**Assignment 2**

1. Discuss all the factors which are necessary for the quality of data.

2. Explain nominal, ordinal and numerical attributes by giving one example for each.

3. Discuss the concept of mean, median, node, trimmed mean, weighted mean , mid range, quantile, quartile, IQR for the given data distribution and suppose that the values for a given set of data are grouped into the following intervals. The intervals and corresponding frequencies are as given below**.**

**age frequency**

1-10 100

10-20 400

20-30 300

* 1. 2500

40-50 800

50-60 150

Compute an approximate median value for the data.

4. Discuss all the methods to fill in missing values and smoothing noisy data.

5. Explain all the issues arise during data integration?