Reading Guide (Chp 1)

Dan Maxwell

May 3, 2017

Section Notes for *R in Action*

1.2 Obtaining and Installing R
You must install base R before installing RStudio. In fact, RStudio won't work without base R installed.

1.3 Working with R In this section, Kabacoff assumes that the reader already knows what a function is and how they are called. As defined by Webopedia, a software function is "a named section of a program that performs a specific task." By now, you should have already read the section on functions, written by Peter Dalgaard. A link to that document is provided in Canvas.

1.3.1 Getting Started
Because we are using the RStudio integrated development environment (IDE), what you see in the RStudio console will differ from Figure 1.3.

1.3.3 The workspace
For newcomers, the getwd () and setwd () functions are critically important because R sources files from your current working directory. Inevitably, students forget where they're at and then get an error message when trying to load a file with one of the read functions.

```
setwd("c:/Informatics")  # Set the working directory -- This is now the default directory
getwd()                  # Get the working directory -- Where are we at?
```

```
## [1] "c:/Informatics"
```

1.7 Working with large datasets
Appendix F is worth the read, especially if you think you'll be working with very large datasets. Memory management in R is unique in that function arguments are passed by *value* and not by *reference*. What this means, in practical terms, is that R makes a copy of each object you pass as an argument to a function. So if you pass a dataframe to a function, you now have two copies of that dataframe in memory, one for the function and another for the program that called it. Because of this, R will gobble up a lot of memory when your datasets (dataframes) are really large.