Intro to R Pre-Assignment

Brian's three lectures will use R. It will be useful to have R installed on your machine and know a little about it. This assignment should be completed by Tuesday (2/28/12) before class.

Getting help on getting started

Feel free to contact Nathan by any of the methods below:

Skype: nathan.vanhoudnos
Email: nathanvan@gmail.com
Call/Text: 217 239 3578

Skype's screen sharing feature is particularly useful for helping fix issues. Instructions on screen sharing are here:

>   https://support.skype.com/en-us/faq/FA10215/How-do-I-share-my-screen

Installing R and RStudio

R is an interpreted programing language that makes it easy to think about statistics instead of thinking about programming. Unlike other programming languages, R is commonly used by typing commands one-at-a-time in an interactive session.

RStudio is a program that facilitates this interactive use of R. We'll install both.

Getting R

R is free. It is distributed by a consortium that calls itself CRAN (Comprehensive R Archive Network). The CMU Statistics department hosts a CRAN mirror, so getting R from campus is particularly fast.

R's homepage is here: http://www.r-project.org/. Feel free to poke around if you like. To install it, pick the appropriate link form the options below. Download the file and install it. The defaults will be fine, just click through it.

>   Mac: http://lib.stat.cmu.edu/R/CRAN/bin/macosx/R-2.14.1.pkg
>   Windows: http://lib.stat.cmu.edu/R/CRAN/bin/windows/base/R-2.14.1-win.exe

Getting RStudio

RStudio is also free. It makes it easy to use R for both one-off projects and extended development.

Its homepage is http://rstudio.org/. Feel free to watch the screencast and poke around. To install it go here:

>   http://rstudio.org/download/desktop

That page will automagically detect your operating system and suggest the appropriate version. Download the suggested file and install it. Once again, all of the defaults will be fine.

RStudio should automatically find your R installation and get you ready to roll.

Completing the Pre-Assignment

The zip which contained this PDF also contains a .R file and a .csv file.

Unzip the archive to a reasonable location on your file system. Open the .R file with a double click. RStudio should open, R should start, and you should get a screen that looks like this:
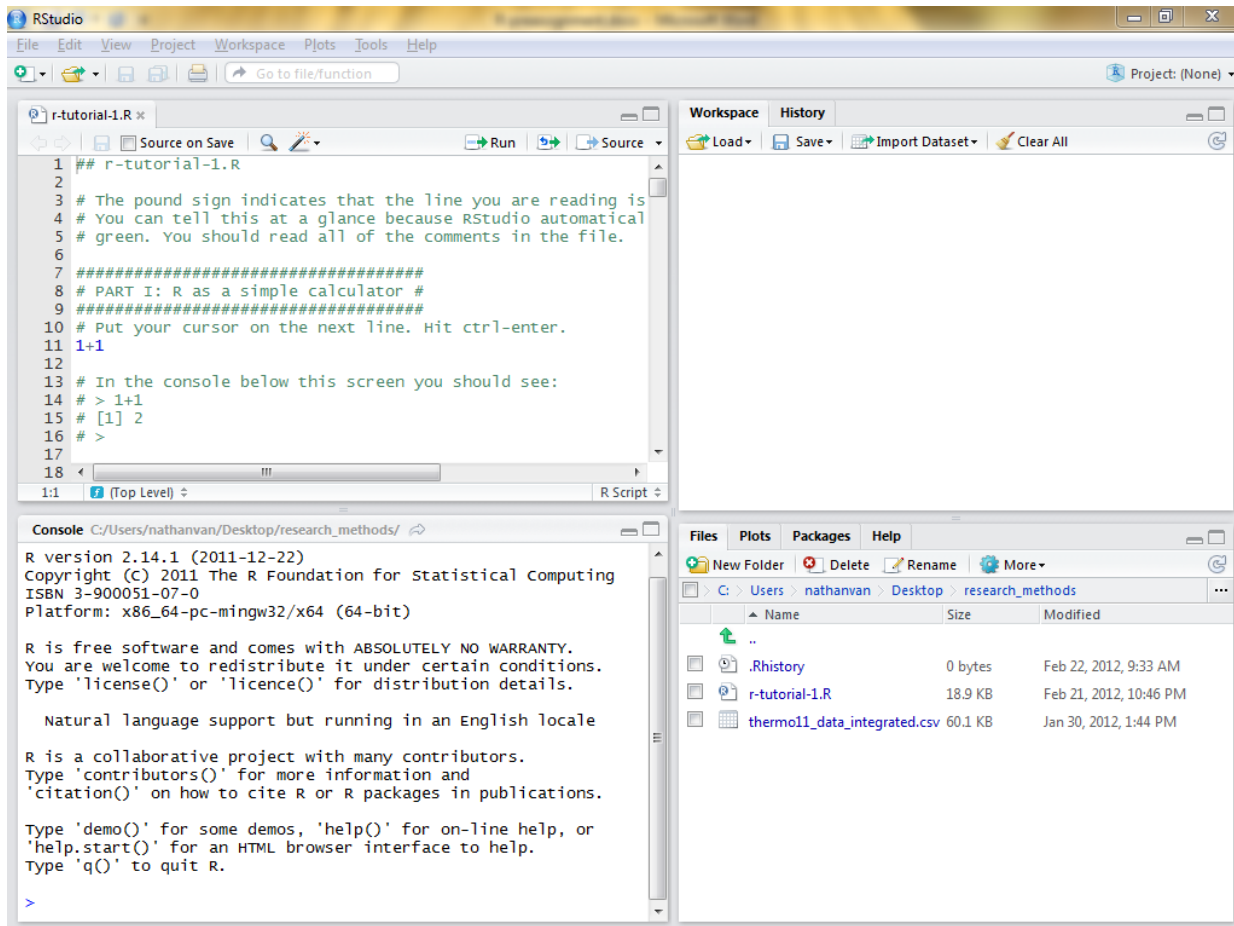
The top left sub window contains the text of the file you just opened.
The bottom left sub window contains the R console. This is where many of your results will appear.
The top right sub window can essentially be ignored for now.
The bottom right sub window, which currently shows the file system structure, is where the plots you make will appear. It also is where the help will appear.

Resize the sub windows as appropriate for your screen. Now do the assignment by reading the top left sub window and doing what it says.

What to do if you are curious to learn more

I can only really recommend web based resources because that's what I am most familiar with. A few good ones:

- http://www.r-bloggers.com/   Their daily digest email (or RSS feed) is useful for learning both how to use R and what kinds of interesting things people are doing with it.
- Quick-R: http://www.statmethods.net/ This is essentially an e-book that goes over basic R stuff quickly. An example of their page on Frequencies and Crosstabs
    - http://www.statmethods.net/stats/frequencies.html
- Stack-Overflow: http://stackoverflow.com/ This is a question and answer site with a thriving R community. Google indexes the questions so you might end up here on your own. Don't be afraid to ask a question here, responses come within 15 minutes. An example of a question I asked:
    - http://stackoverflow.com/questions/9100965/can-boxplot-in-base-r-display-na-when-the-grouping-factor-has-missing-values
- Cross-Validated: http://stats.stackexchange.com/ This is another Q&A site in the same network, but focuses on statistical questions, not software questions. This is good place to ask about high level things. People are friendly and helpful. An example of a question with a particularly good answer:
    - http://stats.stackexchange.com/questions/13314/is-r-squared-useful-or-dangerous