R: FAQs & Beginner Resources by Topic

22.12.2018

Curated and organized by Noushin Nabavi, PhD.

##Getting Set-Up with R & RStudio

- **Download & Install R:**
  - https://cran.r-project.org
  - For Mac: click on **Download R for (Mac) OS X**, look at the top link under **Files**, which at time of writing is **R-3.2.4.pkg**, and download this if compatible with your current version mac OS (Mavericks 10.9 or higher). Otherwise download the version beneath it which is compatible for older mac OS versions. Then install the downloaded software.
  - For Windows: click on **Download R for Windows**, then click on the link **install R for the first time**, and download from the large link at the top of the page which at time of writing is **Download R 3.2.4 for Windows**. Then install the downloaded software.
- **Download & Install RStudio:**
  - https://www.rstudio.com/products/rstudio/download/
  - For Mac: under the **Installers for Supported Platforms** heading click the link with **Mac OS X** in it. Install the downloaded software.
  - For Windows: under the **Installers for Supported Platforms** heading click the link with **Windows Vista** in it. Install the downloaded software.

##Beginner Resources by Topic

- **Exercises in R: swirl (HIGHLY RECOMMENDED):**
  - http://swirlstats.com/students.html
- **Data Prep**:
  - Intro to dplyr: https://cran.rstudio.com/web/packages/dplyr/vignettes/introduction.html
  - Data Manipulation (detailed): http://www.sr.bham.ac.uk/~ajrs/R/index.html
  - Aggregation and Restructing Data (base & reshape): http://www.r-statistics.com/2012/01/aggregation-and-restructuring-data-from-r-in-action/
- **Data Types intro**: Vectors, Matrices, Arrays, Data Frames, Lists, Factors: http://www.statmethods.net/input/datatypes.html
- **Using Dates and Times**: http://www.cyclismo.org/tutorial/R/time.html
- **Text Data and Character Strings**: http://gastonsanchez.com/Handling_and_Processing_Strings_in_R.pdf
- **Data Mining**: http://www.rdatamining.com

- **Data Viz**:
  - ggplot2 Cheat Sheet (RECOMMENDED): http://zevross.com/blog/2014/08/04/beautiful-plotting-in-r-a-ggplot2-cheatsheet-3/

  - ggplot2 theoretical tutorial (detailed but RECOMMENDED): http://www.ling.upenn.edu/~joseff/avml2012/
  - Examples of base R, ggplot2, and rCharts: http://patilv.com/Replication-of-few-graphs-charts-in-base-R-ggplot2-and-rCharts-part-1-base-R/
  - Intro to ggplot2: http://heather.cs.ucdavis.edu/~matloff/GGPlot2/GGPlot2Intro.pdf
- **Interactive Visualisations**:
  - Interactive graphics (rCharts, jQuery): http://www.computerworld.com/article/2473365/business-intelligence/business-intelligence-106897-how-to-turn-csv-data-into-interactive-visualizations-with-r-and-rchart.html

---

- **Statistics**:
  - Detailed Statistics Primer: http://health.adelaide.edu.au/psychology/ccs/docs/lsr/lsr-0.3.pdf
  - Beginner guide to statistical topics in R: http://www.cyclismo.org/tutorial/R/
- **Linear Models**: http://data.princeton.edu/R/gettingStarted.html
- **Time Series Analysis**: https://www.otexts.org/fpp/resources
- **Little Book of R series**:
  - Time Series: http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/
  - Biomedical Statistics: http://a-little-book-of-r-for-biomedical-statistics.readthedocs.org/en/latest/
  - Multivariate Statistics: http://little-book-of-r-for-multivariate-analysis.readthedocs.org/en/latest/

- **RStudio Cheat Sheets**:
  - RStudio IDE: http://www.rstudio.com/wp-content/uploads/2016/01/rstudio-IDE-cheatsheet.pdf
  - Data Wrangling (dplyr & tidyr): https://www.rstudio.com/wp-content/uploads/2015/02/data-wrangling-cheatsheet.pdf
  - Data Viz (ggplot2): https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf
  - Reproducible Reports (markdown): https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf
  - Interactive Web Apps (shiny): https://www.rstudio.com/wp-content/uploads/2015/02/shiny-cheatsheet.pdf

---

Specialist Topics

- **Google Analytics**: http://online-behavior.com/analytics/r
- **Spatial Cheat Sheet**: http://www.maths.lancs.ac.uk/~rowlings/Teaching/UseR2012/cheatsheet.html
- **Translating between R and SQL**: http://www.burns-stat.com/translating-r-sql-basics/
- **Google's R style guide**: https://google.github.io/styleguide/Rguide.xml

---

Operational Basics

- **Working Directory**:
  Example on a mac = `setwd("~/Desktop/R")` or `setwd("/Users/CRT/Desktop/R")`
  Example on windows = `setwd("C:/Desktop/R")`

- **Help**:
  `?functionName`
  `example(functionName)`
  `args(functionName)`

```
help.search("your search term")
```

- **Assignment Operator**: <-

FAQs

Getting Your Data into R

1. Loading Existing Local Data

(a) When already in the working directory where the data is

Import a local **csv** file (i.e. where data is separated by **commas**), saving it as an object:

```
#this will create a data frame called "object"
#the header argument is defaulted to TRUE, i.e. read.csv assumes your file has a header row and will ta
object <- read.csv("xxx.csv")

#if your csv does not have a header row, add header = FALSE to the command
#in this call default column headers will be assigned which can be changed
object <- read.csv("xxx.csv", header = FALSE)
```

Import a local tab delimited file (i.e. where data is separated by **tabs**), saving is as an object:

(b) When NOT in the working directory where the data is

For example to import and save a local **csv** file from a different working directory you can either need to specify the file path (operating system specific), e.g.:

```
#on a mac
object <- read.csv("~/Desktop/R/data.csv")

#on windows
object <- read.csv("C:/Desktop/R/data.csv")
```

OR

You can use the file.choose() command which will interactively open up the file dialog box for you to browse and select the local file, e.g.:

```
object <- read.csv(file.choose())
```

(c) Copying and Pasting Data

For relatively small amounts of data you can do an equivalent copy paste (operating system specific):

```
#on a mac
object <- read.table(pipe("pbpaste"))

#on windows
object <- read.table(file = "clipboard")
```

2. Loading Non-Numerical Data - character strings

Be careful when loading text data! R may assume character strings are statistical factor variables, e.g. “low”, “medium”, “high”, when are just individual labels like names. To specify text data NOT to be converted into factor variables, add `stringsAsFactor = FALSE` to your `read.csv/read.table` command:

```
object <- read.table("xxx.txt", stringsAsFactors = FALSE)
```

3. Downloading Remote Data

For accessing files from the web you can use the same `read.csv/read.table` commands. However, the file being downloaded does need to be in an R-friendly format (maximum of 1 header row, subsequent rows are the equivalent of one data record per row, no extraneous footnotes etc.). Here is an example downloading an online csv file from Pew Research:

```
object <- read.csv("https://vincentarelbundock.github.io/Rdatasets/csv/datasets/AirPassengers.csv")
```

4. Other Formats - Excel, SPSS, SAS etc.

For other file formats, you will need specific R packages to import these data.

Here's a good site for an overview: http://www.statmethods.net/input/importingdata.html

Here's a more detailed site: http://r4stats.com/examples/data-import/

Here's some info on the `foreign` package for loading statistical software file types: http://www.ats.ucla.edu/stat/r/faq/inputdata_R.htm

---

Getting Your Data out of R

1. Exporting data

Navigate to the working directory you want to save the data table into, then run the command (in this case creating a tab delimited file):

```
write.table(object, "xxx.txt", sep = "\t")
```

2. Save down an R object

Navigate to the working directory you want to save the object in then run the command:

```
save(object, file = "xxx.rda")

#reload the object
load("xxx.rda")
```
