HKUST BUSINESS SCHOOL 香港科大商學院

ISOM 3390 Fall 2018
Business Programming in R

R and RStudio Installation and R Markdown

This week is all about getting started with R and getting familiar with RStudio, the integrated development environment (IDE) for R.

Install R and RStudio

First, you will need to install R and RStudio. R is a free, open-source programming language for statistical computing. Almost all of our work in this class will be done using R. You will need regular, reliable access to a computer running an up-to-date version of R. RStudio is a free interactive development environment designed for R. It contains a built-in code editor, many features to make working with R easier, and works the same way across different operating systems. Use of RStudio is required for the labs, and strongly recommended in general.

- **Install R:** Go to https://cran.r-project.org/ and follow the link for your operating system.
- **Install RStudio:** Go to https://www.rstudio.com/products/rstudio/download/ and click on the installer link for your operating system.

See this video for step-by-step installation instructions if needed.

NOTE: If you already have R and RStudio installed, you should still visit these links to confirm that you have the most recent versions of these software. The most recent version of R can be found on The R Project for Statistical Computing page and the most recent version of RStudio can be found on the Download RStudio page. Please install the most recent versions before proceeding.

Before you start using RStudio, one key concept is the **working directory**. This is the directory/folder on your computer where you will store project files, data, and code. It's important that you tell RStudio where the working directory is that you will be using so that it knows where to find the appropriate file (the working directory can be any directory on your computer).

The RStudio Console

The console allows you to work with R commands interactively. It receives command input from programmers and evaluate it immediately. An output may or may not be printed to the screen depending on the types of commands that you run. The RStudio console includes a

variety of features intended to make working with R more productive and straightforward. Please type the following commands and evaluate them sequentially in the RStudio console:

```
x <- c(1, 3, 2, 5)
x
x = c(1, 6, 2)
x
y = c(1,4,3)
length(x)
length(y)
x+y
ls()
rm(x,y)
ls()
rm(list=ls())
?matrix
x<-matrix(data=c(1,2,3,4), nrow=2, ncol=2)
x
x<-matrix( c(1,2,3,4), 2, 2)
matrix(c(1,2,3,4),2,2,byrow=TRUE)
sqrt(x)
x^2
```

R Markdown

Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. R Markdown provides an authoring framework for data science. You can use a single R Markdown file to both

- save and execute code
- generate high quality reports that can be shared with an audience

R Markdown documents are fully reproducible and support dozens of static and dynamic output formats.

Download the plain text file that has the extension .Rmd from Canvas. This is an R Markdown file. Open it in the RStudio IDE.

Notice that the file contains three types of content:

- An (optional) YAML header surrounded by ---s
- R code chunks surrounded by ```s
- text mixed with simple text formatting

When you open the file in the RStudio IDE, it becomes a notebook interface for R. An R Notebook is an R Markdown document with chunks that can be executed independently and interactively, with output visible immediately beneath the input. You can run each code chunk by clicking the ▶ icon. RStudio executes the code and display the results inline with your file.

It allows for direct interaction with R while producing a reproducible document with publication-quality output.

You can quickly insert code chunks like these into your file with

- the Add Chunk  command in the editor toolbar
- or by typing the chunk delimiters ```{r} and ```.

Chunk output can be customized, arguments set in the {} of a chunk header. Above, we use five arguments:

- include = FALSE prevents code and results from appearing in the finished file. R Markdown still runs the code in the chunk, and the results can be used by other chunks.
- eval = FALSE prevents running the code in the code chunk.
- echo = FALSE prevents code, but not the results from appearing in the finished file. This is a useful way to embed figures.
- message = FALSE prevents messages that are generated by code from appearing in the finished file.
- warning = FALSE prevents warnings that are generated by code from appearing in the finished.

See the R Markdown Reference Guide for a complete list of chunk options.

Rendering output

To generate a report from the file, use the “Knit” button in the RStudio IDE to render the file and preview the output with a single click. When you render your .Rmd file, R Markdown will run each code chunk and embed the results beneath the code chunk in your final report. The process generates a new file that contains selected text, code, and results from the .Rmd file. The new file can be a finished web page, PDF, MS Word document, slide show, notebook, handout, book, dashboard, package vignette or other format.