```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score


df = pd.read_csv("/content/Housing.csv")


df
```

	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basement	hot
0	13300000	7420	4	2	3	yes	no	no	
1	12250000	8960	4	4	4	yes	no	no	
2	12250000	9960	3	2	2	yes	no	yes	
3	12215000	7500	4	2	2	yes	no	yes	
4	11410000	7420	4	1	2	yes	yes	yes	
...	...	...	...	...	...	...	...	...	
540	1820000	3000	2	1	1	yes	no	yes	
541	1767150	2400	3	1	1	no	no	no	
542	1750000	3620	2	1	1	yes	no	no	
543	1750000	2910	3	1	1	no	no	no	
544	1750000	3850	3	1	2	yes	no	no	

545 rows × 13 columns

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
print(df.head())
```

```
      price  area  bedrooms  bathrooms  stories mainroad guestroom basement  \
0  13300000  7420         4          2        3      yes        no       no
1  12250000  8960         4          4        4      yes        no       no
2  12250000  9960         3          2        2      yes        no      yes
3  12215000  7500         4          2        2      yes        no      yes
4  11410000  7420         4          1        2      yes       yes      yes

  hotwaterheating airconditioning  parking prefarea furnishingstatus
0              no             yes        2      yes        furnished
1              no             yes        3       no        furnished
2              no              no        2      yes   semi-furnished
3              no             yes        3      yes        furnished
4              no             yes        2       no        furnished
```

```
df = pd.get_dummies(df, drop_first=True)
```

```
X = df.drop("price", axis=1)
y = df["price"]
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
df
```

	price	area	bedrooms	bathrooms	stories	parking	mainroad_yes	guestroom_y
0	13300000	7420	4	2	3	2	True	Fal
1	12250000	8960	4	4	4	3	True	Fal
2	12250000	9960	3	2	2	2	True	Fal
3	12215000	7500	4	2	2	3	True	Fal
4	11410000	7420	4	1	2	2	True	Tr
...	...	...	...	...	...	...	...	
540	1820000	3000	2	1	1	2	True	Fal
541	1767150	2400	3	1	1	0	False	Fal
542	1750000	3620	2	1	1	0	True	Fal
543	1750000	2910	3	1	1	0	False	Fal
544	1750000	3850	3	1	2	0	True	Fal

545 rows × 14 columns

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
model = LinearRegression()
model.fit(X_train, y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
y_pred = model.predict(X_test)
```

```
df
```

	price	area	bedrooms	bathrooms	stories	parking	mainroad_yes	guestroom_y
0	13300000	7420	4	2	3	2	True	Fal
1	12250000	8960	4	4	4	3	True	Fal
2	12250000	9960	3	2	2	2	True	Fal
3	12215000	7500	4	2	2	3	True	Fal
4	11410000	7420	4	1	2	2	True	Tr
...	...	...	...	...	...	...	...	
540	1820000	3000	2	1	1	2	True	Fal
541	1767150	2400	3	1	1	0	False	Fal
542	1750000	3620	2	1	1	0	True	Fal
543	1750000	2910	3	1	1	0	False	Fal
544	1750000	3850	3	1	2	0	True	Fal

545 rows × 14 columns

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
mae = mean_absolute_error(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)
rmse = np.sqrt(mse)
r2 = r2_score(y_test, y_pred)


print("\nModel Evaluation:")
print(f"MAE: {mae}")
print(f"MSE: {mse}")
print(f"RMSE: {rmse}")
print(f"R² Score: {r2}")
```

```
Model Evaluation:
MAE: 970043.4039201636
MSE: 1754318687330.6638
RMSE: 1324506.9600914386
R² Score: 0.6529242642153184
```

```
if X_train.shape[1] == 1:
    plt.scatter(X_test, y_test, color='blue', label='Actual')
    plt.plot(X_test, y_pred, color='red', linewidth=2, label='Predicted')
    plt.xlabel("Feature")
    plt.ylabel("Price")
    plt.legend()
    plt.show()
```

```
df
```

	price	area	bedrooms	bathrooms	stories	parking	mainroad_yes	guestroom_y
0	13300000	7420	4	2	3	2	True	Fal
1	12250000	8960	4	4	4	3	True	Fal
2	12250000	9960	3	2	2	2	True	Fal
3	12215000	7500	4	2	2	3	True	Fal
4	11410000	7420	4	1	2	2	True	Tr
...	...	...	...	...	...	...	...	
540	1820000	3000	2	1	1	2	True	Fal
541	1767150	2400	3	1	1	0	False	Fal
542	1750000	3620	2	1	1	0	True	Fal
543	1750000	2910	3	1	1	0	False	Fal
544	1750000	3850	3	1	2	0	True	Fal

545 rows × 14 columns

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
print("\nModel Coefficients:")
coeff_df = pd.DataFrame(model.coef_, X.columns, columns=['Coefficient'])
print(coeff_df)
```

```
Model Coefficients:
                                  Coefficient
area                             2.359688e+02
bedrooms                         7.677870e+04
bathrooms                        1.094445e+06
stories                          4.074766e+05
parking                          2.248419e+05
mainroad_yes                     3.679199e+05
guestroom_yes                    2.316100e+05
basement_yes                     3.902512e+05
hotwaterheating_yes              6.846499e+05
airconditioning_yes              7.914267e+05
prefarea_yes                     6.298906e+05
furnishingstatus_semi-furnished -1.268818e+05
furnishingstatus_unfurnished    -4.136451e+05
```