

R Markdown

Lun-Ching Chang

July 14, 2020

All rights reserved. This note or any portion thereof may not be reproduced or used in any manner whatsoever without the express written permission of the instructor except for the use of brief quotations in a book review.

1. Introduction

R Markdown provides a unified authoring framework for data science, combining your code, its results, and your prose commentary. R Markdown documents are fully reproducible and support dozens of output formats, like PDFs, Word files, slideshows, and more. R Markdown files are designed to be used in three ways:

- For communicating to decision makers, who want to focus on the conclusions, not the code behind the analysis.
- For collaborating with other data scientists, who are interested in both your conclusions, and how you reached them (i.e., the code).
- As an environment in which to *do* data science, as a modern day lab notebook where you can capture not only what you did, but also what you were thinking.

R Markdown integrates a number of R packages and external tools. This means that help is, by and large, not available through `?`. Instead, as you work through this chapter, and use R Markdown in the future, keep these resources close to hand:

- R Markdown Cheat Sheet: available in the RStudio IDE under *Help -> Cheatsheets -> R Markdown Cheat Sheet*
- R Markdown Reference Guide: available in the RStudio IDE under *Help -> Cheatsheets -> R Markdown Reference Guide*

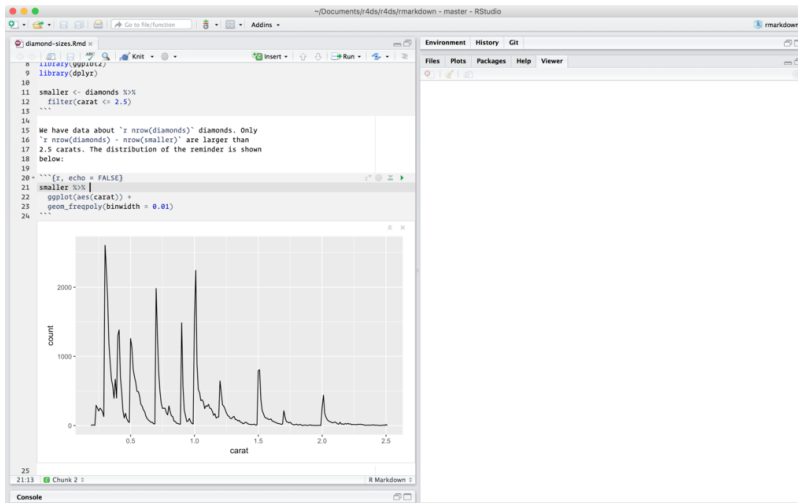
1.1. Prerequisites

You need the **rmarkdown** package, but you don't need to explicitly install it or load it, as RStudio automatically does both when needed.

2. R Markdown Basics

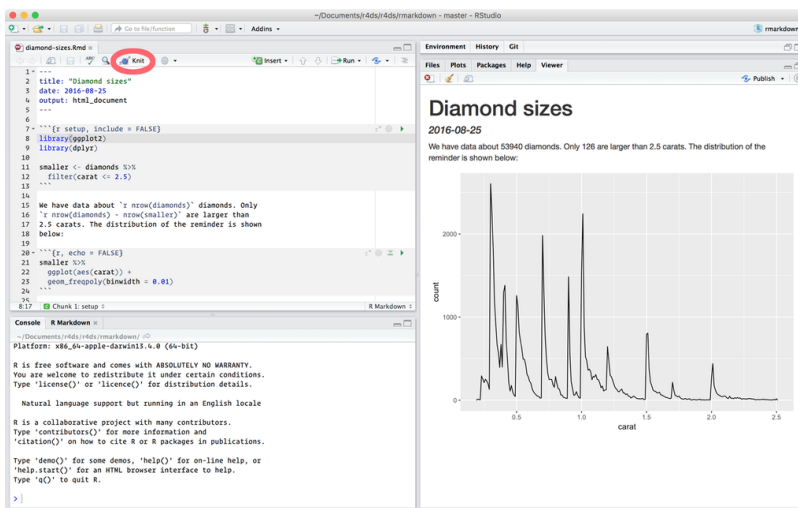
Please download the file: **RMarkdown Basic.Rmd**. This is an R Markdown file, a plain-text file that has the extension *.Rmd*. It contains three important types of content:

1. An (optional) *YAML header* surrounded by `---s`.
2. *Chunks* of R code surrounded by three ```.
3. Text mixed with simple text formatting like `# heading` and `_italics_`.



When you open an *.Rmd*, you get a notebook interface where code and output are interleaved. You can run each code chunk by clicking the Run icon (it looks like a play button at the top of the chunk), or by pressing Cmd/Ctrl-Shift-Enter. RStudio executes the code and displays the results inline with the code:

To produce a complete report containing all text, code, and results, click “Knit” or press Cmd/Ctrl-Shift-K. You can also do this programmatically with `rmarkdown::render("1-example.Rmd")`. This will display the report in the viewer pane, and create a self-contained HTML file that you can share with others.



When you *knit* the document R Markdown sends the *.Rmd* file to **knitr**, which executes all of the code chunks and creates a new Markdown (.md) document that includes the code and its output. The Markdown file generated by **knitr** is then processed by **pandoc**, which is responsible for creating the finished file. The advantage of this two-step workflow is that you can create a very wide range of output formats.



There are many types of output you can produce with R Markdown. There are two ways to set the output of a document:

1. Permanently, by modifying the the YAML header:

```

title: "Viridis Demo"
output: html_document

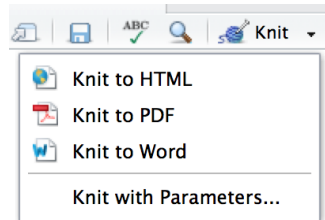
```

2. Transiently, by calling `rmarkdown::render()` by hand:

```
rmarkdown::render(  
  "diamond-sizes.Rmd",  
  output_format = "word_document"  
)
```

This is useful if you want to programmatically produce multiple types of output.

The easiest way to choose the output format is RStudio's knit button renders a file to the first format listed in its **output** field. You can render to additional formats by clicking the drop-down menu beside the knit button.



To get started with your own *.Rmd* file, select *File -> New File -> R Markdown.* in the menu bar. RStudio will launch a wizard that you can use to pre-populate your file with useful content that reminds you how the key features of R Markdown work.

R Markdown is still relatively young, and is still growing rapidly. The best place to stay on top of innovations is the official R Markdown website: <http://rmarkdown.rstudio.com>.

For all the homeworks through out this semester, please use R markdown to generate your homework (Uploading file on Canvas will only accept “.Rmd” file).

References

Wickham, Hadley, and Garrett Golemund. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. 1st ed. O'Reilly Media, Inc.