Data Inquiry Lab: Introduction to R

This introduction will span approximately 75 minutes, plus additional time for questions and answers. The workshop is intended for those starting at 'R knowledge=0'. If you have suggestions for additional R workshop topics, please let me know.

What is R? R is both an application for data analysis and graphics, like SPSS, but also a flexible programming language. Unlike proprietary software, R is free to use and popular all over the world. It is arguably the *lingua franca* of statistical programming languages and worth your time to learn. Like most worthwhile endeavors, it is not easy. It takes patience and persistence. But you can do it! And we'll get started together.

Learning Objectives

By the end of this workshop, you should be able to

1. Open RStudio. Identify the Console, Script, Environment, and Plots pane.
2. Load a script file, run lines from it, edit and save the script file. Set a working directory.
3. Load a Workspace containing an R dataframe, edit the dataset, and save the Workspace.
4. Install an R 'package'; import data from .csv (comma separated value) file format and .xlsx (Excel) file format.
5. Type your own R code to use the API (application programming interface) of the World Bank to download and analyze global development indicators.
6. Create a 'GapMinder' style plot.
7. Create univariate visualizations with two different R packages.
8. Identify books, websites, and additional sources for further learning and help.

R Installation

We will use the R + RStudio installation on campus computers, in the "Statistics" Department applications folder. *While in the workshop we will use lab computers, Murphy's Law seems to always be in play when using these networked applications.* Download R to your own computer, later. To get started with R on your own computer, you need to do one of the following:

1. Download R from `https://cran.r-project.org/`. Download the "base" version for Windows or the most recent version for Macs. The DIL introductory R guide explains this process in greater detail at `http://www.gvsu.edu/datainquirylab/r-resources-8.htm`.

   (a) Next you'll want to download RStudio, from https://www.rstudio.com/products/rstudio/download/#download, which is technically an "integrated development environment", meaning it is an R interface with lots of additional functionality. After downloading RStudio, open it to use R.

2. Another option is to download Microsoft's "enhanced" version of R, at https://mran.microsoft.com/open/. It is not as user-friendly for beginners as RStudio, so you'd want to review these instructions https://mran.microsoft.com/documents/getting-started/.

   (a) A complement working in base R or Microsoft's version is to use Jupyter Notebooks (http://jupyter.org/), which is a web-based application for typing R code, visualizations, reports, and other text. It looks great, but has a steeper learning curve than RStudio. Note that RStudio's 'Rmarkdown' system has the same report functionalities as Jupyter.

Cloud computing with R at GVSU

GVSU maintains a server for using RStudio in the cloud. You need to request an account prior to logging on at http://rstudio.stat.gvsu.edu. Contact Carl Strebel at strebelc@cis.gvsu.edu and see http://www.gvsu.edu/bigdata/.

Guide to R packages

See current R packages organized by subject heading at https://cran.r-project.org/web/views/.

help with R commands

Check out the RStudio Cheat Sheets at https://www.rstudio.com/resources/cheatsheets/. Other cheat sheets at http://cran.r-project.org/pub/R/doc/contrib/Short-refcard.pdf and common R commands at http://www.personality-project.org/r/r.commands.html.

GVSU library R resources

e-books Through the R library catalog, http://www.gvsu.edu/library/, search for "R statistics"; many books are available free. A good place to start is searching the catalog for "R for Dummies".

secret stash of R e-books Try the Springer series of R books, not all of which seem to appear from the library's catalog search: http://ezproxy.gvsu.edu/login?url=http://link.springer.com/

In your free time

Try an online R course. All the major MOOC providers, such as Coursera and EdX have R based classes with applications in many different subject areas. A free online introduction to R is https://www.datacamp.com/courses/free-introduction-to-r.

Where to go for help

What to do if you are stuck on an error message or unsure how to proceed?

1. Google. Copy and paste your error message into it.
2. Also try looking through http://stackoverflow.com/questions/tagged/r. Submit your own question; prepare yourself for initial rudeness, such as "lol n00b...this question has been asked already" type responses.
3. Even better, stop by the Data Inquiry Lab on campus, at the Mary I. Pew Library, during the Winter 2017 semester, 11am - 3 pm, Tuesday through Thursday. No appointment needed. Or make an appointment at http://www.gvsu.edu/datainquirylab.
4. We post R resources at http://www.gvsu.edu/datainquirylab/r-resources-8.htm.

Where to go for inspiration

R graph gallery http://www.r-graph-gallery.com/

popular R packages Google packages for your subject area, or see
https://support.rstudio.com/hc/en-us/articles/201057987

http://Plot.ly web visualization integration with R https://plot.ly/r/