Refining your RMarkdown file

Data Wrangling and Husbandry

04/06/2020

RMarkdown is quite flexible, but I find it a little unclear just what it can and cannot do. If you need precise formating, and know LaTeX, you might prefer to create `*.Rnw` or `*.Rtex` files instead of `*.Rmd` files. There is a *very* brief introduction at http://kbroman.org/knitr_knutshell/pages/latex.html and some examples at http://yihui.name/knitr/demo/minimal/

However, there are lots of refinements available in RMarkdown.

Math in RMarkdown

- Many LaTeX expressions will work in RMarkdown. Put `$ $` around inline math and `$$ $$` around displayed math.
- http://www.calvin.edu/~rpruim/courses/m343/F12/RStudio/LatexExamples.html has examples, although you no longer need the `latex` prefix. See also this page and this set of examples
- For example, `$x^2$` gives x^2,
- This is a linear regression model: $Y = \beta X + \epsilon$
- This is an integral

$$\int_a^b f(x)dx$$

- This is a linear regression model:

$$Y = \beta X + \epsilon$$

- `$$\lim_{x_+\to\infty} e^{-x^2} = 0$$` gives

$$\lim_{x \to \infty} e^{-x^2} = 0$$

Remember inline R

For a reproducible document, it's necessary to refer to results via code rather than explicitly. For example, if you've found that the three most frequently purchased widgets are `c("Alpha", "Beta", "Gamma")`, then rather than writing "The three most purchased widgets are Alpha, Beta, and Gamma", write "The three most purchased widgets are *backtick* r c("Alpha","Beta","Gamma") *backtick*" (but use an actual back tick) which will give

The three most purchased widgets are Alpha, Beta, Gamma.

This approach is only worthwhile if instead of `c("Alpha", "Beta", "Gamma")` you refer to some R object

Use `kable()` and `xtable()` to format tables

You can use `kable()`, in the `knitr` package, and the more complex `xtable()`, in the `xtable` package. I believe both require `results='asis'` in the chunk options.

```
select(mtcars, mpg, cyl, hp)[1:5,]
```

```
##                    mpg cyl  hp
## Mazda RX4         21.0   6 110
## Mazda RX4 Wag     21.0   6 110
## Datsun 710        22.8   4  93
## Hornet 4 Drive    21.4   6 110
## Hornet Sportabout 18.7   8 175
```

```r
library(knitr)
kable(select(mtcars, mpg, cyl, hp)[1:5,])
```

	mpg	cyl	hp
Mazda RX4	21.0	6	110
Mazda RX4 Wag	21.0	6	110
Datsun 710	22.8	4	93
Hornet 4 Drive	21.4	6	110
Hornet Sportabout	18.7	8	175

```
library(xtable)
print(xtable(select(mtcars, mpg, cyl, hp)[1:5,]), comment =
```

mpg

cyl

hp

Mazda RX4

21.00

6.00

110.00

Mazda RX4 Wag

21.00

6.00

110.00

Datsun 710

YAML

The `Knit` button in RStudio actually performs several tasks, which can be thought of as:

- knit: Rmarkdown to markdown
- pandoc: markdown to html

The intermediate file might be different (say a LaTeX file), and pandoc (or LaTeX) might produce a Word file, a pdf, or many other possibilities.

The opening lines of a RMarkdown file, as inserted by RStudio, are known as YAML (YAML Ain't Markup Language).

```
---
title: "Refining your RMarkdown file"
author: "Jason Klusowski"
date: "04/06/2020"
output: ioslides_presentation
---
```

they are actually typically giving directions to pandoc

RStudio has a page, http://rmarkdown.rstudio.com/formats.html, with lots of examples of options. The R Markdown Reference Guide has a summary. Here's a recent example I've used

```
---
title: 'Executive Summary of Monthly Progress Reports: Marc
date: "Reports submitted 04/03/2018. Summary prepared by St
header-includes:
- \usepackage{array}
- \usepackage{ragged2e}
- \usepackage{fancyhdr}
- \usepackage{booktabs}
- \usepackage{float}
- \pagestyle{fancy}
- \fancyhead[CO,CE]{March 2018 Data}
- \fancyhead[LO,LE]{\thepage}
output:
  pdf_document:
    keep_tex: yes
classoption: landscape
```

Chunk options

The R Markdown Reference Guide also has a summary of chunk options

I find the most useful to be

- `eval = FALSE` which will prevent the code chunk from running
- `include = FALSE` which will run the chunk but not include it
- `echo = FALSE` which will not show the code in the chunk
- `message = FALSE` which suppresses messages
- `warning = FALSE` which suppresses warnings
- `cache = TRUE` which will cache the results

You can set options throughout the document with the `knitr::opts_chunk$set()` command

Tip on slides

If you are using a slide format,

---

or

***

will start a new slide (as will a new header starting with ##)

In class exercise

Start with the example RStudio provides when you select File > New File > R Markdown. Experiment with adding math in text, with different chunk options, and with different pdf options. Can you simultaneously produce html and pdf files?