# R: Some Helpful Vocabulary

We often think about a data analysis program as a single program that is developed by one company and is self-contained. You install the program once and it has everything you need to run it. The various components of the program are well-integrated and the conventions are consistent across the different parts of the program.

R was built as an open-source platform that allows people to easily collaborate. Anyone can rapidly develop complex analysis using a very flexible programming framework. This analysis is shared with other users either through sharing code (scripts), or through packaging that code and sharing them as programs (packages).

As a result, it is more useful to think about R as an operating system rather than a program. Just like Windows makes it possible for your computer to run a bunch of programs that were not written by Microsoft, R makes it possible to run a bunch of programs that were developed by members of the community. In this way R is as much a community as it is a program. The "R Environment" is a set of protocols that people use to share their work. Here is some vocabulary that will help you make sense of the R paradigm:

R Environment

↑

Packages

↑

Functions

↑

Arguments
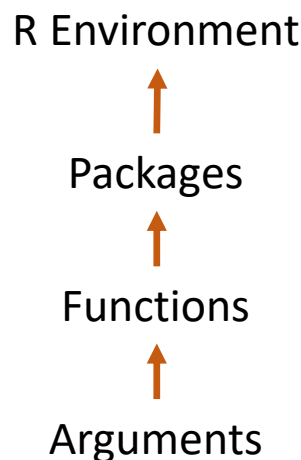
The R Environment

In software terms, an environment is the term used for the things that are needed to run an application. On a laptop, the environment for a specific desktop application might include the operating system, a database, and a compiler. R is called an environment because it has all of the elements necessary to run programs written for R.

Package

R programs are called "packages" and they are loaded by the "library" command. The term comes from the fact that each package (program) is comprised of a bunch of functions. They are organized into a library, and this library is shared by packaging it all together into a single entity.

Function

A function is a small program that does one specific task or calculation. For example, it could be a program that calculates the average of a bunch of numbers. It could be a program that changes the temperature in Celsius to the temperature in Fahrenheit. It could also be a program that creates a graphic. Any task that you do in R is accomplished by a function, and your analysis will consist of a bunch of functions used to manipulate your data.

Argument

Each function is like a recipe that needs ingredients in order to work (see the example below). The ingredients are either data that you give to the function, or specific parameters it needs in order to run. For example, if you want to calculate a mortgage payment you need to know the interest rate (a parameter), otherwise the calculation will not be possible.

Script

A script is a short program that accomplishes some analysis using several functions.

Integrated Development Environment (IDE)

A software application that provides a comprehensive set of facilities needed to effectively develop and test code. R Studio is one of many IDE's used by R programmers. It is unique in that it was developed specifically for R.

CRAN

Comprehensive R Archive Network, or CRAN, is the online repository that houses all packages written for R. There are currently over 7000 packages available for download. You can browse all packages here: http://cran.r-project.org/web/packages/

Mirror

A local server used to download R packages.

Path

The address of a file in the directory structure.

Object-Oriented Programming

A programming paradigm that breaks chunks of code into "objects" to make them re-usable and robust.

Class

The type of object in the R environment.

Mode

The underlying data structure of a class.