* y – response variable / target variable / dependent variable / predicted variable
* X - X1, X2, X3, X4, X5 – features of the data / independent variables / dimensions of the data / predictor variables
* Numerical data
  + Continuous distribution (floating point)
  + Discrete distribution (integral numbers)
    - Factor type – only x number of levels eg. Passenger class (having levels 1/2/3)
* Character data
  + Factor type / categorical type eg. gender (only 2 levels m / f)
  + Character type eg. nam
* When my response variable (y) is continuous distribution – regression problem eg. prediction of sales of a company, prediction of sales by each sales rep in my team, prediction of ds scores of each learner in my batch
* Classification problems are when the response variable is categorical in nature
  + Binary classification problem – response variable is binary categorical

eg.

* + - prediction if the tumour is malign or benign
    - Prediction if the person will default on loan or not
    - Prediction if my existing customer will churn or not
    - Prediction if the customers who have put my product in the wishlist will purchase the product or not
  + Multi-class classification problem – when response variable is having multiple categories eg.
    - What rating will the customer give to my app (1/2/3/4/5)
    - Product rating / food rating / movie rating
    - Will my customers prefer to purchase Diet Coke / Regular Coke / Coke sweetened with Jaggery
    - Tourists will choose which destination – Goa / Pune /Shimla
    - What dish the customer will order in my restaurant
* Big Data – volume, velocity(real time data – on internet) and variety of data (excel sheet / html file / json files /pdf files / doc files / image (pixels) / video /audio data / text data
* Structured data
* Semi-structured data – there is a structure to this data, but that structure is not in rows and cols format like html data, json, xml files – with simple coding that can be converted to a structured format
* Unstructured data – image, video, audio, text