IOM 530: Applied Modern Statistical Learning Methods
Assignment 1 (Due Sep 5, 2013)

Questions:

Exercise 8 in Chapter 2 on page 54 of *An Introduction to Statistical Learning in R*.

Guidelines for assignment submission:

1. This exercise has multiple parts. Please, answer each part separately and in a brief way. Be direct to the point!
2. Type each question before you answer it, and provide a clear separation between each part.
3. All relevant computer output should be provided unless noted otherwise.
4. Print your homework, and submit it at the beginning of the class. Make sure that it is stapled, and your name is typed on it.
5. Attach your R code as an Appendix. Make sure to provide comments on what your code is doing. Keep it clean and clear!

Hints:

1. In part (a), when you read the data into R, make sure to check if the data has a header or not.
2. In part (b), you don't need to use the *fix()* function to view the loaded dataset. Instead, and since we are using rstudio IDE, we can view the data by clicking on the data name (college) in the "Workspace" window in rstudio. I am saying this, because sometimes the *fix()* function might crash rstudio especially if you are using Macs. Another option would be to use the command *View(college)*.
3. Parts (a) and (b) are for data manipulation (i.e. there is no need to include any output in the report for submission). You will be mainly answering questions from part (c).
4. For part (c-iii), make sure to annotate the plot (title, x-axis, and y-axis).
5. If you want to learn more about a certain R function, you can use the command "?". For example, if you want to learn more about the plot() function, type the command "?plot", and a help document will pop up.
6. In part (c-v), when you use the command par(mforw(2,2)), the plotting screen should split into 2x2 = 4 parts. To go back to the original setting, run the command par(mforw(1,1)).