

A-State R User Group
Spring 2023 – 28 March 2023
Basics of R

1. What is R?
 - a. A programming language and software that interprets it
2. What is R Studio?
 - a. Software to write scripts and interact with the R environment
3. Has everyone installed R and RStudio?
 - a. If not... https://intro2r.com/install_r.html
4. RStudio interface
 - a. Script (only appears when you create a new script) – editable, savable, record keeping, sharable, etc...
 - b. Console (where R evaluates all the code you write) – not editable once code has been sent
 - c. Environment/history/connections – objects you’ve created, history of commands you’ve sent, connect to external databases
 - d. Files/plots/packages/help/viewer – files on your computer, plots will be displayed here, packages show you your installed and loaded packages, help documentation, viewer (graphics displayed by some packages)
 - i. The base installation of R comes with many useful and standard packages
 - ii. Many thousands of specialized packages are available to install and use. We talked about some of them at our last session. We’ll install and use a few today.
 1. Can install packages from CRAN, Bioconductor, GitHub, maybe others?
5. Let’s start an R Project
 - a. Working directories
 - i. the default location (PATH) where R will look for files you want to load and save any files you create. It is shown at the top of the console tab or can use `getwd()` to display it. You can also set it by `setwd()` or `Session -> set wd -> define the path`
 - b. File names – keep a consistent naming convention and stick to it! Do not use spaces!
6. Let’s do some basic math in the **console**
 - a. ``2 + 2``
 - b. ``11-4``
 - c. ``sqrt(4)``
 - d. the `[1]` is just R telling you that this is the first observation of the results (more important when you have many lines!)
 - e. The ``>`` tells you R is ready for you to enter code
 - i. Try ``11-`` and run it, what do you see now? Should see a “+” which just means you didn’t finish telling R what you wanted (i.e., the code is incomplete) so R gives you another chance to finish it. If you don’t want

to finish it (e.g., because of a mistake) press the ESC key on your keyboard.

7. Let's start an R script
 - a. The console is useful for quickly checking things, maybe if you just want to get a quick result, etc. but almost everything you'll want to do in R is going to be based on a script
 - b. Project documentation
 - i. Comments (#), R does not read these
 - ii. Name, date created, date modified, purpose of script, packages needed to run it (and versions), all your code, and then `xfun::session_info()` to include everything needed to run the script, your machine info, etc... helps with de-bugging
 - iii. Save this right away!
 - iv. We'll learn more about RMarkdown in a few weeks, which is another option for project documentation and reproducibility...
 - v. Can also create a README text file (file -> new file -> text file) for metadata, data collection, data equipment, whatever else is important information about the analyses, etc.
 - vi. Organization based on #####
 - c. Almost everything else you'll want to do in R is related to R objects
 - i. Objects can be numbers, strings, outputs of plots, summary of statistical tasks, one line to run an entire code, etc.
 - ii. Let's make some objects
 - iii. Naming objects is VERY important and is hard to do, especially when you are unsure of the downstream analyses you'll need!
 1. Make them specific
 2. Don't name objects the same name as a built-in function such as ``sqrt`` or ``data``
 - d. Let's talk about vectors and a function called `c()`, which means concatenate
 - i. Functions are always the name of the function followed by parentheses, even if the parentheses are empty.
 - ii. Arguments are located within the parentheses. Arguments sort of customize the use or behavior of a function and each argument is separated by commas
 - iii. Let's make a vector and a few functions on it
8. Let's look at some data and import our own
 - a. There are several 'toy' (pre-loaded) datasets already installed in R, one is called ``iris``
 - i. Basic viewing, subsetting, base R plotting
 - b. Import the 'flowers.csv' file
 - i. Basic data manipulation, ggplotting
9. If we've made it this far on time, can play around with the dataset(s) or ask questions