Data Analytics

**Lecture 6**

**Regression**

* Lm functions for will hand-over a model which will be considered as a predictive functions.
* Create a formula object and it will start with the name of the response functions.

**Response 🡨 Y = b0 x + b~ --🡪 predictor**

**Response ~ Predictor1 + Predictor 2**

Note - Signs can be changed based on the model . It can be replaced by \* . this it totally depends on the model which is implemented and the variables which are considered for regression.,

* To get the best predictor – we minimize the residue i.e. the value which we get by taking difference of ideal data and data we get after.
* Adding more and mode variables can give you worst model. Try to figure out the relevant variables to get the best prediction.

Y = b1x1 + b2x2 + b3x3 + b0

* Predictor can also be categorical.

**Adjusted R2** – Basically it’s a model where the residue has been squared so that the their will be optimum number of predictors which can minimize the residue.

Lecture 7

**Probability**

* If you do repeat the same process many times, you will get different results each time. For example flipping a coin several times 100 – if in first 100 times you can get 50 tails and 50 heads on the other hand you can get different results on second 100 attempts.
* Rule of probability – Unusual event have probability of less than 5%.