Introduction to R

Christopher Maronga

February 21, 2019

R Packages

An R package is a collection of R functions and/or data with a well compiled code(s) that enable R users perform even more tasks with R(expand functionality). Examples of R packages include dplyr, ggplot2 etc. There are currently over 13,000 contributed packages in the Comprehensive R Archive Network(known as CRAN). learn more

Common commands

- install.packages("package name") # to install a package
- library(package name) # load an R package into the workspace for use, also `require(package name)`

Many more commands are available for dealing with packages

Managing data using R and Basic R procedures/commands

The primary goal of any data collection exercise (i.e. finance or even health research) is to have a sound output dataset, and this can only be achieved via good data management practices.

All data analyses rely 99.9% on data that has been well management for effective reporting.

R is the perfect platform to perform all those tasks. It provides a wide range of packages to manage your data as well operate on it. Basic R commands you encounter in R includes:-

- `str()/dim()`
- `summary()/table()`
- `nrow()/ncol()`

NOTE: These are all base R functions, we will encounter most of them in practice along the course and explain the distinction between each of them and when to use them.

The Concept of global enviroment and variable refrencing

You can reference variables in any dataset in R using the following procedures:-

- use of `$` operator (dollar sign)
- use of **`with()`** function
- use of **`attach()/detach()`** strategy
- use of pipe operator **%>%** *we will dwell on this on later chapters*

Let demonstrate and explain the first three ways of easening our working with dataset variables in R using real examples