Why use the R Language?

R is a dialect of the S language, and has come to be – by far – the dominant dialect. S started as a research project at Bell Labs a few decades ago; it is a language that was developed for data analysis, statistical modeling, simulation and graphics. However, it is a general purpose language with some powerful features – it could (and does) have uses far removed from data analysis. It should be used for many of the tasks that spreadsheets are currently used for. If a task is non-trivial to do in a spreadsheet, then almost always it would more productively (and safely) be done with R.

Given its qualities, R has become the preferred computing environment for a large part of the statistical community. When a new statistical method is invented, chances are it will be implemented first in R. In March 1999 John Chambers – one of the originators of S at Bell Labs – was presented the ACM Software System Award. It stated, "S has forever altered the way people analyze, visualize, and manipulate data." Previous winners of this award include Unix, TeX and the World-Wide Web. John is now a member of R Core (the group that produces R).

Installing R

- Go to the main R website: http://www.r-project.org
- Click on the CRAN button on the left side.
- Select a mirror near you from the menu that is offered.
- Select which type of machine you have (Linux, MacOS X, Windows).
- Select base and click to download the current version of R.
- Follow the prompts to finish installation (generally, keeping the default settings is preferred).

RStudio

RStudio is an open-source Integrated Development Environment (IDE) that provides a more consistent user experience to R. There are many great features of RStudio over "basic" R, including:

- Consistent windowing between sessions (customizable by the user)
- Point-and-click exploration of data frames and other data objects
- Importing data files through dialog box functionality
- Customizable code syntax highlighting, auto-complete, and Help menu access from the code editor
- Ability to see all installed packages, turn on packages using a checkbox, and download libraries (and their dependencies) without having to write any code

Installing RStudio

RStudio installs like any other program for Windows or Mac OSX. As far as I can tell, there are no advantages to using RStudio in either environment, both the Windows and OSX versions seem to work equally well. The most important consideration is that RStudio is just an "add-on" so-to-speak, it does not include R itself. So be sure to go to one of the Comprehensive R Archive Network (CRAN) sites to download R first.

- Go to the website: www.rstudio.com
- Click the button "Download RStudio"
- On the next page, scroll down about half a page and click the button "Download RStudio Desktop"
- On the next page, click on the appropriate link (for your system) to download the installer.

Importing Data to RStudio

Now that you have R and RStudio on your computer, you can start working with data. It is very cumbersome to enter data in the R console manually if the data set is large. We can download datasets from many different file types into RStudio. You need to have the dataset saved on your computer in order to import the data into RStudio.

- Open RStudio
- In the upper right frame, select the tab called "Environment". The default is to have this selected at start-up.
- Click on the button "Import Dataset" and select "From Text File" on the drop down menu.
- A window will open allowing you to browse your computer and select the data file. It is good practice to save them all in a single folder.
- Once you have selected a file, a pop-up will open asking your to select different options. RStudio does its best to anticipate file type, so you do not generally need to change any of these.
- Click "Import" and the file is loaded into RStudio and you may begin working with it. You should see it appear in the "Environments" tab in the upper right frame.