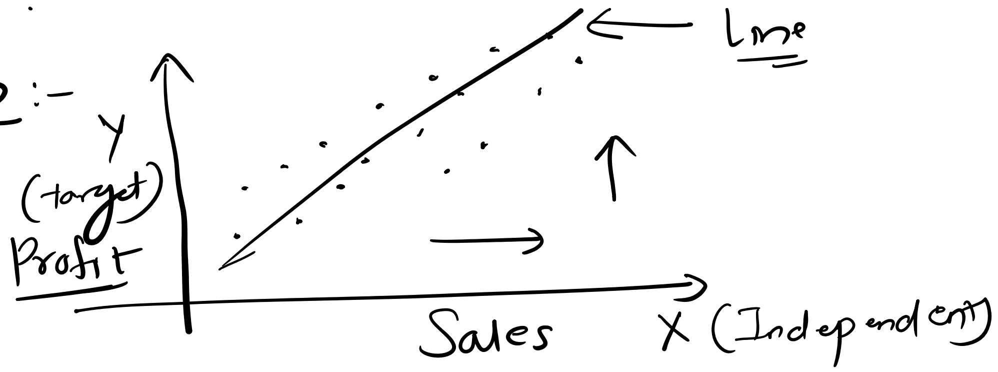
15 March 2023 19:41

Regression :-

Goal :- To predict a continuous target variables based on the available independent variable.

Linear Regression :-

Simple Straight line



Linear Regression :- To establish a relation between the target and independent variable using a straight line.

$$\boxed{y = b_0 + b_1 x}$$

y = target variable
x = Independent variable

$b_0 \rightarrow$ intercept
$b_1 \rightarrow$ Slope

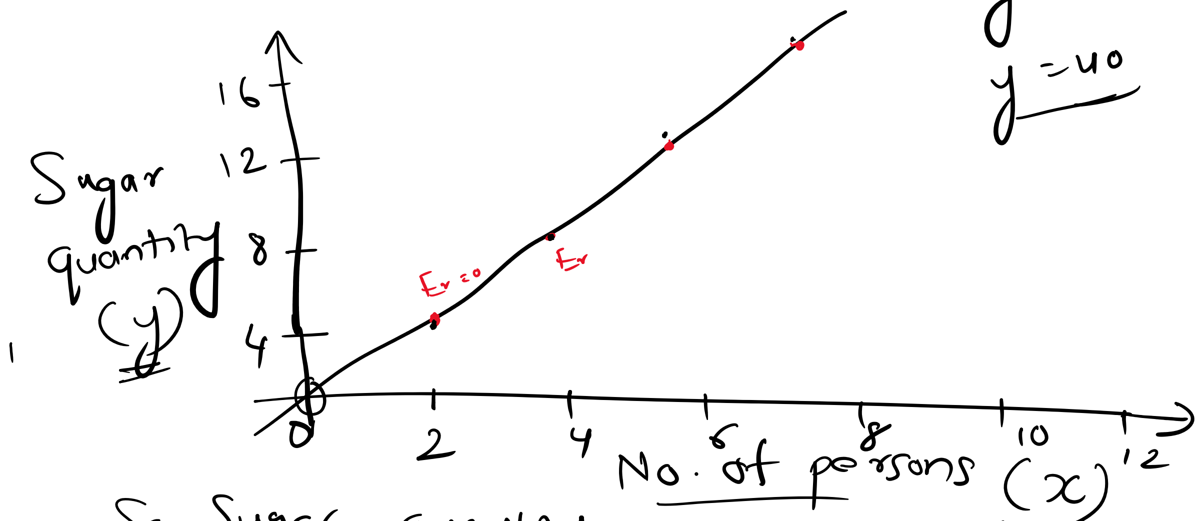
15 March 2023 19:51

eg:- Let consider an example

$x = \underline{\underline{20}}$

$y = 2 * 20$

$\underline{y = 40}$



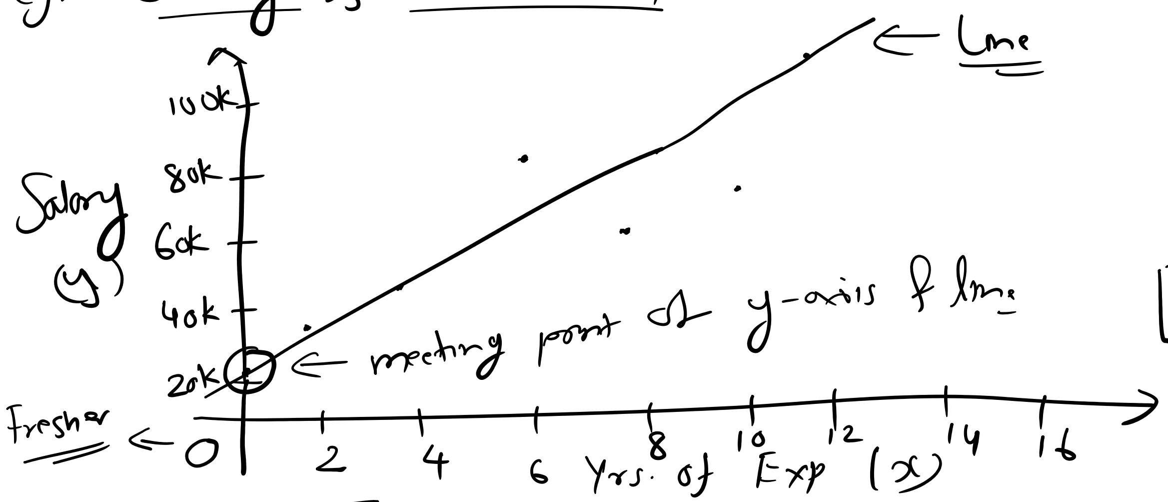
2 x times the sugar of No. of persons

$\boxed{y = 2x}$

~~Su~~ Sugar quantity = 2 * No. of persons.

15 March 2023 19:58

eg:- Salary & Yrs of Exp



Freshers Salary = 25K

$$\boxed{y = b_0 + b_1 x}$$

b_0 = intercept

b_1 = Slope

Salary = $\boxed{25K}$ + (b_1) yrs. of Exp

Slope figure out

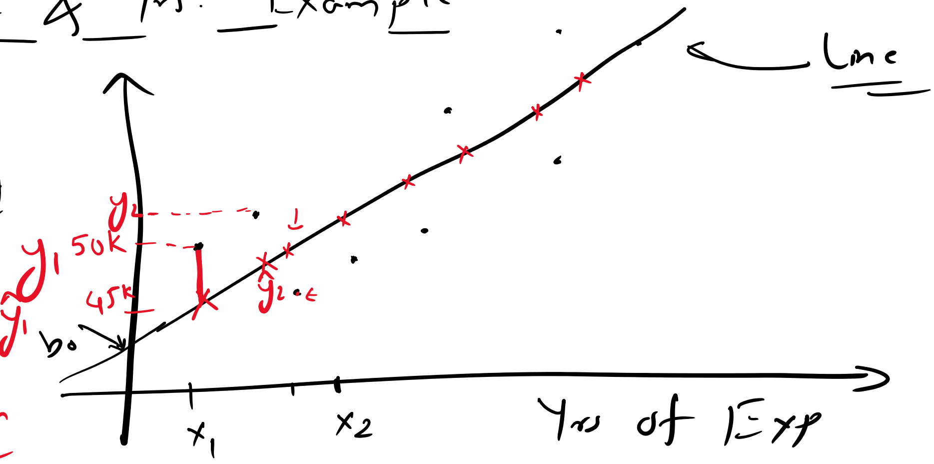
Intercept (b_0) :- The meeting point of Linear Regression Line & Y-axis is called 'intercept'.

Tea ex :- $y = 2x \Rightarrow y = \underbrace{2}_{b_1 \text{ slope}}x + \underbrace{0}_{\text{Intercept}}$

15 March 2023 20:09

Salary & Yrs. Example

Salary



50k - 45k = Error

Error = 5k

• ← represent actual salary values

× ← represent the predicted value.

y_1 = actual salary

$\hat{y}_1$ = predicted salary

$$\boxed{\text{Error} = y_1 - \hat{y}_1}$$

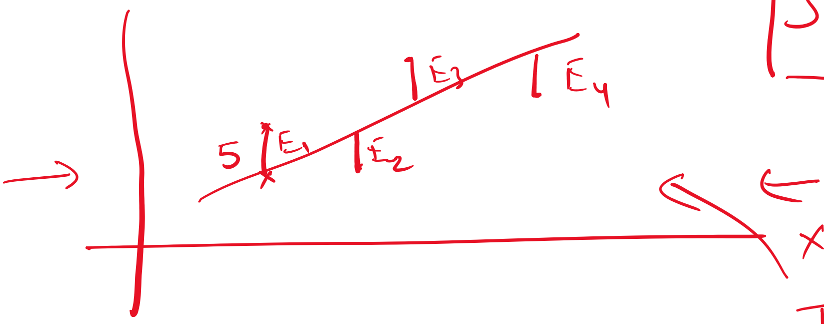
$$\boxed{\text{Error} = y_2 - \hat{y}_2}$$

$$\boxed{\text{Error} = y_3 - \hat{y}_3}$$

$$\boxed{\text{Error} = y_i - \hat{y}_i}$$

where $i = 1$ to n

$$\boxed{\text{Sum of Error} = \sum_{i=1}^{n} (y_i - \hat{y}_i)}$$



$$\text{Error} = \cancel{5} + \cancel{(-5)} + \cancel{(8)} + \cancel{(-8)}$$

$$\boxed{\text{Error} = 0}$$

$E_1 = 5$

$E_2 = -5$

$E_3 = 8$

$E_4 = -8$

15 March 2023 20:25

$$\boxed{\text{Squared Error} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

$E_1 = 5$ $E_3 = 8$

$E_2 = -5$ $E_4 = -8$

$$= (5)^2 + (-5)^2 + (8)^2 + (-8)^2$$

$$= 25 + 25 + 64 + 64$$

$$\boxed{\text{Squared Error} = 178}$$

$$\boxed{\text{Mean Squared Error} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

15 March 2023 20:30

Sq. Error = 178

$$MSE = \frac{178}{4}$$

$$\boxed{MSE = 44.23}$$

* Mean Squared Error is called as the cost function for Linear Regression problems

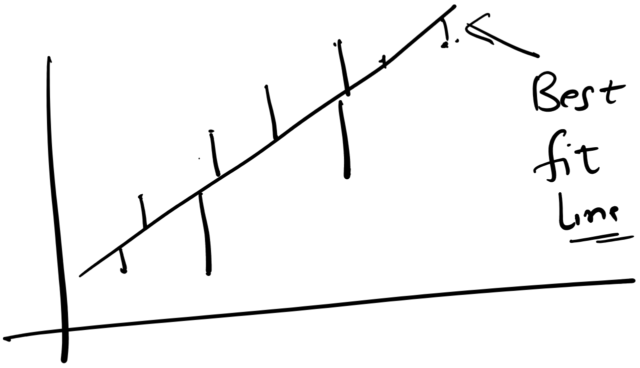
15 March 2023 20:39

Cost function :- ① It is a function used to measure the performance of a model on any data.

② It measures the error between actual & predicted values & outputs that in the form of a number.

For LR, the cost function is Mean Squared Error

For LR, the MSE should be minimum.



Cost function - performance

For LR, cost function = MSE

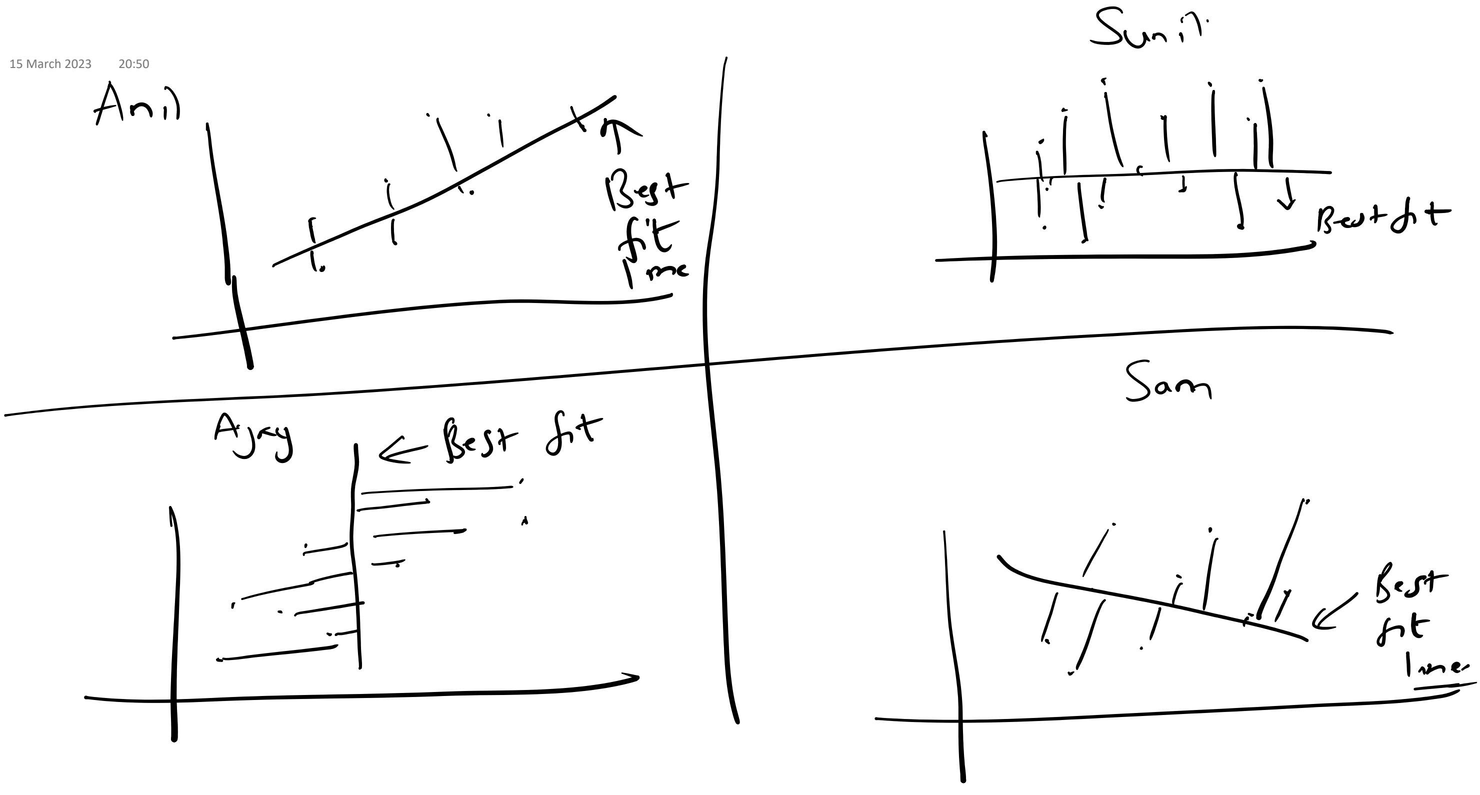
MSE = Error of LR

The entire aim of LR model is to find a Best fit line. by <u>minimizing MSE</u>

Q How to find the best fit line?

→ By minimizing MSE.

Q. Why minimize MSE?


Anil
Best fit line
Sunil
Best fit
Ajay
Best fit
Sam
Best fit line

15 March 2023 20:54

Types of Linear Regression

1. Simple Regression.
   - (a) 1 Independent variable
   - (b) 1 target variable
2. Multiple Regression
   - (a) Many Independent variables
   - (b) 1 target variable.

15 March 2023 20:56

3. Polynomial Linear Regresson
4. Lasso Regression.
5. Rigde Regressim

When we face a problem of overfilting & underfitting

* <u>Note</u> :- For applying Linear Regression, there should be a linear Relationship between the variables.

If no Linear Relationship, then we cannot apply Linear Regression Models.