1 Overview and Course Objectives

This course is designed to get you over the basic hurdles you will face when beginning to learn R. It will cover some of the basic tasks that you face as quantitative researchers and will put you in a position to extend your knowledge of R in whatever direction is required to meet your own needs.

This course is has both lecture and lab components. As with anything in the Summer Program, and perhaps even more so here, the labs are optional. The Labs will be lead by the course TA, Kelly Gleason, who is very knowledgable about R as well as a helpful and patient instructor. Here’s what Kelly has to say about the labs:

1.1 Labs

In the aim of supplementing the Introduction to R nightly lectures with a more hands-on learning experience, we will offer a series of optional lab sessions which will incorporate many of the techniques covered in step with the class. The lab is structured to be
relatively unguided, providing participants the opportunity to begin digging into the R Statistical computing environment at your own pace. During each lab session I will hand out exercises for you to complete independently (or collaboratively with those seated around you in the lab). I will be on hand at this time to answer questions, help with problems, and hopefully ease the almost inevitable explosive frustration experienced when your code will not run due to the missing comma or misplaced capital letter. The labs will be offered at two different times during each day to accommodate most participants’ busy schedules and will be identical in content. Of course, if you attend an earlier lab and wish to receive additional help, you are more than welcome to attend the second lab on the same day.

In addition to providing a setting in which you may practice what you are learning in the R lectures, I invite you to also use these sessions for general R support for any problems you may be facing with R code for homework assignments from the workshops you are attending. While I will almost certainly NOT be able to speak to the substantive content of the results you are generating for the multitude of advanced models you are implementing, I may be able to sort out the error message that has kept you stuck on the same problem for hours. Please also take advantage of my office hours for the same purposes. I specifically set up Sunday office hours to assist participants that are having problems with code on homework assignments that are due on Mondays.

One final note: While I strongly encourage your attendance in both the Introduction to R lectures and the lab sessions, it has not escaped our attention that the overwhelming majority of participants enrolled in ICPSR workshops are faced with massive time constraints with all of the work you must keep up with day to day. If you have to miss an R lecture here and there, or attend lab infrequently, please do not let that dissuade you from coming in for help during any future lab session or office hours! We are aware that that there are participants beginning the summer session with some working knowledge of R and only start to face problems as coding in the homework assignments for your classes becomes more difficult. Feel free to avail yourself of the resources we can offer at any time.

1.2 Lectures

In a departure from previous iterations of the course, I have added an extra week (well, four days really) of instruction. The four days of the course (Tuesday-Friday of week 1) will be aimed at true beginners - those people who want to learn R, but do not have much or any experience at all with a code-based environment like R (as opposed to a menu-based environment, like SPSS). This is intended to provide those who need a slower, gentler introduction to the software a chance to get their feet wet before we deal in more detail with data and models. Those who have some prior experience with R or who have experience with another code-based language (e.g., Sas or Stata) may want to skip the first week and start coming on Monday of the second week. Essentially, if you are confident you can learn R, but just need a little push, start in the second week. If you are less confident about your ability to learn R or you need a bigger push, start
coming in the first week.

The lectures will also hopefully be applied enough that you will immediately see the utility of what we talk about, though that will almost certainly not always be the case. Students with laptops are welcomed to bring them to class to follow along with the notes as much as they can. The materials for the course will be made available as will all of the code to generate all of the results on the course webpage: http://www.quantoid.net/ICPSRR.html as well as on the Summer Program Z drive. Both the instructor and teaching assistant will be available outside class to help students with any specific problems.

There is no formal assessment in the course (i.e., assignments), though there will periodically be exercises on which students can work outside of class to develop their knowledge and understanding of R.

Outline

1. Introduction to Week 1 (Tuesday, June 24)
   - Finding, Downloading and Installing R.
   - First looks - dealing with the user interface.
   - Changing the default look.
   - Where are R’s files stored?

2. Packages, Functions and Syntax (Wednesday, June 25)
   - Finding and installing packages on CRAN.
   - Installing packages from other sources: Git, R-forge, Source
   - The search path and how R finds your functions.
   - Understanding function syntax (argument types and options).

3. Input and Output (Thursday, June 26)
   - Understanding the object oriented nature of R.
   - Processing output - object classes and their print, summary and plot methods.

4. HELP! (Friday, June 27)
   - How to figure out where things are going wrong.
   - Warnings, errors and other messages.
   - Getting help from other people.
   - Getting help from the web.
5. Getting to Know **R** and Data Types. (Monday, June 30 and Tuesday, July 1)
   - R Basics (working with an object oriented language)
   - Vectors and matrices
   - Reading in data from other statistical programs
   - Different data types in R.
   - Basic data management (recoding, transforming, generating new variables)
   - Filtering/subsetting with logical expressions

6. Basic Statistical Routines (Wednesday, July 2)
   - Factors and contrasts
   - Cross-tabs
   - Linear Models
   - GLMs

7. Intro to Traditional Graphics. (Thursday, July 3)
   - Overview of different graphics systems
   - Making graphs in the traditional graphics environment.

8. Basic Function Writing (Monday, July 7)
   - Writing basic functions
   - Repeated calculations - loops and apply
   - Merging Datasets

9. Lattice Graphs (Tuesday, July 8)
   - Difference between lattice graphs and traditional graphs
   - Making graphs in the Lattice system.

10. The Effects Package (Wednesday, July 9)
    - Effects for linear models with interactions
    - Effects for non-linear models

11. **ggplot2** (Thursday, July 10)
    - Model-based graphing
    - Difference between **ggplot2** and the other graphical packages

12. R and **LaTeX** with Sweave (Friday, July 11)