Statistical Computing (140.776) Homework 1

Due Thursday, Sep. 15

1. Download sortdata.txt used in Lecture 2. Define `z<-log(x[,2]-x[,1])`. Replace all NaN in *z* by 0. Then compute (1) mean of *z*; (2) $z_{(1)}+z_{(3)}+z_{(5)}+\ldots+z_{(999)}$.

2. Solve the linear equations below using R:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 8 \\ x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = -1 \\ x_1 + 4x_2 + 9x_3 + 16x_4 + 25x_5 = 2 \\ x_1 + 8x_2 + 27x_3 + 64x_4 + 125x_5 = 9 \\ x_1 + 16x_2 + 81x_3 + 256x_4 + 625x_5 = 3 \end{cases}$$

3. Download student.rda used in Lecture 3.

(1) Convert the data into a data frame that contains name, major, age and first two grades of each student.

(2) Save the data frame into a tab-delimited text file.

(3) Reload the text file you saved. Using the reloaded data, compute the max grade for each student. Then compute the average of the max grade of students with mathematics major.

4. Load the "CO2" dataset which is available in R and write code to reproduce the figure shown below. The figure should be saved automatically by R into a single PDF file. Print your code and figure and submit the printed results.

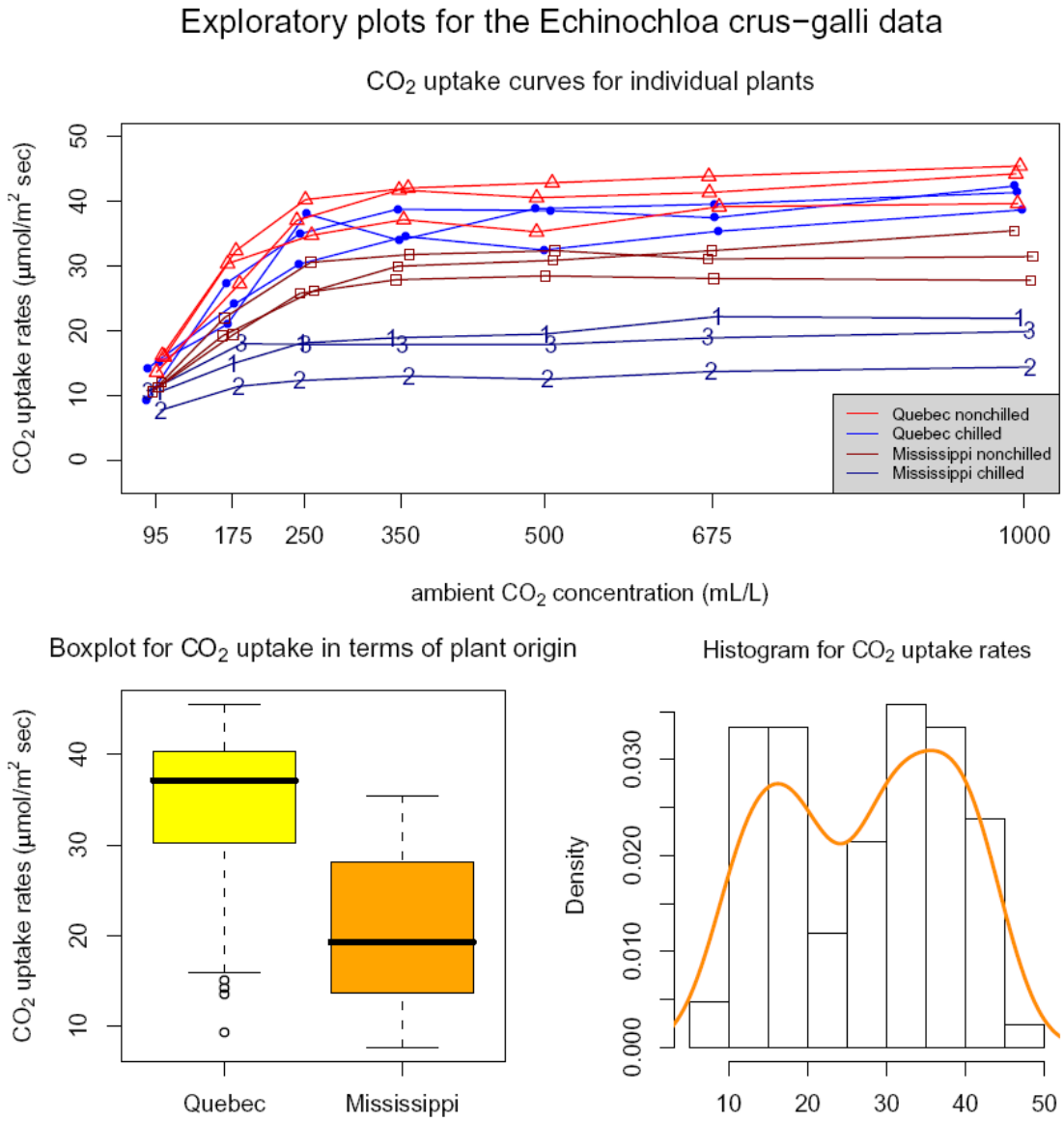
Hint:

(a) You can learn about the data by typing "?CO2".

(b) You can load data by typing "data(CO2)"

(c) The following functions might be useful: par(), layout(), plot(), lines(), points(), legend(), axis(), mtext(), expression(), density(). Try to read their help documents.

(d) Consider to use jitter() function to add a small amount of noise to the data, thus have fewer points overlap.



Just proving that we can write things here too!