Multiple Linear Regression

Example Dataset

Suppose we want to predict **Salary** based only on **Years of Experience**.

Years of Experience (X)	Salary (Y)
1	25,000
2	30,000
3	35,000
4	45,000
5	50,000

Experience	Education	Skills Score	Salary
3	Bachelor	60	35k
3	Master	80	50k

Problem with linear regression

📌 Geometric Understanding

- Simple Linear Regression → 2D Line
- Multiple Linear Regression (2 features) → 3D Plane
- 3+ features → Hyperplane

Multiple Linear Regression is a statistical method used to predict **one dependent variable (Y)** using **two or more independent variables (X_1, X_2, X_3 ...)**.

General Formula:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + ... + b_nX_n$$

Where:

- Y = Target (dependent variable)
- $X_1, X_2, ..., X_n$ = Features
- b_0 = Intercept
- $b_1, b_2, ..., b_n$ = Coefficients (weights)

2️⃣ Example Dataset

Suppose we want to predict **House Price**.

Size (sqft) X_1	Bedrooms X_2	Age (years) X_3	Price (Y)
1000	2	10	200k
1500	3	5	300k
1800	3	2	360k
2000	4	1	400k

Model becomes:

$$Price = b_0 + b_1(Size) + b_2(Bedrooms) + b_3(Age)$$

Each coefficient tells:

- b_1 → How much price increases per sqft
- b_2 → How much price increases per bedroom
- b_3 → How much price changes per year of age

use ordinary least square method for finding theb coefficient

Using matrix form:

$$\beta = (X^T X)^{-1} X^T Y$$

Where:

- X = Feature matrix
- Y = Target vector
- β = Coefficient vector (b_0, b_1, b_2...)

✅ Example: Multiple Linear Regression (2 Features)

Suppose we have this dataset:

X_1	X_2	Y
1	1	6
1	2	8
2	2	9

We want to find:

$$Y = b_0 + b_1 X_1 + b_2 X_2$$

🔵 Step 1: Create Matrix X

We must include the intercept column (1's).

$$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

First column → intercept

Second → X_1

Third → X_2

Step 2: Create Y vector

$$Y = \begin{bmatrix} 6 \\ 8 \\ 9 \end{bmatrix}$$

Step 3: Compute X^T

Transpose means rows become columns.

$$X^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

(In this example it looks similar because of symmetry.)

Step 4: Compute $X^T X$

$$X^T X = \begin{bmatrix} 3 & 4 & 5 \\ 4 & 6 & 7 \\ 5 & 7 & 9 \end{bmatrix}$$

Step 5: Compute $(X^T X)^{-1}$

Inverse of that matrix becomes:

$$(X^T X)^{-1} = \begin{bmatrix} 5 & -1 & -2 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

Step 6: Compute X^TY

$$X^TY = \begin{bmatrix} 23 \\ 32 \\ 40 \end{bmatrix}$$

Step 7: Final Calculation

$$\beta = (X^TX)^{-1}X^TY$$

Multiply matrices:

$$\beta = \begin{bmatrix} 5 & -1 & -2 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 23 \\ 32 \\ 40 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

🎯 Final Model

$$Y = 3 + 1X_1 + 2X_2$$

So:

- $b_0 = 3$
- $b_1 = 1$
- $b_2 = 2$

Data Point 1:

$X_1 = 1, X_2 = 1$

$$Y = 3 + (1)(1) + (2)(1)$$

$$Y = 3 + 1 + 2 = 6$$

Actual Y = 6 ✅

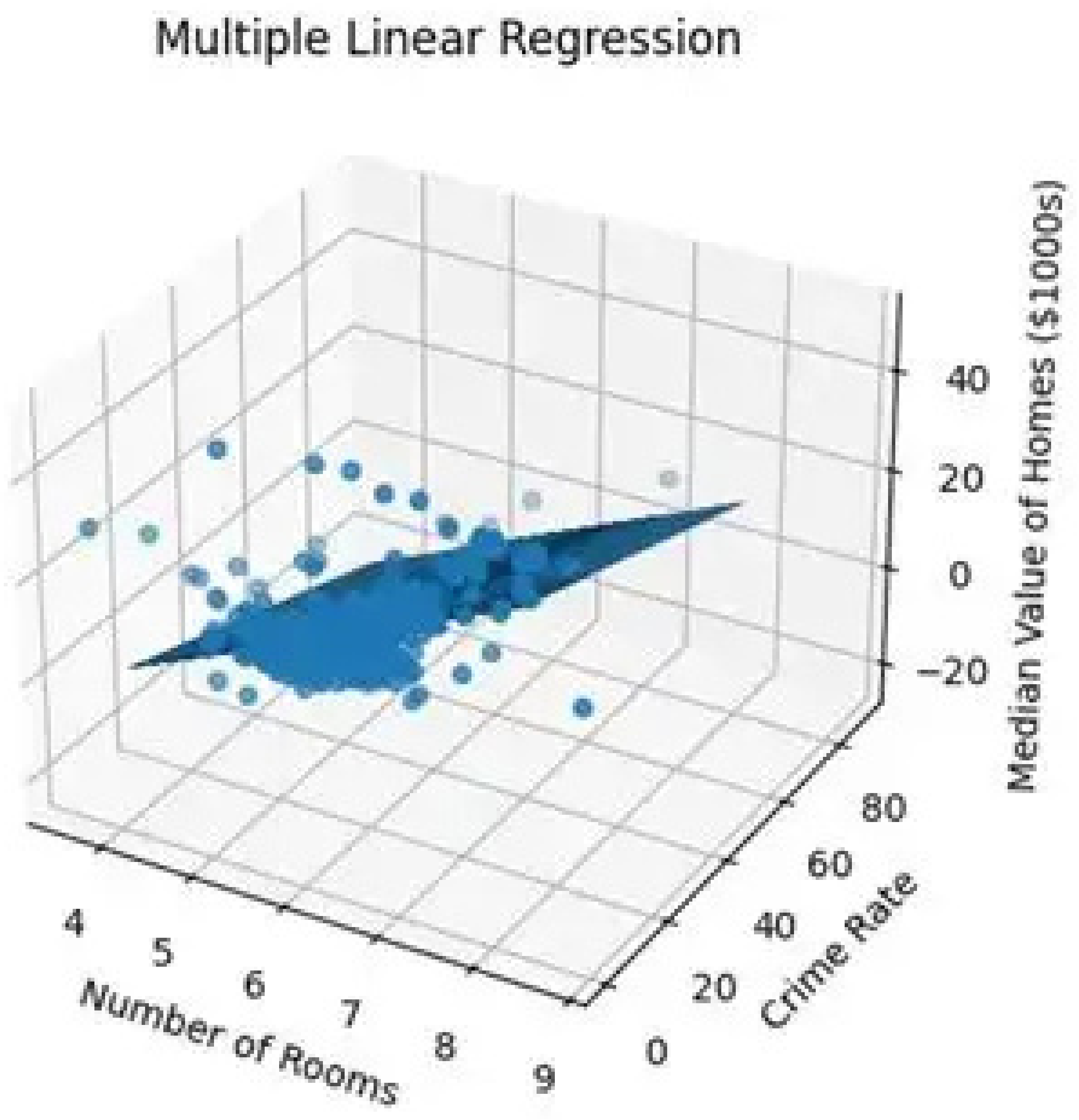
Data Point 2:

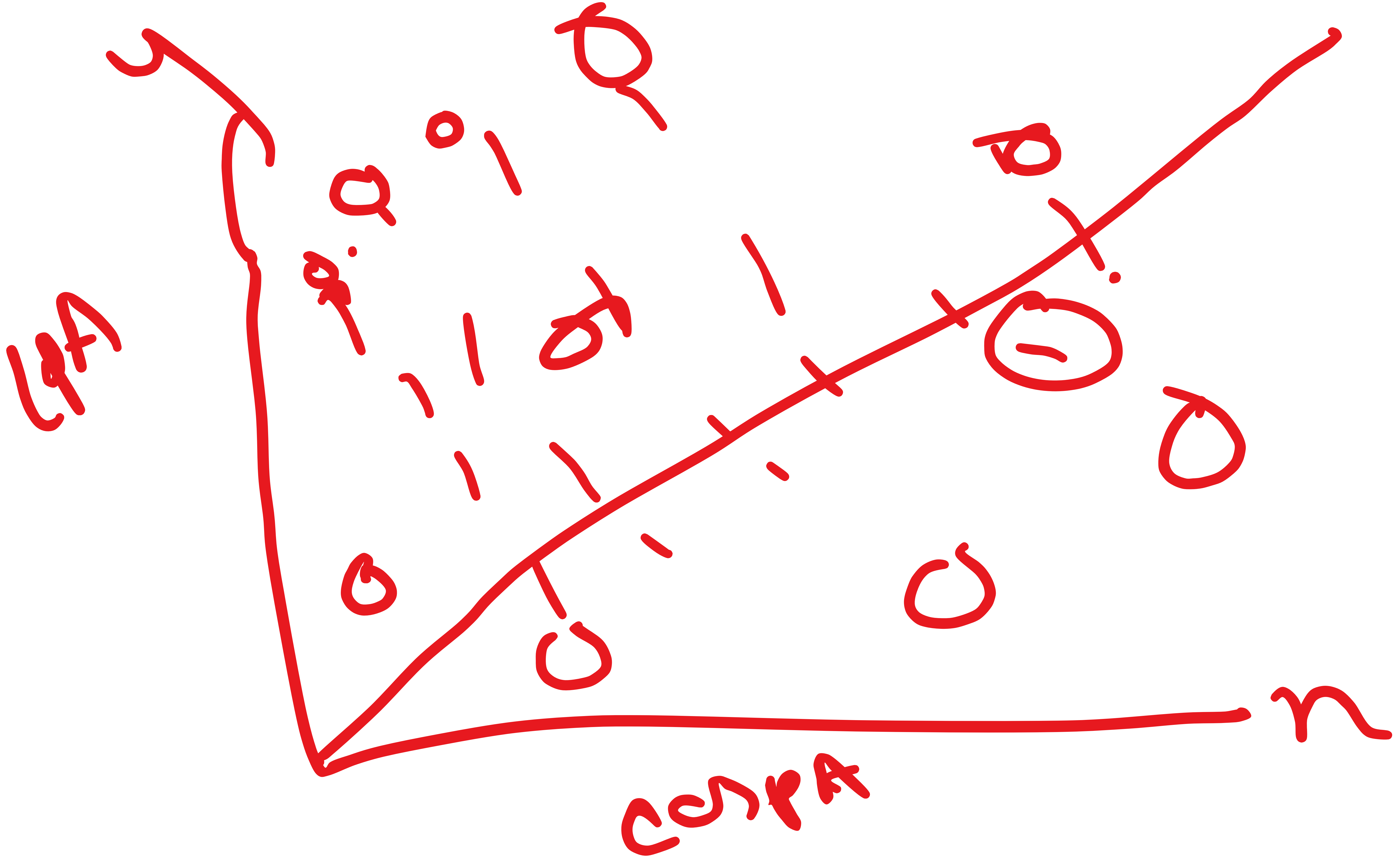
$X_1 = 1, X_2 = 2$

$$Y = 3 + (1)(1) + (2)(2)$$

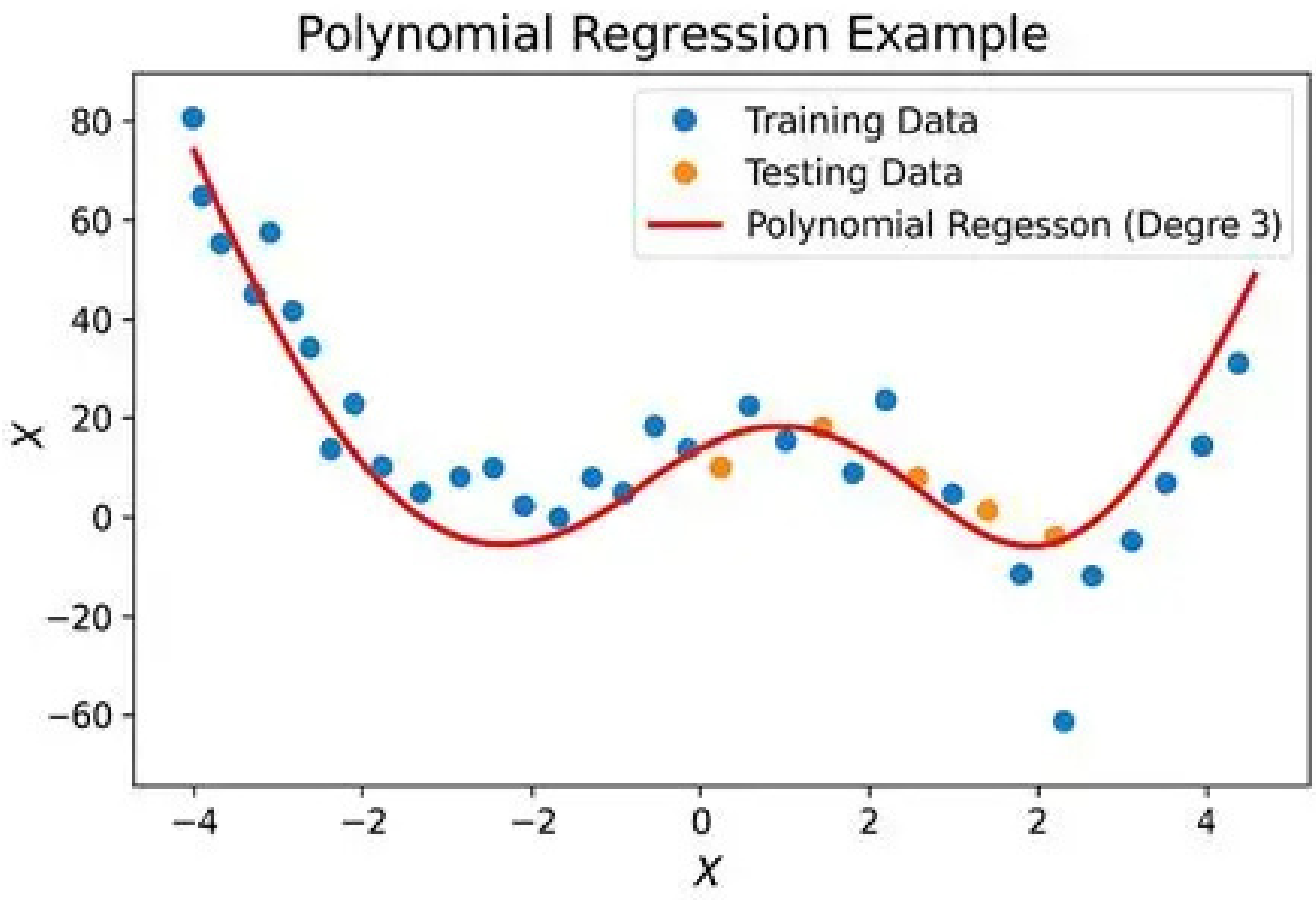
$$Y = 3 + 1 + 4 = 8$$

Actual Y = 8 ✅


Multiple Linear Regression
Median Value of Homes ($1000s)
40
20
0
−20
80
60
40
20
0
Crime Rate
4
5
6
7
8
9
Number of Rooms



y


Polynomial Regression Example
Training Data
Testing Data
Polynomial Regesson (Degre 3)
80
60
40
20
0
−20
−60
X
−4
−2
−2
0
2
2
4
X

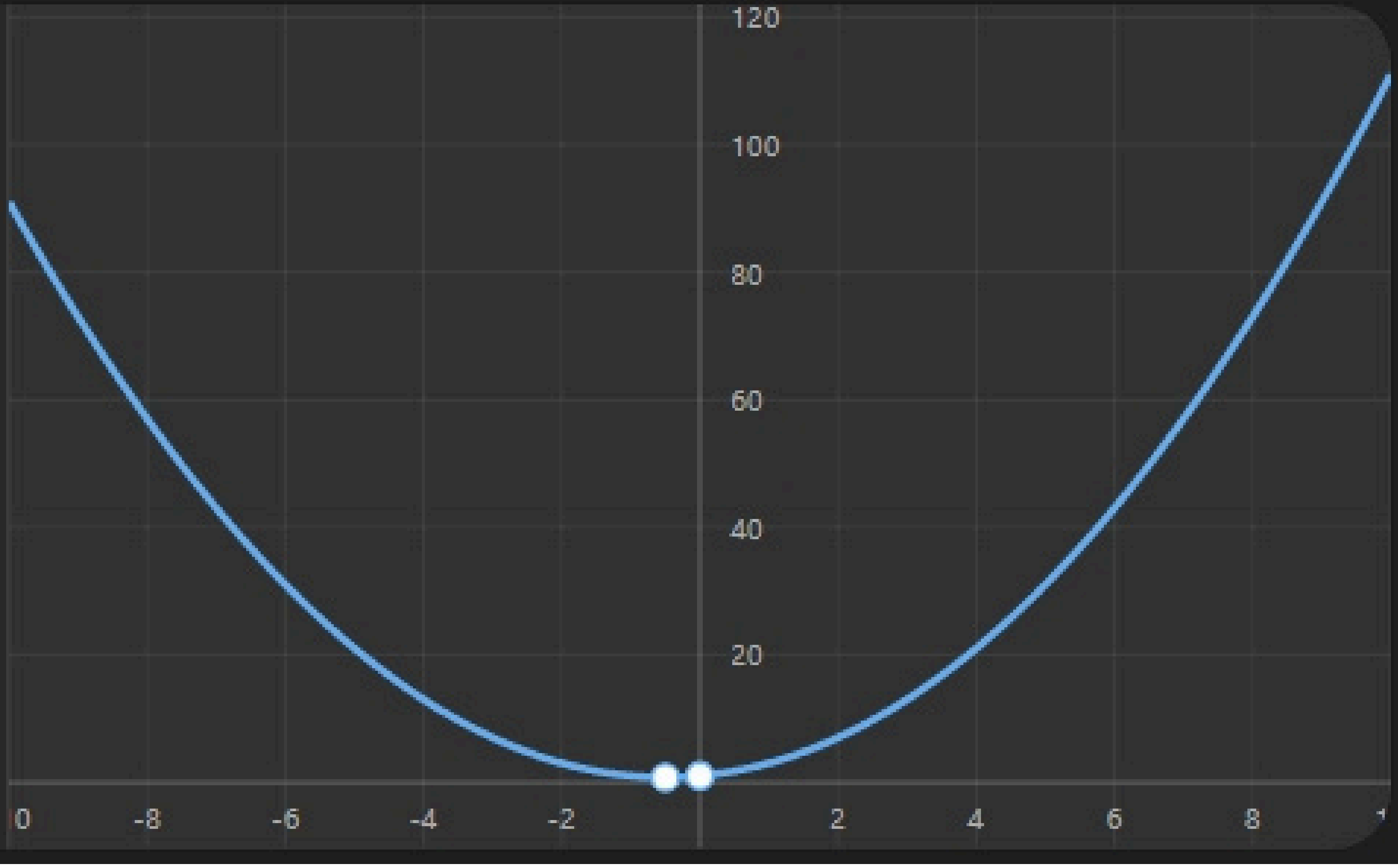
For **degree 2** (quadratic):

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

$beta_0$

$beta_1$

$beta_2$



. Polynomial regression also adds **interaction between columns** → columns can **work together** to bend the curve.

- Example: x1*x2 allows the effect of x1 and x2 combined to influence y.