

R Tutorial

Jake Feldman

August 17, 2015

1 Installing R

The first step is installing R. You can download the latest version at:

- For Windows: <https://cran.r-project.org/bin/windows/base/>
- For Mac: <https://cran.r-project.org/bin/macosx/>

Next, I would highly recommend downloading RStudio which is a tool that allows you to be more productive in R as well as manage R's many packages in a simple manner. RStudio can be downloaded at <https://www.rstudio.com/products/rstudio/download/>. Once you have downloaded R and RStudio, open RStudio. On the left is a panel known as the console, think of it as a fancy calculator. This is where we will do everything. Type `1+1` in the console to check that everything is working...you should get 2.

2 Reading in Excel Worksheet

This tutorial will work with the data from the file `Applicants.xls` so download this file and save it wherever you would like. Let's assume that you have saved it to your Desktop. Before we can read the data into R we have to tell R where our excel is. In other words, we have to change our working directory to location of `Applicants.xls`. There are two ways to do this. The first is directly through the console and uses the command `setwd()` which stands for set working directory. The input for command is the file path to the spreadsheet. So, in our case this would be `setwd("~/Desktop")`, where the `~` is a shortcut for your home directory. The second way is through the panel on the bottom right of RStudio, where you should see a collection of folders. Navigate to the location of `Applicants.xls` by clicking on the appropriate folders. Once you are in the correct place, on the top menu bar go `Session → Set Working Directory → To File Pane Location`. This will automatically give you the correct `setwd()` command.

Before we issue the command to read in the data we will have to install a package called `gdata`. You can view all the packages you have installed through the bottom right window. Click the packages tab and search `gdata`. If you have it installed simply check the box. If not, go to `Tools → Install Packages` in the top menu bar and then search for `gdata`. Once it is downloaded you will see it under packages in the bottom right portion of RStudio so all you have to do is check the box.

Now we are ready to read in the excel data into an R data frame, which is a data structure (special way to store the data) that allows us to easily filter and sort the data. To read in the data from sheet one into a data frame type `df = read.xls("Applicants.xlsx", sheet = 1, header = TRUE)`, since the data is in the first sheet. You will often see `< -` used in place of the equal sign in R, you can treat them as the same thing. Now we have all the data loaded in to a dataframe which have given the name `df`, lets see how to manipulate the data.