Installation

Dr. Aline Coutinho

What is R? And why should it matter to you?

R is a programming language, increasingly used for statistical analysis and data visualization. There are four main reasons R is a suitable and appealing program:

1. R has been increasingly used in academia, think tanks, newsroom data teams (e.g. click here to see what the BBC data team is doing), and governmental statistical agencies to process quantitative data and visualization. Many industries now require a certain proficiency in R.
2. R is free!
3. The R community is very accomodating, and many of us are trying to make it more inclusive. There are several online resources where you can get your questions answered. It is not uncommon to meet random people who are willing to help you solve practical or technical problems. If you have a Twitter account, for example, you can pose a question with famous hashtags. Chances are someone will provide support. Here are some commonly used hashtags: #rtats, #rladies.

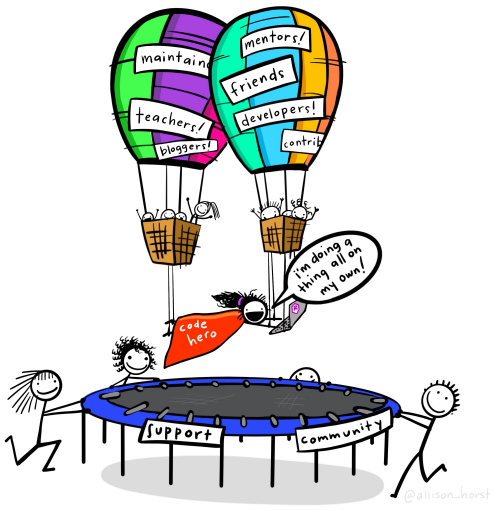

Figure 1: The R community, figure by Allison Horst

4. And last, but not least, R is a powerful tool to analyze data and generate compelling data visualization.

Where can you download R?

First, you have to visit the R project website: click here.

The homepage will ask you to choose a CRAN mirror*. It is advisable that you choose the location closest to you. Because I live in Ottawa, I chose http://cran.utstat.utoronto.ca/. You can find this link on the left side of your window, under Canada.

*CRAN refers to Comprehensive R Archive Network.

After you choose yor mirror, just follow the steps necessary to download any program. If you have problems, please let me know.

But what about RStudio? Do you really need it? Where do you download it?

Once you have installed R in your computer, you will get the usual R GUI (Graphic User Interface). It is absolutely possible to work with R GUI. However, I highly recommend you also download RStudio. Why? RStudio is an open-source IDE, that is, an Integrated Development Environment. It facilitates working with R because it is more user-friendly. RStudio integrates all the tools you use with R into a single environment. It makes easier and faster to navigate to files and functions, start or open existing project files, and creating HTML, PDF, Word documents and slide presentations seamlessly.

You can download RStudio here.

There is an alternative if you don't want to download RStudio. If you have access to the Internet, you can use RStudio through your browser. All you need is to use a cloud-based option called RStudio Cloud: link here. However, you must make an account. If you choose this alternative, all your calculations will be done in a server. I have nerver used it, but I've heard it works fine. It is just a bit slower.

What to expect when you open RStudio?

When you download R and R Studio, and open the software, you will find an interface like this:

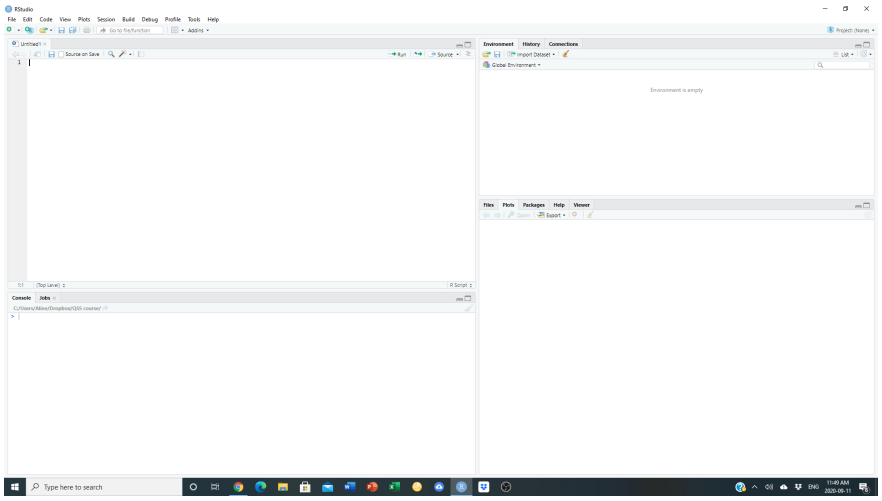

Figure 2: The RStudio interface

1) You will write your codes on the top left window, called *the code window*.
2) The bottom left window is called *the console*, and it is here we see what is happening in R.
3) The top right window is called *the environment*. There, you will find your datasets, the objects you create (we will talk more about it later on), the values of these objects, and so on.
4) Finally, the bottom righ is *the output* of your codes. It is there you will find information about packages, help requests, and the files you have.

You are now all set. Remember, there is a learning curve. Don't give up! You will soon realize you will never stop learning. But the journey pays off handsomely.

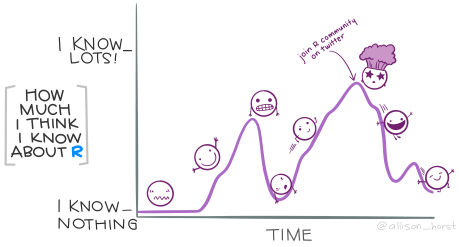

Figure 3: R Knowledge Rollercoaster, by Allison Horst

Additional material

Grolemund, Garrett, and Hadley Wickham. 2017. R for Data Science. First. Sebastopol, CA: O'Reilly Media. https://r4ds.had.co.nz/.

Ismay, Chester, and Patrick C. Kennedy. 2016. Getting Used to R, RStudio, and R Markdown. https://rbasics.netlify.com.