🏠 Regression problem

Predict a house's price based on its features (square footage, number of bedrooms, location).

Import modules

```
import pandas as pd
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
```

Sample data

```
## sample data
data = {
    'sq_footage': [1500, 2000, 1800, 2500, 1600],
    'bedrooms': [3, 4, 3, 5, 3],
    'price': [300000, 450000, 380000, 550000, 320000]
}
df = pd.DataFrame(data)
df
```

	sq_footage	bedrooms	price
0	1500	3	300000
1	2000	4	450000
2	1800	3	380000
3	2500	5	550000
4	1600	3	320000

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

Prepare data

```
## prepare data
X = df[['sq_footage', 'bedrooms']] # Features ตัวแปรต้น
y = df['price'] # Target variable  ตัวแปรเป้าหมาย
```

Split data

```
## split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Train model

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

```
LinearRegression
LinearRegression()
```

Prediction

```
# Make a prediction on new data
# ทำนายราคาจากข้อมูลใหม่
new_house = [[2200, 4]]
predicted_price = model.predict(new_house)
```

```
/usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but Line
  warnings.warn(
```

Print result `new_house`

```
## print result [sq_footage:2200, bedrooms: 4]
print(f"Predicted Price: ${predicted_price[0]:,.2f}")
```

```
Predicted Price: $477,857.14
```

```
Start coding or generate with AI.
```