This Python code demonstrates a basic machine learning workflow for predicting house prices using a linear regression model. Below is a documentary explanation of each part of the code:

1. ***Im*****port necessary libraries** :

- `pandas` for data manipulation
- `train_test_split` from `sklearn.model_selection` for splitting the dataset into training and testing sets
- `LinearRegression` from `sklearn.linear_model` for building a linear regression model
- `mean_squared_error` from `sklearn.metrics` to calculate the mean squared error
- `matplotlib.pyplot` for creating a scatter plot

2. **Load the dataset** :

- Reads a CSV file named 'house_data.csv' into a Pandas DataFrame called 'data.'

3. **Handle categorical variables:**

- Converts categorical variables ('Location' and 'Zip_Code') into numerical form using one-hot encoding.

4. **Select features and target variable**:

- Separates the dataset into features (X) and the target variable (y). The target variable is 'Price,' and the features are all other columns.

5. **Split the data**:

- Splits the data into training and testing sets using `train_test_split`. It reserves 30% of the data for testing and

Uses a random seed (random_state) for reproducibility.

6. **Initialize and train the linear regression model**:

• Creates an instance of the LinearRegression model and

Trains it using the training data (X_train and y_train).

7. **Make predictions** :

• Uses the trained model to make price predictions on the

Testing data, storing them in the ‘predictions’ variable.

8. Create a DataFrame for actual vs. predicted prices :

• Compiles the actual prices from the test data and the

Predicted prices into a Pandas DataFrame called ‘results.’

9. **Save the results to a CSV file**:

• Saves the ‘results’ DataFrame to a CSV file named

‘predicted_prices.csv’ without including the index.

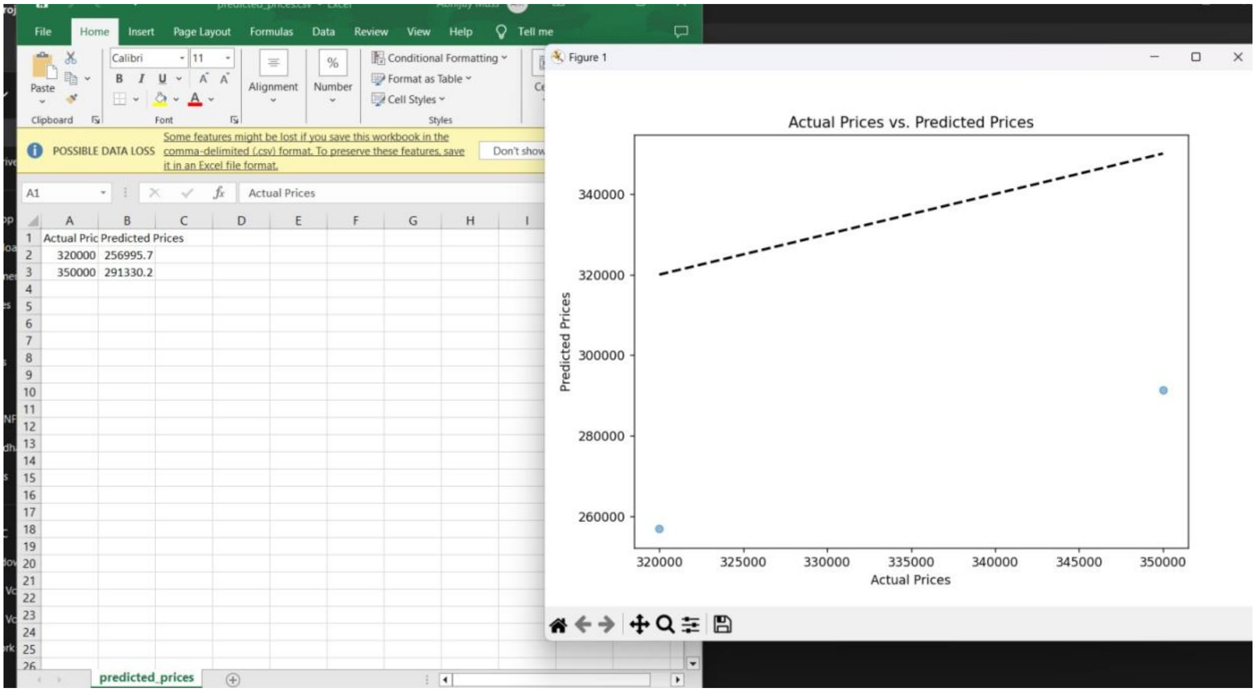
10. ***Generate a scatter plot:***

•Creates a scatter plot using Matplotlib to visualize the

Relationship between actual prices and predicted prices.

•The diagonal line represents perfect predictions; data points Should be close

OUTPUT :



POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format. Don't show

A1 Actual Prices

	A	B
1	Actual Pric	Predicted Prices
2	320000	256995.7
3	350000	291330.2