Reading Guide

Dan Maxwell

May 3, 2017

Section Notes for *R in Action*

2.2 Data Structures
Data structures and/or datatypes are incredibly important so read this section slowly and carefully. Many times when R raises an error or you are getting unexpected results, the problem can be traced back to the datatype. For example, a function might be expecting a number as one of its arguments, and you pass it a character string instead. I often use `class()` to see what the underlying datatype is. Kabacoff does not discuss `class()` here but presents it later in the book.

```
char_string <- "This is a string of characters."

class(char_string)  # Use the class function to ascertain the datatype.
```

```
## [1] "character"
```

```
nums <- c(1, 2, 3, 4, 5, -2, -4)  # A set of numbers.

class(nums)
```

```
## [1] "numeric"
```

2.2.2 Matrices Note well: matrices can "contain only one datatype." And because of this, data frames are much more common in R.

2.2.4 Data frames
The `attach()` function can create issues, especially in situations where you forget to `detach()` a dataframe after use. The problem is that R allows you to attach a dataframe more than once. When this happens, you cannot know which dataframe you're working with as the names are the same.

The <<- operator can also be used within a function to reference external variables, but this practice should be avoided.

2.3.9 – 2.3.12 These sections can be skipped, unless you're trying to import data in specialized formats. Specifically, section 2.3.11 is somewhat dated in that ODBC (Open Database Connectivity) technology dates back to the 1990's and is used less frequently today. If you're using MySQL, you'll want to use the RMySQL package to access the database directly. You can now find R packages to directly access most of the major relational databases.