**Experiment 3**

**Aim: To perform various data exploration operations and data cleaning using python**

References:

1)<https://www.analyticsvidhya.com/wp-content/uploads/2016/08/Data-Exploration-in-Python.pdf>

2) <https://www.kaggle.com/pmarcelino/comprehensive-data-exploration-with-python>

3) https://www.bu.edu/tech/files/2017/09/Python-for-Data-Analysis.pdf

Steps

1) Load the libraries

2) Download the data set from kaggle/ other sources

3) Read the file –select appropriate file read function according to data type of file( refer link 1)

4) Display attributes in the data set.

5) Describe the attributes name, count no of values, and find min, max, data type, range, quartile, percentile, box plot and outliers. (link 1 and link2-descibe function)

6) Give visualization of statistical description of data – in form of histogram, scatter plot, pie chart

7) Give correlation matrix

8) Identify missing values and outlier and fill them with average. (link 3)

**Post Lab: Assignment2**