1-01 Introduction & Preliminaries

CSI 500

Course material derived from:
An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics
Version 3.4.3 (2017-11-30)
https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf

Introduction to R

- What is R?
  - an integrated set of tools, data, and operators designed for mathematical and statistical applications.
  - A set of graphical tools designed for visualization of numeric data
- Related Software
  - R is derived from the S-Plus statistical language developed at AT&T
  - Syntax and semantics are similar
- R and Statistics
  - R includes a large set of classical statistical tools, including distributions and standard statistical tests
  - Most implemented in the base system, others available as add on "packages"

Introduction to R

- R and Windows
  - R is designed to be used interactively from a workstation
  - Runs on Linux, Windows, Mac
- Interacting with R
  - R is designed to be used in a "read-evaluate-print" style
  - R command prompt is the ">" character
  - to quit R session, type in "q()" at the prompt. Parens are necessary - q is a function call
  - R is case sensitive, so "Q()" won't work…
- R online help
  - R has a built-in help facility, invoked by calling "help(object)" or "?object"

Introduction to R

- R commands
  - R expects commands to be typed at the prompt
  - Multiple commands can be entered on a line, separated by a ";" character
  - Comments indicated by "#" character - ignores rest of line
  - No facility for multiline block comments like C or Java /* … */
  - Commands exceeding line length of screen result in a continuation prompt "+", which indicates R is expecting more typing
  - R commands are generally limited to 4095 characters
- R command recall and correction
  - R will recall previous commands using the up-arrow key
  - Cursor can be moved using right and left arrow keys, deleted with the DEL key

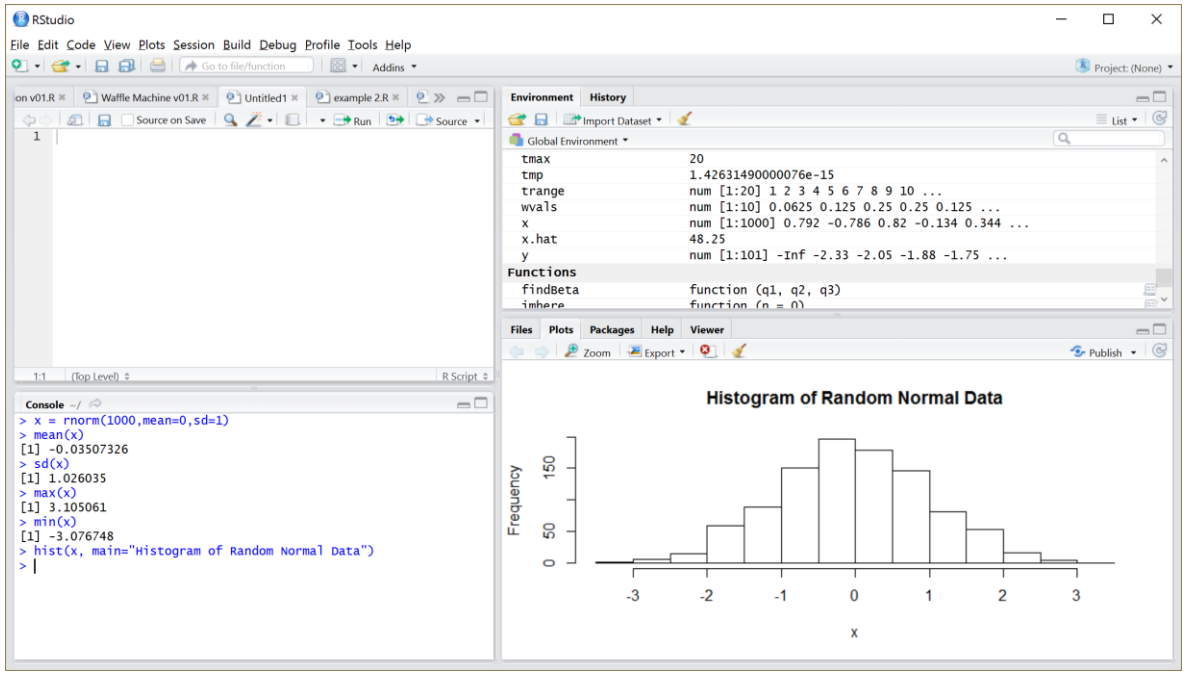
Introduction to R

- Executing commands from or diverting output to a file
  - A set of R commands can be executed from an external file using the source("filename") command. Quote are required as filename is a String.
  - All output can be redirected to a file using the sink("filename") command. Quotes are required as filename is a String.
  - To restore interactive output, enter sink() without a filename.
- Data permanency and removing objects
  - R sessions typically involve creating and updating objects. To view objects in your workspace, type objects(). The ls() command does the same thing.
  - to remove objects, use the rm(obj-1, obj-2, …) command.
  - Objects are stored, so the next time you start R it knows about all your prior work

Quick demonstration

- Start up Rstudio
  - At the command prompt in the lower left, type in these commands
  - Don't worry if you're not following along - this is a quick demo

```
x = rnorm( 1000, mean=0, sd=1 )

mean( x )

sd( x )

max( x )

min( x )

hist( x, main="histogram of random data" )
```



Summary

- R is a mathematical and statistical analysis system
- R is designed for interactive use in data analytics and statistics
- R has extensive built-in mathematical, statistical, and graphics capabilities