
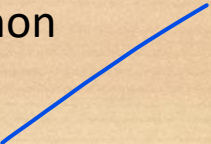


# Linear Regression with **Multiple** Variable

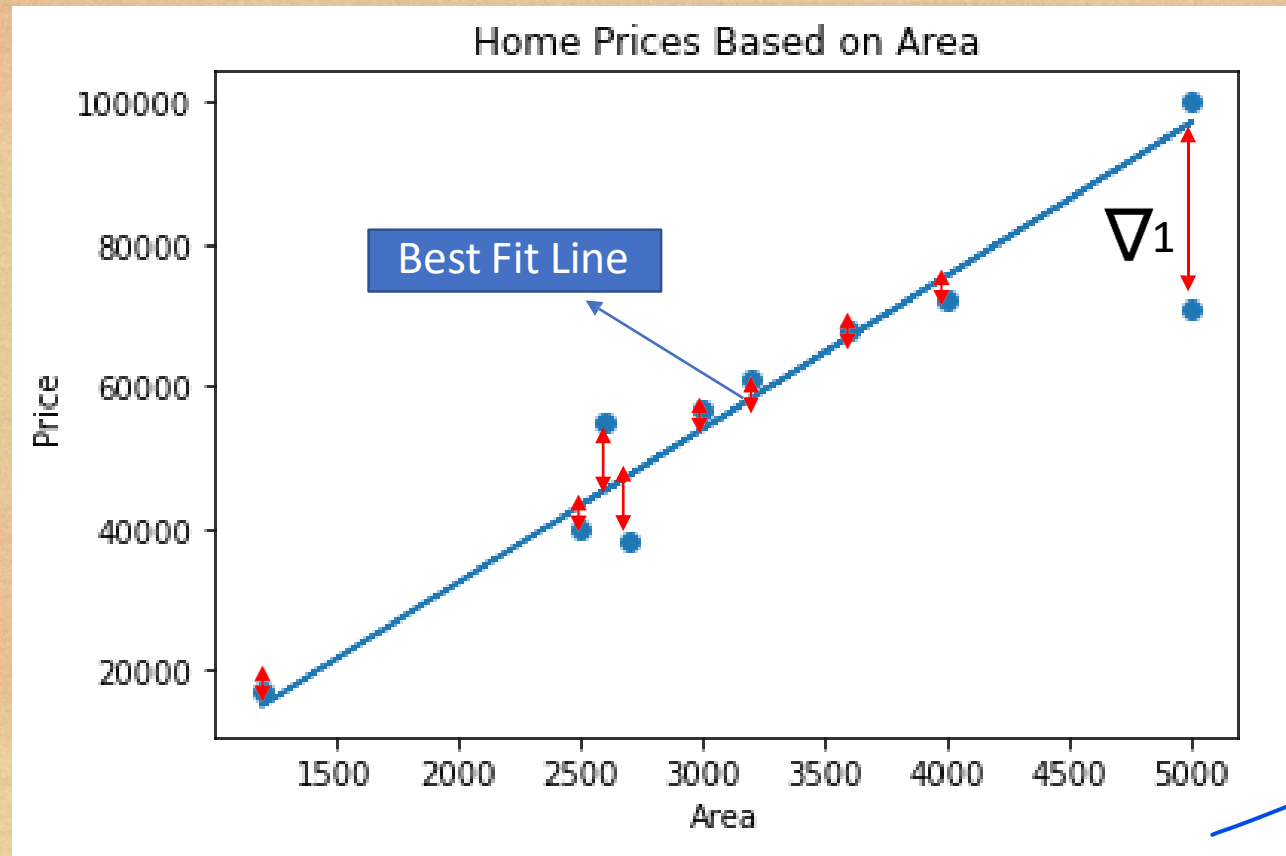


## Overview:

- Single Variable Linear Regression
  - Multiple Variable Linear Regression
  - Single vs Multiple
  - Cost Function
  - Gradient Decent
  - Accuracy
    - R2 Value
  - Implementing with Python
- 



## All about Single Linear Regression



$$y = mx + b ; \text{ or, } Y = 21.43 * X + 4980.13$$

Coefficient = 21.43  
Intercept = 4980.13

# All about Single Linear Regression

---

.



## All about **Single** Linear Regression

$$y = mx + b + \text{error}$$

$$y = b + mx + \text{error}$$

The diagram shows the equation  $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$  with the following labels and arrows:

- Dependent Variable** points to  $Y_i$ .
- Population Y intercept** points to  $\beta_0$ .
- Population Slope Coefficient** points to  $\beta_1$ .
- Independent Variable** points to  $X_i$ .
- Random Error term** points to  $\epsilon_i$ .

Below the equation, two curly braces are used for grouping:

- A brace under  $\beta_0 + \beta_1 X_i$  is labeled **Linear component**.
- A brace under  $\epsilon_i$  is labeled **Random Error**.



# Linear Regression with **Multiple** Variable

Multiple

The diagram illustrates the multiple linear regression equation: 
$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$$
 with the following annotations:

- Y**: response, dependent variable, observation, 'y-variable' (indicated by a red arrow from the red circle around Y).
- $x_1$** : predictor, 'x-variable', independent variable, explanatory variable (indicated by a green arrow from the green circle around  $x_1$ ).
- $\beta_2$** : coefficient (indicated by an orange arrow from the orange circle around  $\beta_2$ ).
- Linear Predictor**: The sum of the terms  $\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$  is grouped by a blue bracket and labeled "linear predictor".
- $\varepsilon$** : random error, "noise" (indicated by a purple arrow from the purple circle around  $\varepsilon$ ).



## Linear Regression with Multiple Variable

Single

$$y = b_0 + b_1 * x_1$$

Multiple

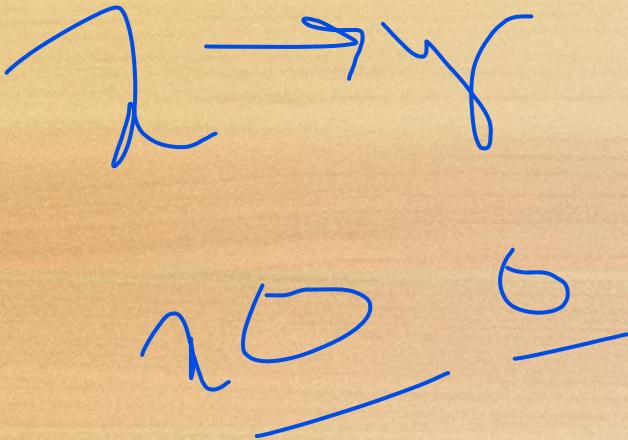
Dependent variable (DV)      Independent variables (IVs)

$$y = b_0 + \underline{b_1 * x_1} + \underline{b_2 * x_2} + \dots + \underline{b_n * x_n}$$

C



# R Squared Value


$$\text{R Squared Value} = \frac{\sum_{i=0}^{n-1} (\text{Predicted Value (Yp)} - \text{Mean Value } (\bar{Y}))^2}{\sum_{i=0}^{n-1} (\text{Actual Value (Y)} - \text{Mean Value } (\bar{Y}))^2}$$

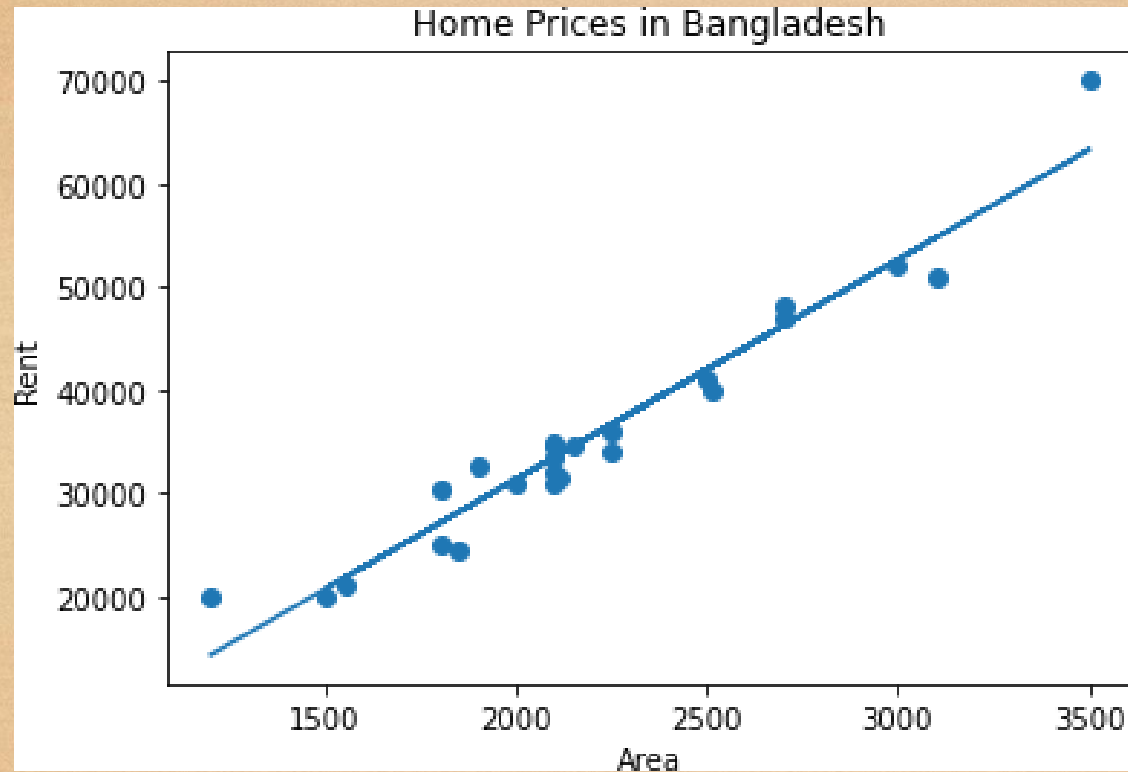
= Something \* 100

Accuracy = Something %

Abu Bakar Siddique Mahi







```
reg.score(xtest,ytest)
```

```
0.8296548051728814
```

`Y_pred = reg.predict(xtest)` **#Predicted Y**

**R2 Score:** `from sklearn.metrics import r2_score`

`Score = r2_score(ytest, Y_pred)`