A brief introduction to RStudio and RMarkdown

Kenneth Benoit

30/09/2019

Before we start our first guided coding session, let's become a bit more familiar with RStudio (which we will use to interact with R) and RMarkdown.

Why R?

- Currently used by most statisticians and social scientists interested in data analysis; and also becoming one of most popular languages in Data Science.
- Open source: makes it highly customizable and easily extensible through "packages" (over 10,000 and counting!).
- Powerful tool to conduct automated text analysis, social network analysis, and data visualization, with packages such as quanteda, igraph or ggplot2.
- Command-line interface and scripts favors reproducibility.
- Excellent documentation and online help resources.

RStudio

We will be using RStudio to interact with R, and write our annotated R code using Markdown.

RStudio is an open-source integrated development environment (IDE) that enhances the capabilities of the R programming language. The main advantage of RStudio with respect to other graphical interfaces, such as R GUI (the default), is that it integrates a powerful built-in text editor as well as other tools for plotting, debugging, and workspace management. ## RMarkdown

Markdown is a simple formatting syntax to generate HTML or PDF documents. In combination with R, it will generate a document that includes the comments, the R code, and the output of running such code.

You can embed R code in chunks like this one:

```
1 + 1
```

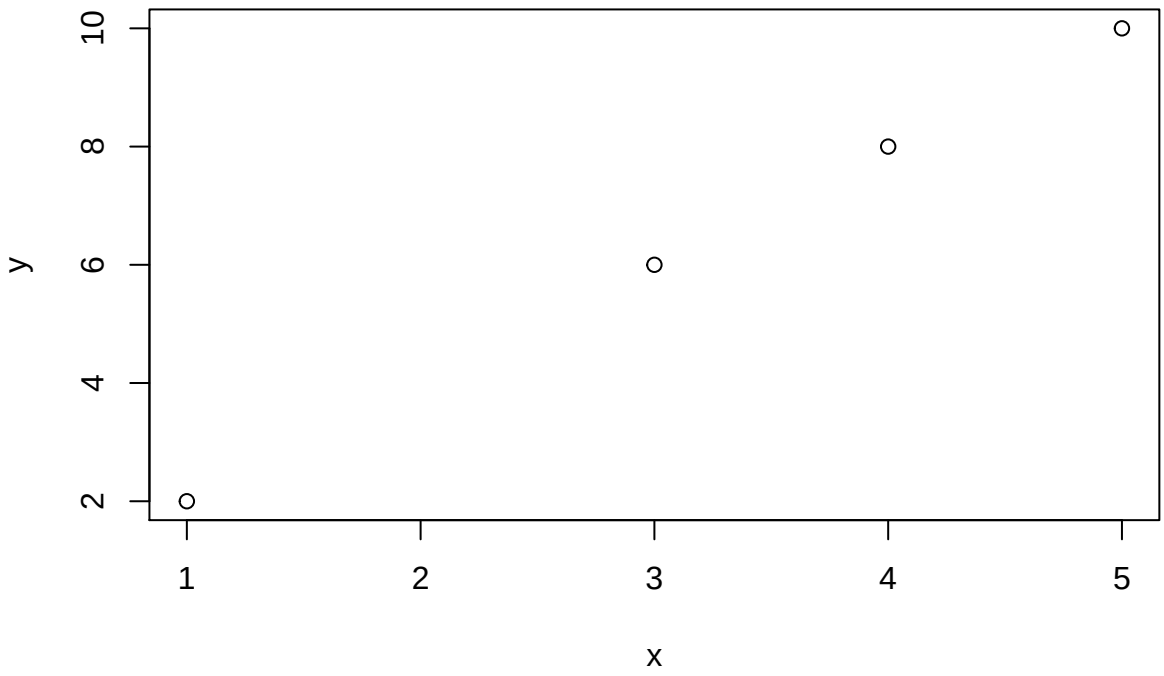
```
## [1] 2
```

You can run each chunk of code one by one, by highlighting the code and clicking `Run` (or pressing `Ctrl + Enter` in Windows or `command + enter` in OS X). You can see the output of the code in the console right below, inside the RStudio window.

Alternatively, you can generate (or **knit**) an html document with all the code, comment, and output in the entire `.Rmd` file by clicking on `Knit HTML`.

You can also embed plots and graphics, for example:

```
x <- c(1, 3, 4, 5)
y <- c(2, 6, 8, 10)
plot(x, y)
```



If you run the chunk of code, the plot will be generated on the panel on the bottom right corner. If instead you knit the entire file, the plot will appear after you view the html document.

Using R + Markdown has several advantages: it leaves an "audit trail" of your work, including documentation explaining the steps you made. This is helpful to not only keep your own progress organized, but also make your work reproducible and more transparent. You can easily correct errors (just fix them and run the script again), and after you have finished, you can generate a PDF or HTML version of your work.

RMarkdown is compatible with all other types of Markdown syntax:

- For example, you can have bullet points
- Or you can also add hyperlinks
- As well as other types of such as **bold** or *italics*

We will be exploring R through R Markdown in the lab session. For more details and documentation see http://rmarkdown.rstudio.com.