Department of Statistics, School of Mathematical Sciences
Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon
M.Sc. (Statistics) I Practical Sheet: MST 506
Topic 2 Use of dplyr package

---

1) Consider a following data.

Manager	Date	Country	Gender	Age	q1	q2	q3	q4	q5
1	10/24/14	US	M	32	5	4	5	5	5
2	10/28/14	US	F	45	3	5	2	5	5
3	10/01/14	UK	F	25	3	5	5	5	2
4	10/12/14	UK	M	39	3	3	4		
5	05/01/14	UK	F	99	2	2	1	2	1

   a) Create a data frame of the above data by named `employee_data`.
   b) Create two summary variables, `total_score` and `mean_score`.
   c) Recode `M` and `F` to `Male` and `Female`.
   d) Rename `Manager` and `Gender` variable as `ID` and `Sex` respectively.
   e) Sort the data by `Sex` variable and then total score within `Sex`.
   f) Create a new data frame containing the `mean_score` with respect to `ID`.
   g) Create a new variable containing `Male` with `total_score` above 10.

2) Use the `mtcars` data set from `MASS` library. Write a R code to solve the following.
   a) Rank the cars that have more than 100 horsepower by their gas mileage in descending order.
   b) Create a data set by selecting `mpg`, `hp` and `qsec` columns only.
   c) Create a new variable called `DPC` which is `disp/cyl`.
   d) Summarize the data by calculating mean and s.d. of `mpg`.
   e) Group data by number of cylinders and calculate average `mpg` and `hp` within each group.
   f) Modify the data by filtering the data where `mpg` is greater than 20 then arrange in descending order with respect to `hp` and then select `mpg`, `hp` and `qsec` variables.
   g) Create a new variable `mpg_category` that categorizes cars as "High" and "Low" when `mpg > 25` and `mpg < 25` respectively.
   h) Create a new variable that represents a scaled version of `mpg`.
   i) Rename the variables `mpg` as MilesperGallon, `hp` as Horsepower, `qsec` as QuarterMileTime.
   j) Code `0` and `1` as `Automatic` and `Manual` respectively to values of `am` variable.