

Linear Regression

- ❖ Linear regression is a way to identify a relationship between two or more variables and use these relationships to predict values of one variable for given value(s) of other variable(s).
- ❖ Linear regression assume the relationship between variables can be modelled through linear equation or an equation of line

$$\text{Dependent/Regressed variable} \leftarrow \mathbf{y = w_0 + w_1 X} \rightarrow \text{Independent/Regressor variable}$$

↑
Slope

↓
Intercept

Multiple Regression

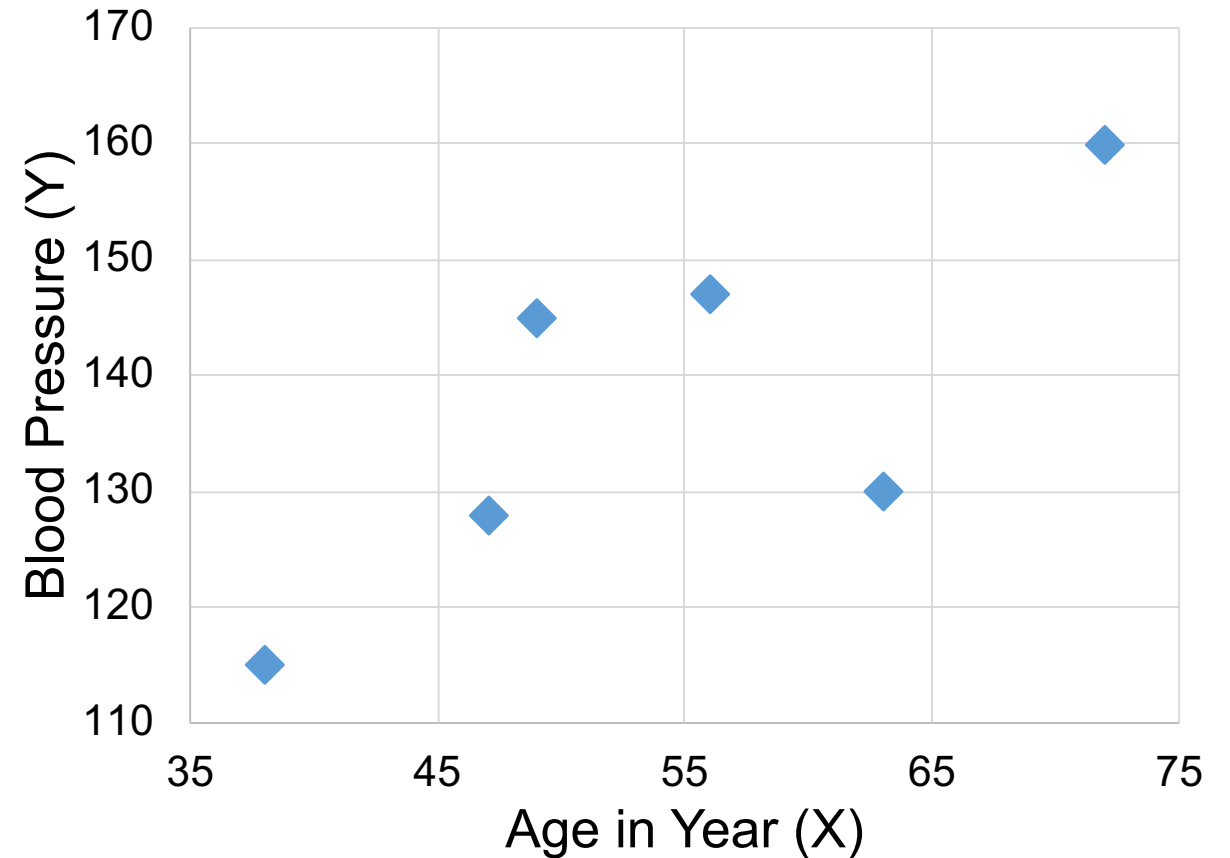
- ❖ Last slide showed the linear regression model with one independent and one dependent variable.
- ❖ In Real world a data point has various important attributes and they need to be catered to while developing a regression model. (Many independent variables and one dependent variable)

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_3 \dots \dots \dots W_nX_n$$

Regression –Problem Formulation

Let you have given with a data:

Age in Years (X)	Blood Pressure (Y)
56	147
49	145
72	160
38	115
63	130
47	128



Linear Regression

❖ For given example the Linear Regression is modeled as:

$$\text{BloodPressure}(y) = w_0 + w_1 \text{AgeinYear}(X)$$

OR

$$y = w_0 + w_1 X - \text{Equation of line}$$

with w_0 is intercept on Y -axis and w_1 is slope of line

Blood Pressure

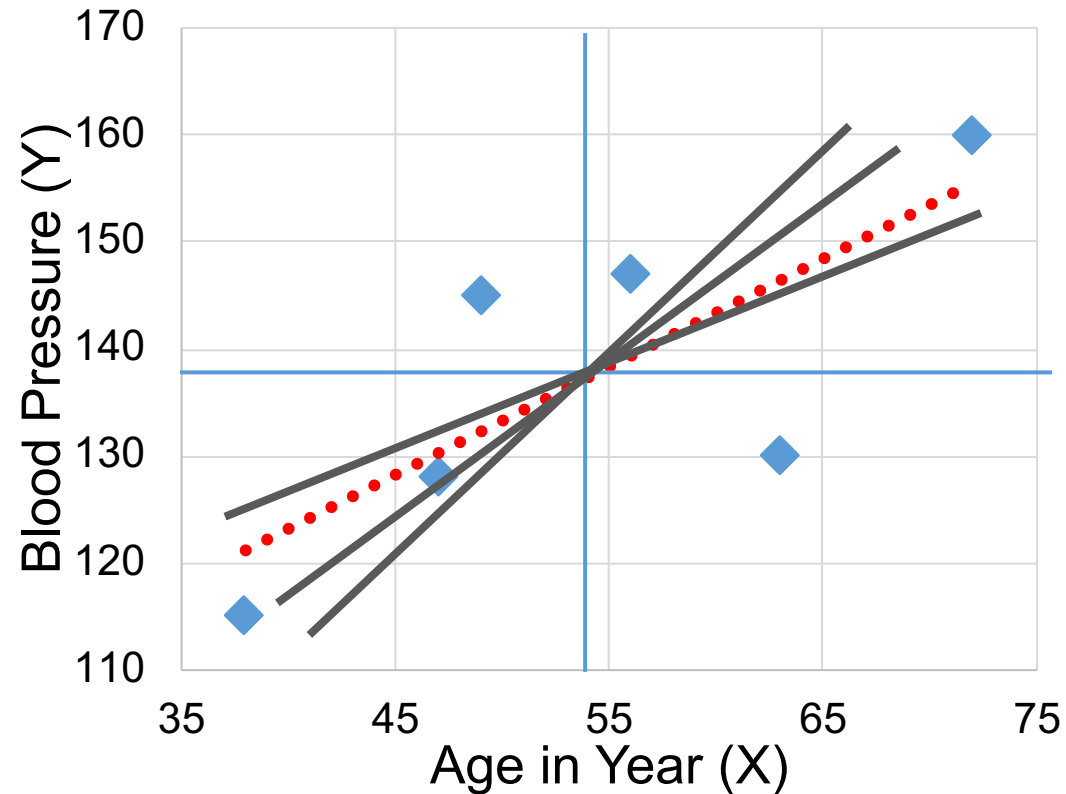
- Dependent Variable

Age in Year

- Independent Variable

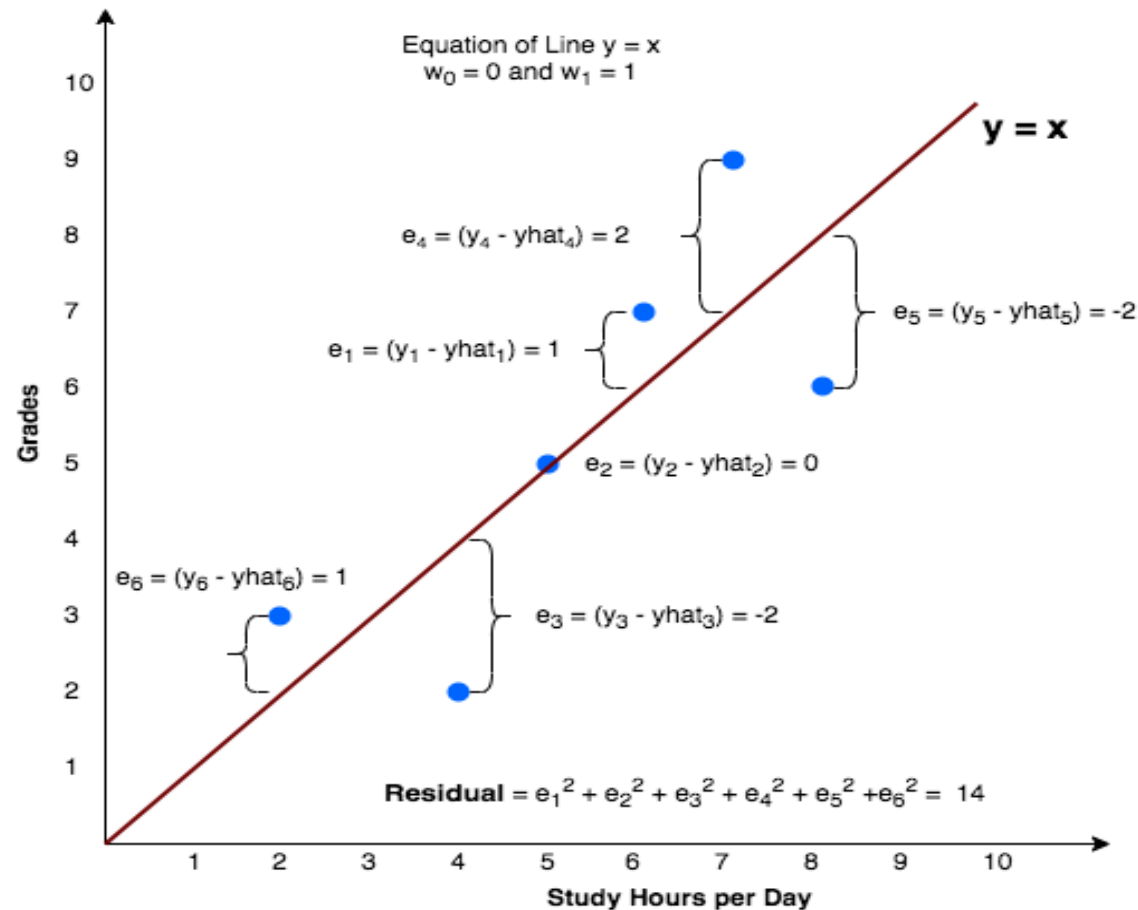
Linear Regression- Best Fit Line

- ❖ Regression uses line to show the trend of distribution.
- ❖ There can be many lines that try to fit the data points in scatter diagram
- ❖ The aim is to find **Best fit Line**



What is Best Fit Line

- ❖ Best fit line tries to explain the variance in given data. (minimize the total residual/error)



What is Best Fit Line

- ❖ Best fit line tries to explain the variance in given data. (minimize the total residual/error)

