**Lecture-2**

1. Before starting data analytics you have to do a lot of data processing that includes data cleaning, visualization and drawing inferences and this overall process is highly iterative till the time you don’t get satisfactory results. In some sense exploratory data analysis without using a statistical model or using prior hypothesis.
2. This essentially does visual analysis of data to formulate a hypothesis that could be tested on newer datasets.
3. First try to visualize data might be a case that some datasets are statistically similar but there can be significant inferences which can be made on basis of visualization.
4. Try to ask as many questions as you can from your data and try to visualize in as many ways it is possible to visualize in order to get more inferences about the data.
5. Repeatability and workability is important that means the ways adopted should be iterated multiple times to validate efficacy of your models.
6. Data Types: Nominal (categorical) are equal or not equal to, Ordinal (O) obey a less than or smaller than relationship, Quantitative (q) can do arithmetic.
7. First row in dataset is semantic which tells what each column means. Each row holds a data entry and each column corresponds to an attribute.
8. Nominal/Ordinal = Dimensions (Describe data and independent variables), Quantitative = Measures (Numbers to be analysed)