")

Mean Squared Error: 3193989923757.62  
Root Mean Squared Error: 1787173.73  
R-squared: 0.26

new\_house\_area = 3000  
  
new\_house\_price = lm.predict([[new\_house\_area]])  
print(f"Estimated price for a new house with an area of {new\_house\_area} square feet: ${new\_house\_price[0]:.2f}")

Estimated price for a new house with an area of 3000 square feet: $3805802.55

/opt/homebrew/lib/python3.11/site-packages/sklearn/base.py:464: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names  
 warnings.warn(