R Basics

August 2, 2018

Presenter: Helen Rogers

Helen.Rogerst@gmail.com

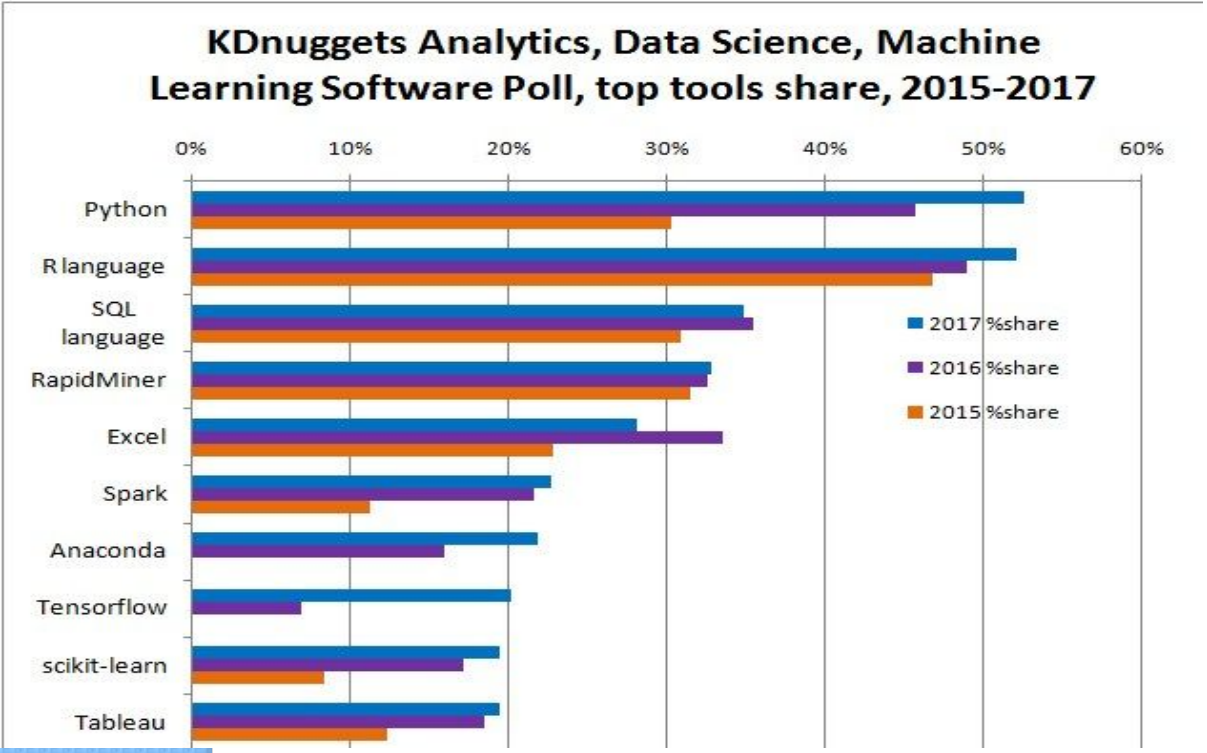
ASQ®

The Global Voice of Quality®

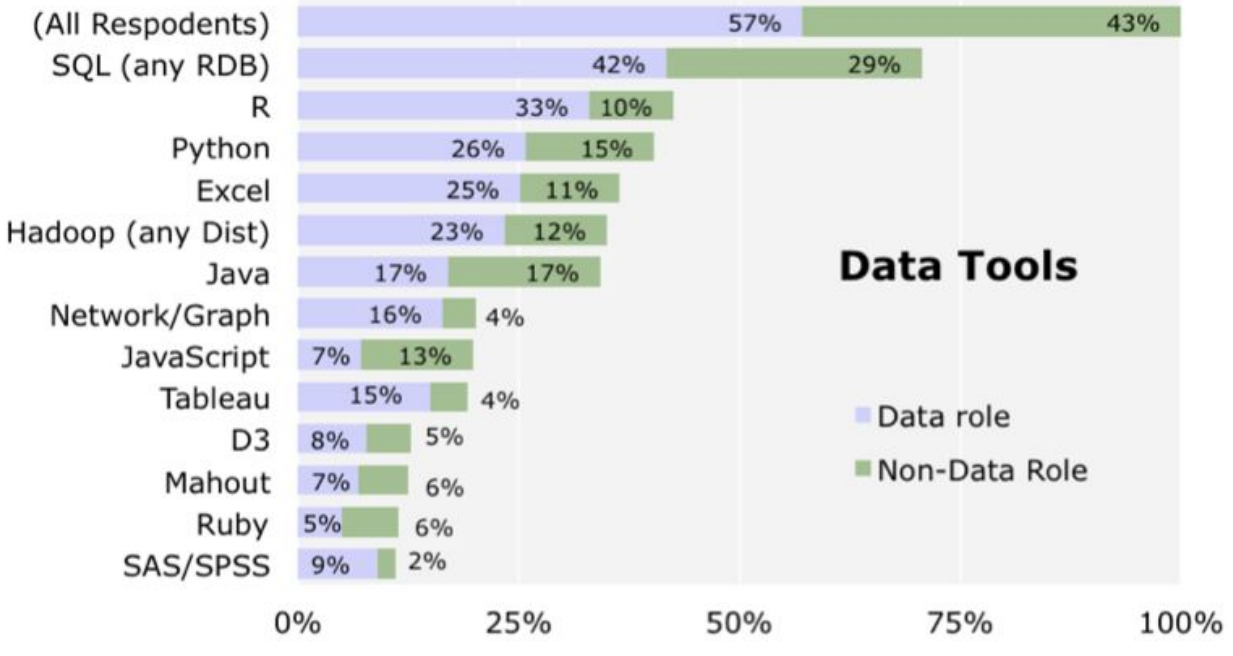
Introduction

- Background
  - About me
  - Intended for users who are unfamiliar with R

- What to expect from this tutorial:
  - Basics of R and data manipulation in R
  - Introduction to R Environment
  - Resources and where to go next



Top: KDNuggets Software Poll:2015-2017
Source
Right: O'Reilly Media: 2013 Data Science Salary Survey
Source

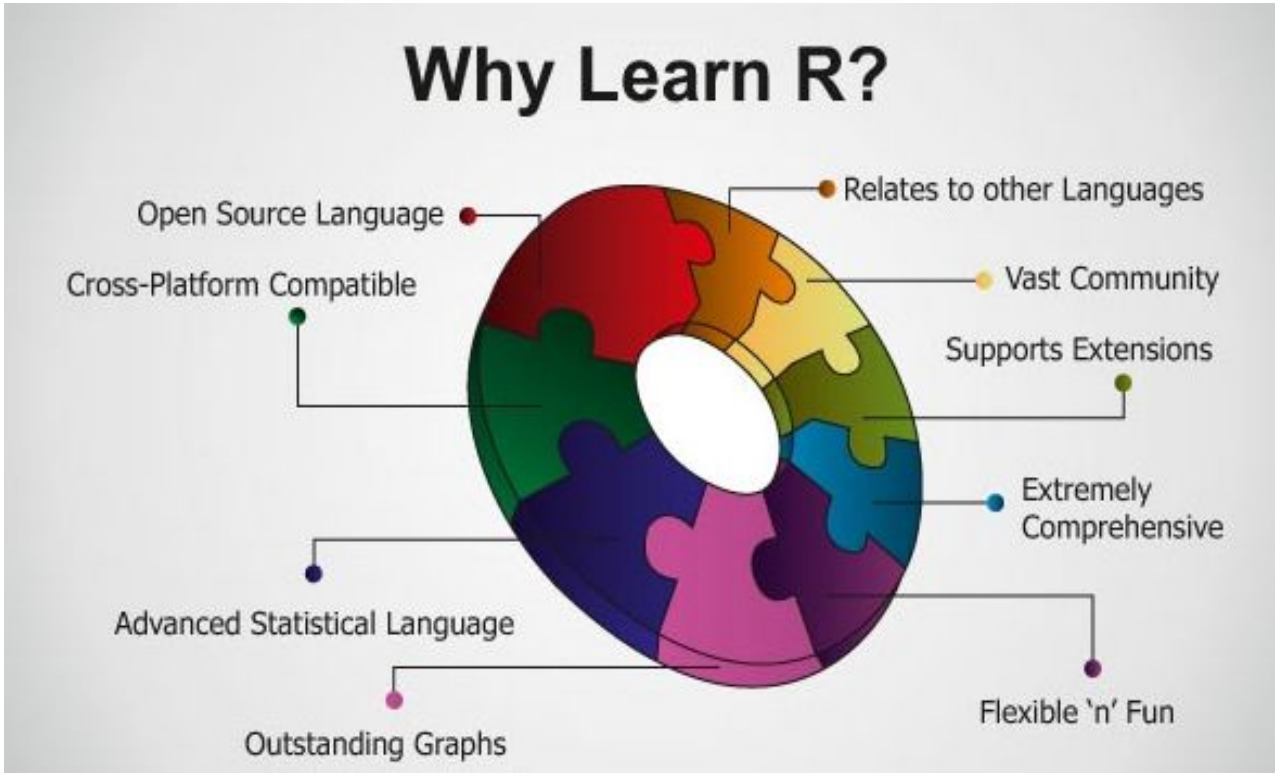


What is R?

- Programming language and environment
- Data management, statistical analysis, and graphics platform

Benefits

- Free, open source, and user friendly
- Reproducible and reusable
- OS Friendly
- Online community
- Maintained



Source

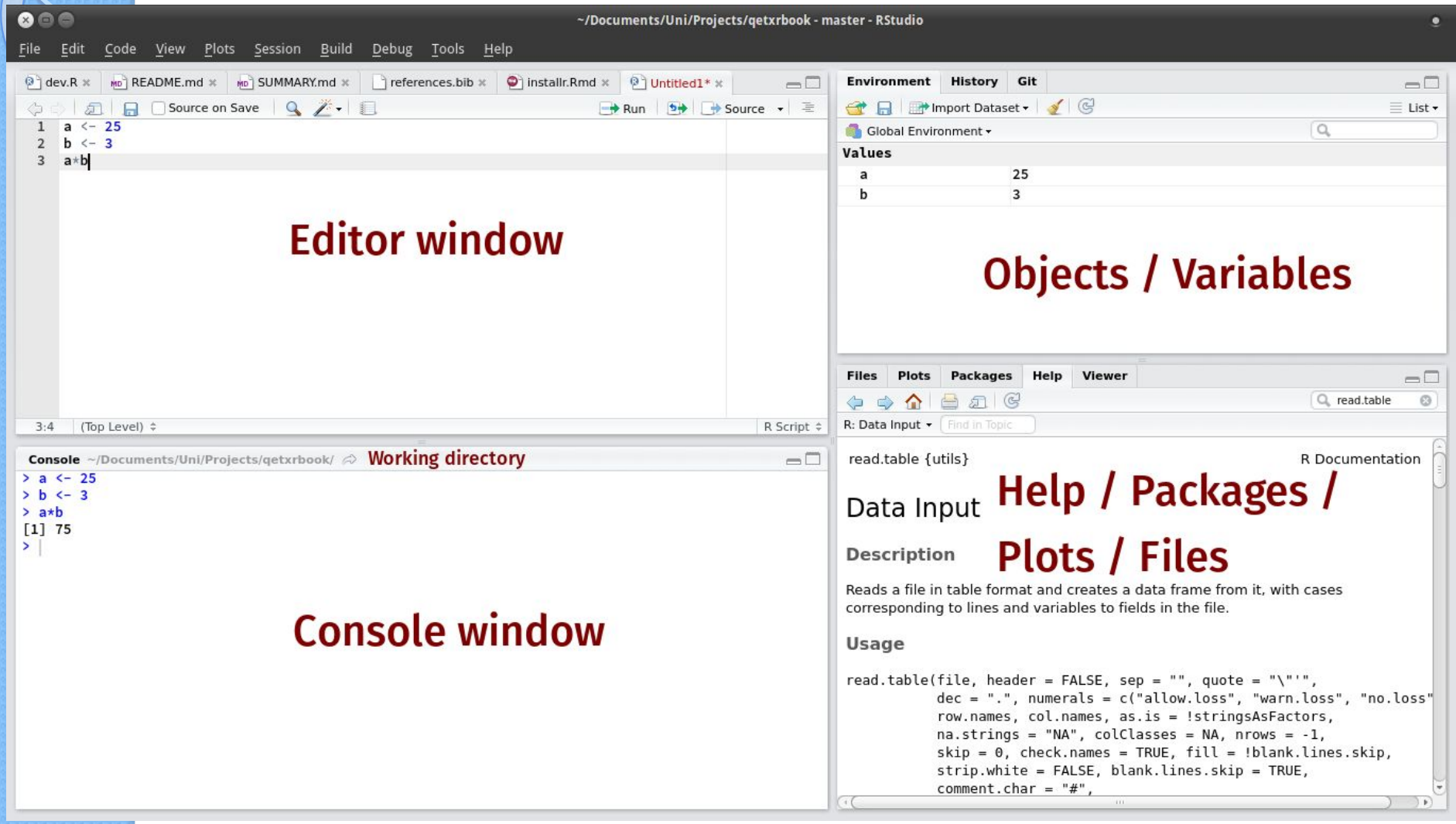
Let's get started - Installation

- R has free mirror links for download available on website
- https://cran.case.edu/
- Should take about a minute to complete installation

Rstudio Desktop Installation

- User-friendly environment for working
- https://www.rstudio.com/products/rstudio/download2/
- Integrated tools for R
  - Edits/highlights Syntax
  - Features - https://www.rstudio.com/products/rstudio/features/

R Studio


~/Documents/Uni/Projects/qetxrbook - master - RStudio
File Edit Code View Plots Session Build Debug Tools Help
dev.R README.md SUMMARY.md references.bib installr.Rmd Untitled1*
Source on Save
Run
Source
1 a <- 25
2 b <- 3
3 a*b
Editor window
3:4 (Top Level)
R Script
Environment History Git
Import Dataset
List
Global Environment
Values
a 25
b 3
Objects / Variables
Files Plots Packages Help Viewer
read.table
R: Data Input
Find in Topic
read.table {utils}
R Documentation
Data Input
Help / Packages /
Plots / Files
Description
Reads a file in table format and creates a data frame from it, with cases corresponding to lines and variables to fields in the file.
Usage
read.table(file, header = FALSE, sep = "", quote = "\"'",
dec = ".", numerals = c("allow.loss", "warn.loss", "no.loss"
row.names, col.names, as.is = !stringsAsFactors,
na.strings = "NA", colClasses = NA, nrows = -1,
skip = 0, check.names = TRUE, fill = !blank.lines.skip,
strip.white = FALSE, blank.lines.skip = TRUE,
comment.char = "#",
Console ~/Documents/Uni/Projects/qetxrbook/
Working directory
> a <- 25
> b <- 3
> a*b
[1] 75
>
Console window

Basics of the Editor

1. 5 + 5
2. 1:200
3. Print(“HelloWorld!”)
4. ?install.packages
5. x <- 1:5
6. y <- sample(c(1:5))
7. g <- c(“HelloWorld!”)

What is a package?

- Customize your R Environment
  - Only need to install packages once
  - Create your own local R library

- Install.packages("ggplot2")
  - Or
- Help Window – Packages – Install

- Click on a package name to learn more
  - Example – GGplot2

Working in the Editor with Packages

- 1. install.packages("ggplot2")
- 2. library("ggplot2")
  - Packages are in library
  - See current list of installed packages
- 3. search()
  - See current list of loaded packages

Excel to R - Example

month	rain_mm	flow_cms
1	128	15000
2	98	12000
3	92	11000
4	77	9800
5	68	7600
6	63	5800
7	76	5500
8	81	5700
9	84	6200
10	122	9500
11	117	15000
12	125	1700

Importing Data

- Very tricky
  - Excel – save your data as csv file
  - CSV is a universal format that is easy to read
- Work in the finder and create an R folder
  - What is the file path?
  - Copy and paste file.example.csv into a folder
- example.csv <- read.csv("~/Desktop/R/file.example.csv", header =)

Success?

- View(example.csv)
- Your file pathway will be unique to your computer
  - Double check your pathway

Additional FunctionS

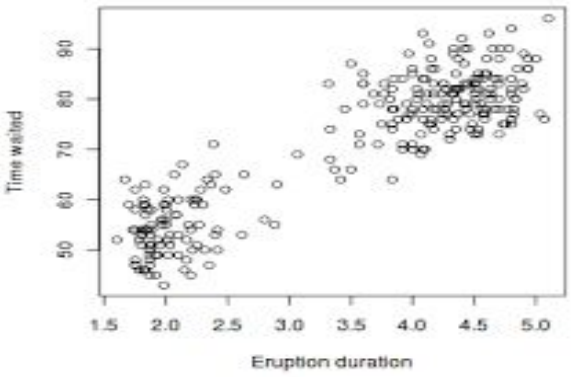
1. Summary(example.csv)
2. Mean(object$variable)
    1. Isolates subset of data
    2. attach()
3. mean(variable)
4. dim()
5. class()
6. as.factor cmd
7. rm()

Practice with R data

1. data()
2. data("mtcars")
3. ? , head, tail
4. lmfit <- lm(mpg ~ wt +qsec, mtcars)
5. summary (lmfit)
6. Install.packages("Broom")
7. broom::tidy(lmfit)

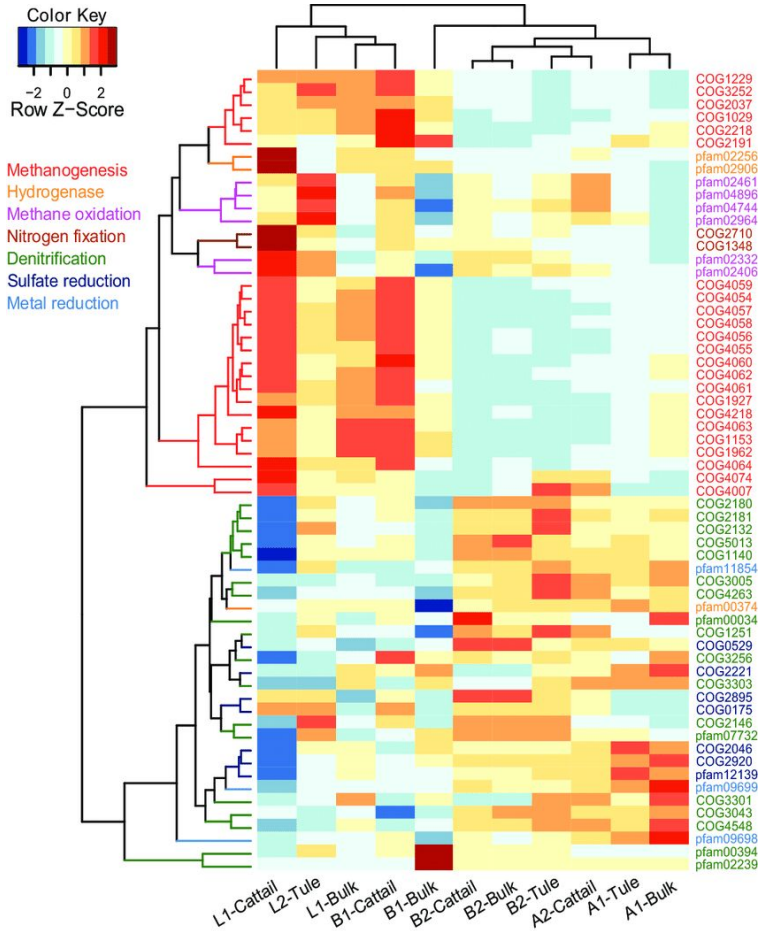
Simple Visualization

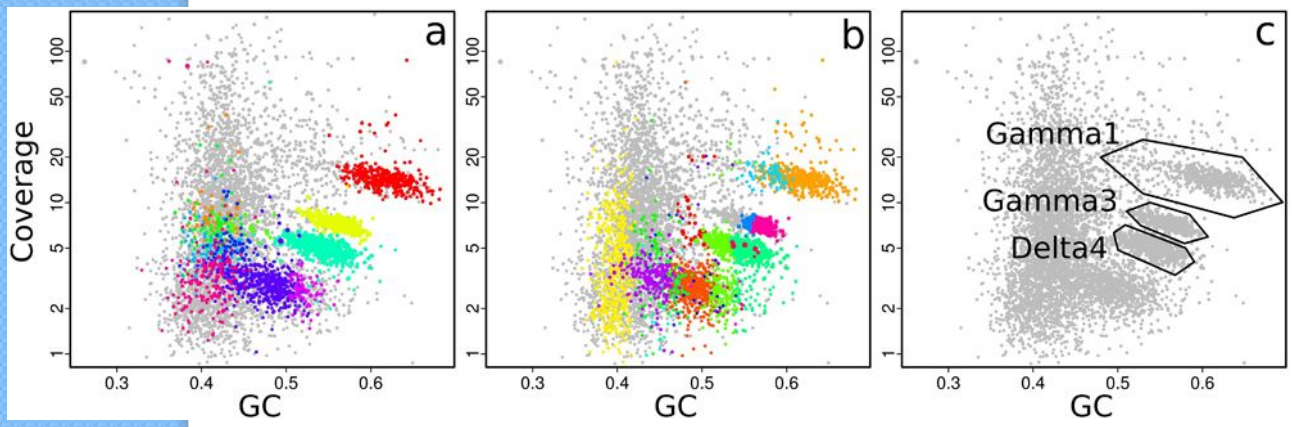
- Td <- broom::tidy(lmfit, conf.int = TRUE)
- Hist(mtcars$mpg)
- ggplot(td, aes(estimate, term, color = term)) + geom_point() + geom_errorbarh(aes(xmin = conf.low, xmax = conf.high))



Conclusion

- Start out small
- Focus on fundamentals
- Utilize resources
- Have fun!
- https://plot.ly/r/



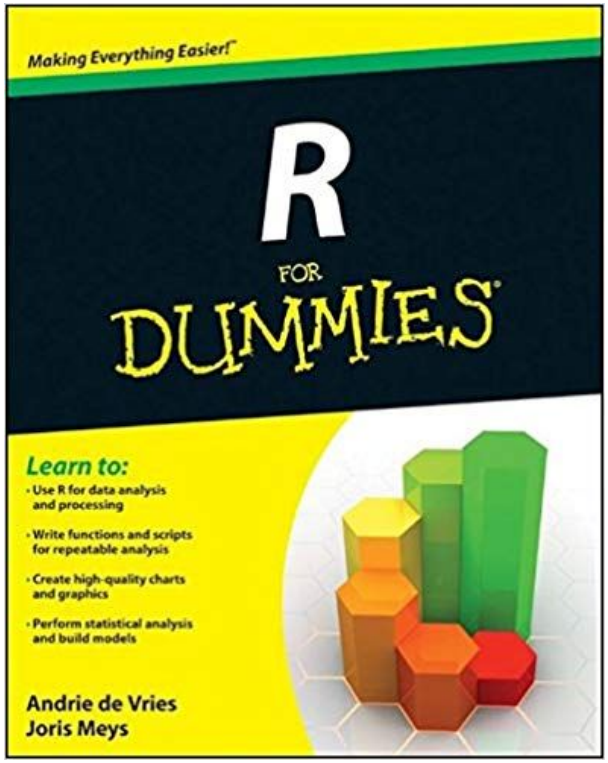


Top: He et al. (2015)

Left: Seah and Vodicka (2015)

Common FAQs

- Is R case-sensitive?
- A command won't work, now what?
- Double check: / not \



Resources

- https://www.r-project.org/doc/bib/R-books.html
- https://leanpub.com/rprogramming
- https://www.r-bloggers.com
- https://simplystatistics.org/
- https://www.coursera.org/learn/r-programming
- https://www.datacamp.com/
- https://swirlstats.com/
- http://adv-r.had.co.nz/