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

<b>→</b> *		nrice	area	bedrooms	bathrooms	stories	mainroad	guestroom	hasement	hot
		pi icc	ai ca	bear ooms	ba cili ooms	3001103	main oaa	gueser com	baschiene	1100
	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 rc	ws × 13 colu	umns							

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

print(df.head())

2

3

4

<b>→</b> ▼		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	
		hotwaterhe	ating	airconditio	ning park:	ing prefa	area furni	İshingstatı	ıs	
	0		no		yes	2	ves	furnishe		
			110		y c s	_	yes			
	1		no		yes	3	no	furnishe	ed	

2

3

2

yes

yes

no

semi-furnished

furnished

furnished

no

yes

yes

no

no

no

```
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

 $\rightarrow$ 

						guestroom_y
00 7420	4	2	3	2	True	Fal
00 8960	4	4	4	3	True	Fal
00 9960	3	2	2	2	True	Fal
00 7500	4	2	2	3	True	Fal
00 7420	4	1	2	2	True	Tr
00 3000	2	1	1	2	True	Fal
50 2400	3	1	1	0	False	Fal
00 3620	2	1	1	0	True	Fal
00 2910	3	1	1	0	False	Fal
00 3850	3	1	2	0	True	Fal
	9960 9960 7500 7420  90 3000 50 2400 90 3620	00       8960       4         00       9960       3         00       7500       4         00       7420       4              00       3000       2         50       2400       3         00       3620       2         00       2910       3	00       8960       4       4         00       9960       3       2         00       7500       4       2         00       7420       4       1               00       3000       2       1         50       2400       3       1         00       3620       2       1         00       2910       3       1	00       8960       4       4       4         00       9960       3       2       2         00       7500       4       2       2         00       7420       4       1       2                00       3000       2       1       1         50       2400       3       1       1         00       3620       2       1       1         00       2910       3       1       1	00       8960       4       4       4       4       3         00       9960       3       2       2       2         00       7500       4       2       2       3         00       7420       4       1       2       2                 00       3000       2       1       1       2         50       2400       3       1       1       0         00       3620       2       1       1       0         00       2910       3       1       1       0	00       8960       4       4       4       4       3       True         00       9960       3       2       2       2       2       True         00       7500       4       2       2       3       True         00       7420       4       1       2       2       True         00       3000       2       1       1       2       True         50       2400       3       1       1       0       False         00       3620       2       1       1       0       True         00       2910       3       1       1       0       False

545 rows × 14 columns

Next steps: Generate code with df Vie

View recommended plots

**New interactive sheet** 

model = LinearRegression()
model.fit(X\_train, y\_train)

**→** 

v LinearRegression (i) (?)
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

Next steps: Generate code with df

545 rows × 14 columns

```
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}")
\overline{\mathbf{x}}
     Model Evaluation:
     MAE: 970043.4039201636
     MSE: 1754318687330.6638
     RMSE: 1324506.9600914386
     R<sup>2</sup> 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

		_
_	⇁	٦

	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



New interactive sheet

print("\nModel Coefficients:")
coeff\_df = pd.DataFrame(model.coef\_, X.columns, columns=['Coefficient'])
print(coeff\_df)



## Model Coefficients:

Houce 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
${\tt furnishingstatus\_semi-furnished}$	-1.268818e+05
furnishingstatus_unfurnished	-4.136451e+05