**Python Assignments**

**Assignment 1:**

1. Using Pandas, load the dataset from <https://www.kaggle.com/uciml/iris/data> to a variable name iris

2. Create a list named headers with all the column header names in the given order.

3. Using the slice operation on headers, extract the column names with index 1 to 4 onto a list called features.

4. Display the first five records of iris

5. Make a scatterplot of the Iris features.

6. What is the range of ‘SepalLengthCm’ in the dataset? What is the second largest value of ‘SepalLengthCm’ in the dataset?

7. Find the mean of all the values in SepalWidthCm using numpy

8. Identify ‘SepalLengthCm’ values less than 5. Create a new column named ‘Length’ , categorise each entry as ‘Small’ or ‘Large’, if less than 5.

9. Group dataFrame by the "Species" column. Make a histogram of the same.

10 .Find the deviation of length for ‘SepalLengthCm’ from the average

11 . Find correlation between columns and display columns with more than 70% percent correlation (either positive or negative).

12. Impute missing values if present using mean of the dataset.

13. Save the current dataFrame out to a new csv file.

**Assignment 2:**

**User movie ratings dataset**

1. Read and merge the dataset into single dataframe

2. Provide the names of 10 movies rated highest.

3. Rank the movie names by their highest average rating scores.

4. Plot movie scores across each genre.

5. Find the group that provides the highest average movie ratings when split into genre groups.

6. Provide a table with the average rating of a movie by each genre group along with the movie title.

**Assignment 3:**

**Numpy**

1. Create a function which creates an n×n array with (i,j)-entry equal to i+j.
2. Create a numpy array which contains odd numbers below 20. Arrange it to a 2x5 matrix. Compute the log of each element.
3. Create a function which creates an n×n random array. Subtract the average of each row of the matrix
4. Create a function which creates an n×n random array. Write a program to find the nearest value from a given value in the array.
5. Write a function to check if two random arrays are equal or not.
6. Create a function to get the n largest values of an array.