Hints for Getting Started with R in Predict 401

If you are wondering where to focus your initial study efforts with R, here are some suggestions.

1. Install R. R can be obtained at *http://cran.r-project.org/*. Instructions are given in the text by Verzani as well as the text by Lander. The instructions in Lander cover Windows, Mac and Linux systems with pictorial illustrations.

   RStudio is recommended for new users of R. It is not required. Everything students need to do in this course can be accomplished using the standard R console with a plain text editor. RStudio is an integrated development environment for R. Installer packages for RStudio are located at *www.**rstudio**.com/*. RStudio is easy to install and use.

2. Read and try out the exercises in the "Quick Start Guide". This guide demonstrates how R can be used as a calculator. It also illustrates the use of a variety of functions that are needed for the midterm exams. For example: summary(), plot(), pairs(), c(), seq(), par(), boxplot(), and hist(). This is not an exclusive list. You should try out sort() and stem(). Learning how to use the functions, and perform the operations shown in the guide will be a good starting point.

3. Once the basic functions mentioned above are understood, try the basic R programming lessons in SWIRL. These lessons are quick and give immediate feedback. They are good for practice and bring in additional aspects of the language.

4. There are three additional topics that need to be studied.

   a. The first is construction of a histogram with a defined starting point and class width. (This can be accomplished using options supplied with the hist() function.)
   b. The second is construction of a frequency table in conjunction with calculation of a grouped mean and variance. (Doing this in R requires a little programming.)
   c. The third is use of a normal quantile plot with a line. (The R functions are qqnorm() and qqline().)

5. After trying out the "Quick Start Guide" and SWIRL, the next step would be to work on the "Lessons in R". These bring out other aspects of the language.

6. During the course it is important to master the following R basics, at a minimum, to be prepared for more advanced courses in the program: (1) how to read a data file into a data frame, (2) how to install R libraries, (3) how to write a for loop, (4) how to write a R function, (5) how to make a basic graphical object, and (6) how to fit an OLS regression model. Beyond these basics, it would be helpful to learn how to subset a dataframe and use aggregate(), apply() and lapply().

7. The data analysis assignments give coding examples and suggestions for their completion. Todd Peterson is available to answer questions.