**CS 40003 : Data Analytics** **12.08.2016**

Problems

Two >CSV files are given. These files contain a tables of data, which are self-explanatory.

With the data, you are to do the following.

- Read the .csv files.
  [Hint: You should read the file into RStudio using R-command “read.csv (<filename>).
  e.g.: table1 = read.csv("D:/DataScience-R/Data1.csv", header = T, sep = ",");

- Use na.rm = T argument for handling missing values in Data2.csv

- Read any entry in any table
- [e.g., table1$Sub3[7] gives marks of Studentid=7 for Sub3.

- Calculate subject wise mean, variance and standard deviation for each subject in each file.

- Calculate the student wise mean, variance and standard deviation for each student in each file

- What is the median of mean values subject-wise?

- Data in which table is more consistent? How, you conclude that?

Submit all R-programs to Moodle course management system

**Last date of submission: 21.08.2016 24:00 hrs.**