COMM 645, LAB 4: DUE OCTOBER 03, 2012

Please e-mail your lab to **comm645@ognyanova.net** before 2pm on Wednesday.

This lab asks you to do basic data reading and processing in R. You will find the data for it in a file called "Lab4Data.csv". The file contains 50 cases including names of people, the date when they took a survey, their age, income, and total years of education.

One of the tasks you will be asked to perform (converting strings into dates) is not listed on the handout for this lab. You will have to use the R help files or search online to find the appropriate function for this.

You do not need to write a report for this lab – just turn in your R script.

Your assignment consists of the following:

1. Read Lab4Data.csv into a data frame in R.
2. Compute the mean and standard deviation of the age and income columns.
3. Compute the correlation between age and years of education.
4. Find the names of all people who have an income level of 7. Compute their average age.
5. Add a new numeric column to the data frame containing only the number 1 for all rows/cases.
6. Add a new column to the data frame containing case IDs: consecutive numbers from 1 to 50.
7. Add a new column to the data. This column should contain logical values: TRUE for people under the age of 40 and FALSE for people age 40 and over.
8. The "Date" column of the data is stored as text (character vector). Convert it to a Date format. Add a new column to the data containing logical values: TRUE for everyone who took the survey before May 1st 2008, and FALSE for everyone who took it on or after that date.
   (Note that you can use comparisons – greater than (>), less than (<), equal (==), etc. – for dates much like you do for numbers).
9. Create a new data frame that contains only the name and income columns from the original data (it should have 2 columns and 50 cases/rows).
10. Save this data frame to a .csv file.