26.Question 3: Linear Regression for Housing Price Prediction

You are a real estate analyst trying to predict housing prices based on various features of the

houses, such as area, number of bedrooms, and location. You have collected a dataset of houses

with their respective prices.

Write a Python program that allows the user to input the features (area, number of bedrooms, etc.)

of a new house. The program should use linear regression from scikit-learn to predict the price of

the new house based on the input features.

Code :

import pandas as pd

from sklearn.linear\_model import LinearRegression

from sklearn.preprocessing import OneHotEncoder

from sklearn.compose import ColumnTransformer

from sklearn.pipeline import Pipeline

# Load data from Excel

data = pd.read\_excel(r"C:\Users\jampa\Downloads\q26\_05.xlsx")

# Clean column names safely

data.columns = [str(col).strip() for col in data.columns]

# Check required columns

required\_columns = ['Area (sqft)', 'Bedrooms', 'Location', 'Price ($)']

for col in required\_columns:

if col not in data.columns:

raise ValueError(f"Missing required column: {col}")

# Features and Target

X = data[['Area (sqft)', 'Bedrooms', 'Location']]

y = data['Price ($)']

# Preprocessing for categorical data

preprocessor = ColumnTransformer(

transformers=[('cat', OneHotEncoder(), ['Location'])],

remainder='passthrough'

)

# Create pipeline

model = Pipeline([

('preprocessor', preprocessor),

('regressor', LinearRegression())

])

# Train model

model.fit(X, y)

# Get user input for prediction

area = float(input("Enter area (in sqft): "))

bedrooms = int(input("Enter number of bedrooms: "))

location = input("Enter location (e.g., Downtown, Suburb, Rural): ")

# Create input for prediction

new\_house = pd.DataFrame([[area, bedrooms, location]], columns=['Area (sqft)', 'Bedrooms', 'Location'])

# Predict

predicted\_price = model.predict(new\_house)[0]

print(f"\nPredicted House Price: ${predicted\_price:,.2f}")

output :

Enter area (in sqft): 56

Enter number of bedrooms: 5

Enter location (e.g., Downtown, Suburb, Rural): Downtown

Predicted House Price: $153,646.84

Dataset :

|  |  |  |  |
| --- | --- | --- | --- |
| **Area (sqft)** | **Bedrooms** | **Location** | **Price ($)** |
| 1400 | 3 | Downtown | 300000 |
| 1600 | 3 | Suburb | 280000 |
| 1700 | 4 | Suburb | 310000 |
| 1875 | 3 | Downtown | 340000 |
| 1100 | 2 | Rural | 190000 |
| 1550 | 3 | Suburb | 295000 |
| 2350 | 4 | Downtown | 405000 |
| 2450 | 4 | Downtown | 430000 |
| 1425 | 3 | Rural | 220000 |
| 1700 | 3 | Suburb | 310000 |