Simple Linear Regression (Supervised Machines learning Algorithm)

* Regression :- To establish a relationship b/w the two Variables/more than two variables.
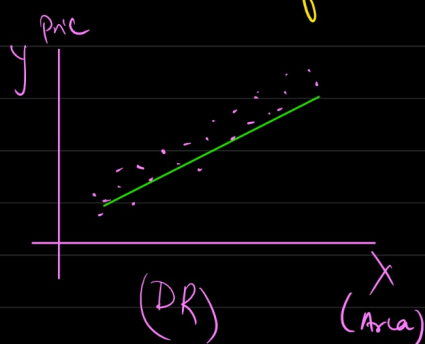
* Linear → It establish a linear relationship.

Supervised ML

Regression   classification

* SLR attempts to determine the strength and characteristics of the relationship between one dependent Variable (y) and another variable independent Variable (X)
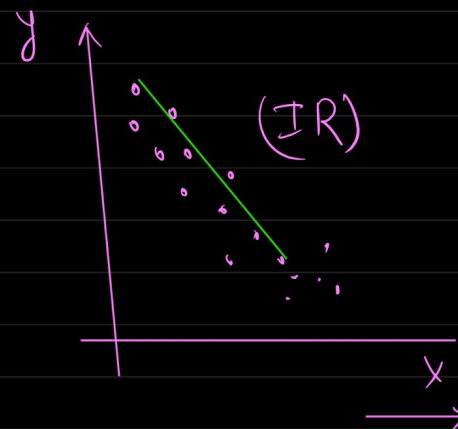
Predict Price of house based on Area of house.

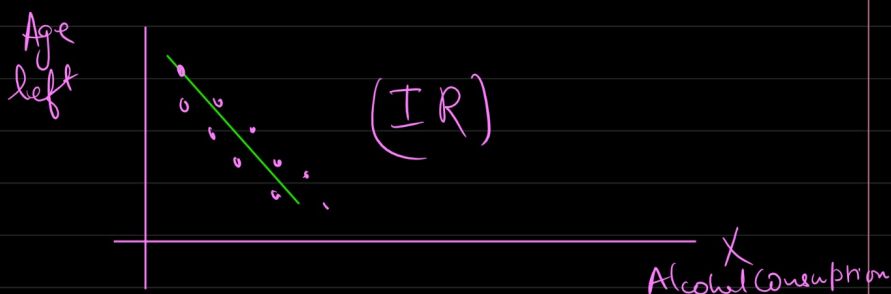| Area of house	Y
1100	50
1200	60
1300	70
—	—
—	—

y Price

(DR)   (Area)

eg   Selling Price of Car wrt age of car.

X age of car	Y selling Price
10	3.1
9	4.1
7	5.3
—	—

y
(IR)
X

eg Age left Vs Alcohol consumption

Age left
(IR)
X
Alcohol consumption

Simple   Linear   Regression
↓   ↓   ↓
Only one X   Captures the   Understand
and y   linear relationship   relationship

* if multiple IV (multiple X's) → multiple linear regression.

* Example

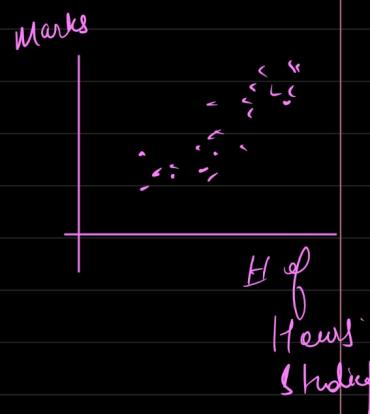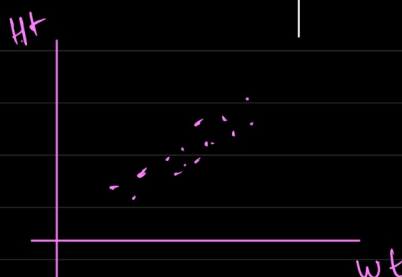| No of Hours studied	Marks obtained
8	72
9	85.5
—	—
—	—
—	—

Marks

H of
Hours
studied

example

Weight	Height
65	168
70	172
—	—
—	—

Ht

wt

* To predict price of a house based on number of room.

y = price of house
X → Number of rooms.

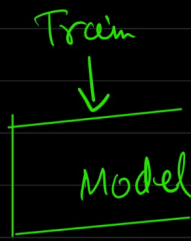| # of rooms (x)	Price of house (y)
1	50
1.5	60
2	65
3	75
4	? → 85L

⇒ ML — To understand the mathematical relationship.

This was 4 dp's that's why you understood the trend. If many dp's you

$\hat{y} = mx + c$

$\hat{y} \rightarrow$ estimator

(You have estimated the price of house base on Area of house, m, c.)

$(slope)\ \hat{m} = \dfrac{y_2 - y_1}{x_2 - x_1}$

c = intercept (where it cuts Y axis; $x = 0$)

Y (price of house)

120
110
100
90
80
70
60
50

$y_{pred} = 52L$

1100 1200 1300 1400 1500 1600 1700 1800 1900 2000

x 1350

Area of house (X) (sq ft)

$\hat{y} = mx + c$

$\parallel$

$\hat{y} = \beta_0 + \beta_1 x$

$\begin{pmatrix} \beta_1 = m \\ \beta_0 = c \end{pmatrix} \rightarrow$ Coefficients

$\boxed{h_\theta(x) = \theta_0 + \theta_1 x}$ $(\theta_1, \theta_1 \rightarrow$ Coefficient$)$

$x \rightarrow$ is dp

$(m, c), (\beta_0, \beta_1), (\theta_0, \theta_1)$

Coefficients.

many book, you will find this format → because we consider a S·L·R line to be hypothesis that it passes from these points