Assignment No:1

**DATA PREPROCESSING**

Filtering Data (library **dplyr**)

Read IRIS dataset and perform the following using filtering function

1. Select all observations having following conditions:
   1. Sepal width<5
   2. Petal length=4
   3. Petal width>3
   4. Sepal width>5 and species =”setosa”
   5. Petal width<=3 or petal width>=4
   6. Sepal width in between 4 and 5
   7. Sepal length near to 3

Arrange Data (library **dplyr**)

Read IRIS dataset and perform the following using arrange function

1. Arrange observations with following conditions:
   1. Ascending order of Sepal length
   2. Descending order of petal width
   3. Ascending order of sepal length and sepal width

Sorting Data (library **dplyr**)

1. Create a data vector and perform various order of sorting
2. Create a data vector having NA and check the usage of na.last

Ranking Data (library **dplyr**)

1. Create a data vector and perform ranking
2. Try various types of ties breaking methods and find the difference

Grouping Data (library **dplyr**)

1. Read iris dataset and group the data with respect to species
2. Find the max, min, mean of each feature in grouped data

Mutate Data (library **dplyr**)

1. Read iris dataset and mutate a new column with log of sepal length
2. Try the usage of **keep** parameter in different ways

Rename Data (library **dplyr**)

1. Use rename function in IRIS dataset
2. Try the usage of **rename\_with** with various functions (toupper, tolower)

SliceData (library **dplyr**)

1. Use slice function in IRIS dataset in different ways