# Regression problem

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

\sunsun-datateathyme\

## Import modules

```
1 Start coding or generate with AI.

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

### ∨ Sample data

```
1 ## sample data
2 data = {
3     'sq_footage': [1500, 2000, 1800, 2500, 1600],
4     'bedrooms': [3, 4, 3, 5, 3],
5     'price': [300000, 450000, 380000, 550000, 320000]
6 }
7 df = pd.DataFrame(data)
8 df
```

<del>_</del>		sq_footage	bedrooms	price	
	0	1500	3	300000	ılı
	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

```
1 ## prepare data
2 X = df[['sq_footage', 'bedrooms']] # Features ตัวแปรตัน
3 y = df['price'] # Target variable ตัวแปรเป้าหมาย
```

### Split data

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

#### Train model

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

```
LinearRegression (1)?
```

#### Prediction

```
1 # Make a prediction on new data
2 # ทำนายราคาจากข้อมูลใหม่
```

```
3 new_house = [[2200, 4]]
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

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

→ Predicted Price: \$477,857.14

1 Start coding or generate with AI.