Final Project

The final project is wide open. Working in groups you will decide on a programming problem, program it and then present your work in a class presentation. There will be five groups (one group of four and four groups of three) and presentations will be about 30 minutes long during the last two weeks of class.

As always, please turn-in a hard copy of your R code along with a copy of your presentation slides. Also submit via e-mail (njc23@pitt.edu) a copy of your R code.

Due Dates

March 29, 2011	E-mail me your group members and a brief description of your project
April 5, 2011	Submit about a one page summary of your prosed project
April 19, 2011	Group Presentations
April 26, 2011	Group Presentations

Projects

Projects should have a clear purpose and objective that is meaningful. Possible projects include writing code that automates a repetitive task or introducing new functions that make programming and analysis with R more convenient. This is not a data analysis project, but you are welcome to program a statistical method and demonstrated it using a dataset.

Below are some ideas of possible projects, you are welcome to take one of these ideas and expand on it for your project. Or if you prefer you can do something related to your research or anything else.

1. *Education* Write a few functions, possibly using animation, (maybe even using the `tcltk` package) that will allow undergraduate students to explore various concepts of statistics. In addition to these example functions, include code that will keep tract of: (1) how often a student calls a function and (2) how much time a student spends between function calls. The idea is to monitor how many examples the student works through and how much time they spend thinking about the results before the go on to the next example. When the student closes R a report should be generated and saved to a file. The student should not be aware that their activity is being recorded. The function `.Last()` will be very useful.

2. *Quarterly Report* Suppose you are responsible for generating quarterly reports every three months. These reports could be patient accrual rates or sales data. Suppose you are only given data for the last three months, but need to include cumulative statistics and graphs. Furthermore, the descriptive statistics and graphs in each report is the same.

   Write a function that will just accept the most recent quarter's data and then use a `for` loop to load all of the previous datasets. Use the functions `textplot()` in the package

gplots (`textplot()` is used for displaying text and matrices on graphs) and `layout()` to create a PDF file that summarizes the results. You could also use `Sweave()` to create the report (`Sweave()` is a tool for inserting R output into a L^AT_EX document).

So that each quarter you simply pass the most recent data to your function and you get your report of the current and past results.

3. *Table 1* Usually a manuscript for a clinical trial includes a Table 1 that gives summary statistics of the demographic information of the patients in the trial. Write a function that automatically creates this table for a given dataset. This function needs to be flexible enough to handle a variety of data types and summary statistics. A formula interface maybe very useful in this case. See `formula()` and `model.frame()`.

4. *Plot Methods* Currently there are plot methods for objects returned by a linear model, objects of class `lm`. However there are not plot methods for objects returned by hypothesis tests like `t.test()` that return an object of class `htest`. Possible plots for a `htest` object could be a histogram for one-sample t-test, side-by-side box plots for two-sample t-tests or variance tests, normal quantile-quantile plot for normality tests etc. The type of plot would be determined by the type of test and the arguments for that test. You may need to create S4 classes and methods, so that it is clear what information is returned by each hypothesis test.

Try not to limit yourself to the functions covered in this class. If you have a great idea and want to know or think there may be tools in R that will help let me know. I may be able to point you in the right direction.