**Numpy, Pandas and Visualisation**

Section A (2 marks each question)

1. Create a 7x7 array with random values using numpy.

2. Given an array with strings -

str\_arr = np.array(['Python', 'Machine learning', 'Sql', 'Excel']).

Use inbuilt numpy functions to print the array in upper and lower case.

3. Create a Dataframe using the following dictionary:

dicts = ({"Subject":["Machine Learning","Python"],

"Mid Exam":[150,80],

"End Exam":[200,70]})

4. Write code to reshape the array: np.array([2,3,6,3.4,6,3])

into a 2X3 array. After reshaping please get the square of all the elements.

Section B (3 marks each question)

5. Read the books.csv and ratings.csv dataset. Merge both the datsets using pandas on the BookId column.

6. Find out the number of authors in the dataset.

7. Find the number of books written by each author.

8. Plot the results of the previous question as a bar chart.

9. Plot the distribution of the average ratings column.

Section C (7 marks)

10. Use the merged datasetto plot the following: (3 + 4 marks)

a) Generate boxplots for all the numerical variables in the dataset.

b) Generate pairplots for the same dataset and point out the variables having a high correlation.