Day 3, Session 1: Installing R and RStudio

Jessica Williams-Nguyen and Brian D. Williamson

EPI/BIOST Bootcamp 2017

26 September 2017

R is a free, open source software package that can be used for data analysis, graphics, and programming.

At its core, R is an interactive, command-driven language: you type a command and R executes it and returns results.

While R is sometimes said to have a steep learning curve, it is relatively easy to get set up with the basics and analyze some data.

Why R?

R has many advantages, including:

- Free and open source!
- Active group of contributors (anyone!)
- Flexible
- Large set of packages for data analysis

However, this comes with some challenges:

- Sometimes packages don't do what they say they do...
  - ...but you can trust basically anything written by the R Core Team, the RStudio Team, Hadley Wickham, or Yihui Xie

RStudio

RStudio is both an integrated development environment (IDE) and a graphical user interface (GUI) for R programming and data analysis. The free version is also open source.

It includes a console, text editor that allows for direct execution of code, as well as tools for importing/exporting data, plotting, file management, and debugging. (We will cover all of these terms later!)

Why RStudio?

The base R GUI is both light and functional. Sometimes, however, we want more than that!

RStudio adds:

- An improved layout of the console, text editor, and other tools
- Support for embedding reproducible research tools
- Support for building your own R packages
- Integrated R help and documentation
- A pretty good debugger

The only downside to RStudio is that it takes a decent amount of memory to run... but for most purposes, this isn't a problem.

Why two programs?

R is a programming software, prepackaged with a GUI. However, R programs can be executed from the command line without an interactive interface.

RStudio is a GUI, and is a helpful tool for working in R. Using RStudio makes it easier to:

- write R scripts to save your work, along with comments for what your code does
- write reports with code embedded (using Rmarkdown)
- organize your data analysis workflow (e.g., reading in data, access help files)

Why two programs?

At the end of the day: you are executing commands/programs in R, but using RStudio as an intuitive interface to the software (much like your operating system is a GUI to the machine language that your computer understands).

The combination of R and RStudio makes reproducible research attainable by everyday users. The RStudio environment has many easy options to facilitate this through R, and a lot of support behind the scenes.

Installing R

1. Go to https://cran.r-project.org/

2. Choose the correct link under Download and Install R
   - Windows users, select install R for the first time
   - Mac users, click the R-[replace with most recent version number].pkg link and install
   - Linux users, I assume you know what you are doing

Installing RStudio

1. Go to `https://www.rstudio.com/`

2. Scroll down until you see the headers for `RStudio`, `Shiny`, and `R Packages` (see figure)

3. Click Download

4. Click the green download button in the column for `RStudio Desktop`

5. Choose the correct installer for your operating system and click the link