Using Rmarkdown and GitHub with RStudio

Part One: Using Rmarkdown

Benoit Liquet and Nicholas Tierney

2018/06/25

Using Rmarkdown and GitHub with RStudio
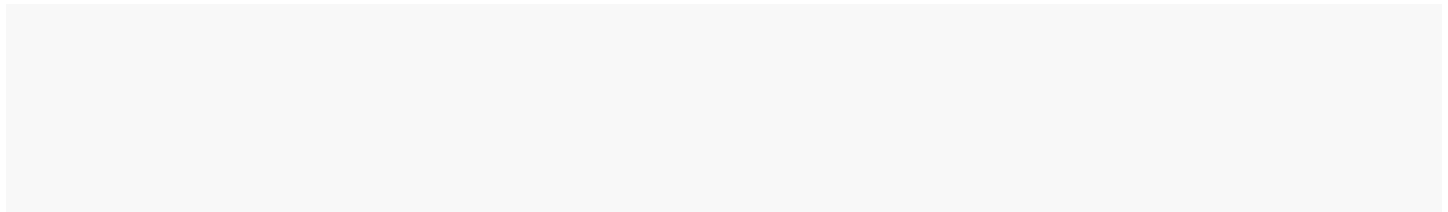
Part One: Using Rmarkdown

Benoit Liquet and Nicholas Tierney

2018/06/25

Please go to this link and click "download"

https://github.com/njtierney/angletr-rmd-gh

You will also need to have these packages installed:

Also, please let me know if you have any questions or concerns before the workshop starts!

Who is Nick?

- PhD in Statistics under Kerrie Mengersen with BRAG
- Currently a research Fellow in Statistics at Monash University with Di Cook and Rob Hyndman at NUMBAT
- Active in the rOpenSci community
- Very Excited about the R language
- Also enjoy:
  - Rock Climbing
  - Drinking Coffee

Who am I ?

- Professor at UPPA.
- Affiliated with ACEMS (Centre of Excellence for Mathematical and Statistical Frontiers), Queensland University of Technology.
- Active member of the LMAP (Laboratory of Mathematics and Its Applications - UMR CNR 5142).
- Also enjoys:
  - Cycling!

Your Turn

Introduce yourselves to TWO other people

Key Elements of the course

Rmarkdown

Github

RStudio

The bad news

There are times when learning this material might be frustrating, and that is a totally normal thing, and expected

The good news:

This is temporary and well worth the effort.

We are here to help!

Why RMarkdown + GitHub?

https://www.youtube.com/watch?v=s3JldKoA0zw

Rstudio

Is amazing and free

An RMarkdown document

An example document

I've skipped ahead - where are we, what is this place?

Let's take it back a notch, and work through these pieces step by step.

How is rmarkdown designed to be used?

As Hadley says in his excellent (and free!) book: R for data science, Rmarkdown is designed to be used:

1. Communicate with decision makers, who want to focus on the conclusions, not the code behind the analysis.

2. An environment in which to   data science, to capture the rich narrative around your code as a modern day lab notebook, where you can capture not only what you did, but also what you were thinking.

3. For collaborating with other data scientists (including future you!), who are interested in both your conclusions, and how you reached them (i.e. the code).

What does rmarkdown look like in RStudio?

- Running chunks
- Creating chunks
- Running a document
- Markdown syntax
- Cheathsheets for rmarkdown:
    - RStudio cheatsheets
    - Rmarkdown cheatsheet

Your Turn:

- Go to

Do the following tasks:

- Add an introduction heading
- Add your name and date to the YAML header
- Describe the data from the airquality dataset using blockquotes (hint:)
- Provide:
  - a bulleted list describing the variables used in the data
  - A direct quote from the ?data set describing the data
  - Make a code chunk
  - Make a library call in the code chunk
  - Make a plot of ozone and solar radiation

Extra:
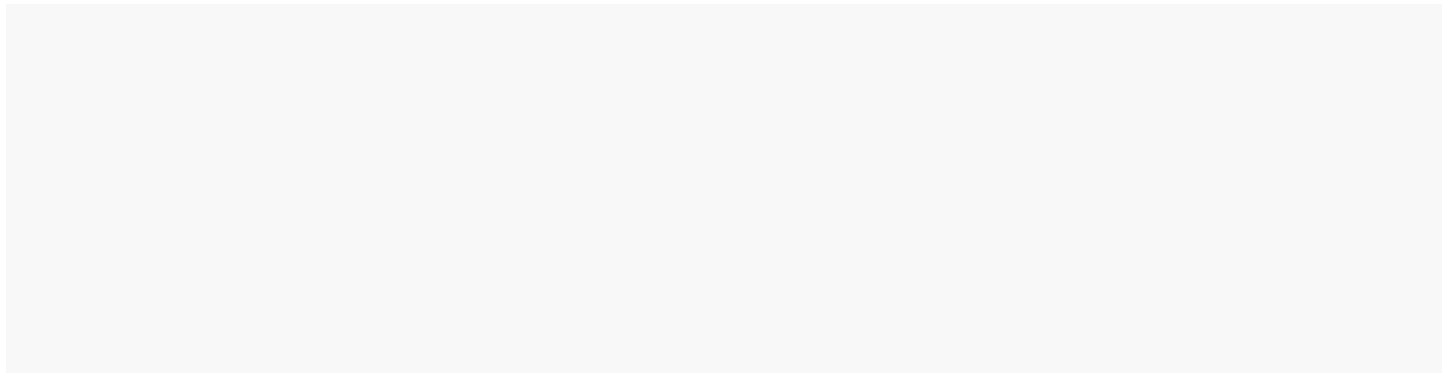
- Make a plot of ozone and temperature

Demonstrate

- Here's one I prepared earlier

What if I want to change the figure?

- Code chunks have options

`

What if I want to keep the images?

Your Turn

Tasks

- Go to

- Change the fig.width, fig.height, and dev of the figure in plot-1

- Add an option to keep the markdown source code (hint: look at the gear icon next to   )
- print the code for one of the figures, but not the other (hint: use   )

Demonstrate

- There are many other chunk options, to learn about them, I recommend:
  - The RStudio cheatsheet
  - The official documentation
  - R4ds section on rmarkdown

Tables, inline code, Equations

- You can print tables in rmarkdown, and refer to values inside text
- You can also print equations using LaTeX syntax wrapped in .

Your Turn

- Go to

    - Make a table
    - Make some inline code chunks
    - Make an equation describing the linear model used

Demonstration

Aside: There are many many options for tables

- knitr::kable is simple
- pander
- xtable
- the list goes on!
- https://github.com/ropenscilabs/packagemetrics
- https://ropensci.org/blog/blog/2017/06/27/packagemetrics

Setting global options

- What if you have 100 chunks that you want to set the figure for?
- Setting all of these options can be a bit annoying - but there is a better way

Your turn

- Go to

- set a global option to:

    - hide all of the code
    - set the graphics device
    - set the DPI

Using citations in a document

- bibliographies and citations are super important in
- You can cite work by putting it in a .bib file in the same directory as your rmarkdown document
- [@Tag]
- This is then printed at the end of the document, so you normally # References at the end

Your turn

- go to

    - refer to the packages that you use and add a citation
    - http://rmarkdown.rstudio.com/authoring_bibliographies

How to troubleshoot an Rmarkdown document

Sometimes you have an error in your rmarkdown document. These can be difficult to track down sometimes, so here is our recommended workflow:

- develop code in chunks and execute the chunks until they work, then move on
- knit the document regularly to check for errors
- if there is an error:
  - recreate the error in an interactive session:
  - restart R
  - run all chunks below
  - find the chunk that did not work, fix until it does
  - run all chunks below
  - explore working directory issues
    - remember that the rmarkdown directory is where the .Rmd file lives

Your turn

Open

Get it to run

My philosophy on Rmarkdown

- Keep it simple
- focus on content
- Then move it to word or LaTeX if you need to

Demonstration: Going further with Rmarkdown

Using rmarkdown and word

Using rmarkdown and LaTeX

Simple, when it works

Extending word templates:
http://rmarkdown.rstudio.com/articles_docx.html

Templates

- Templates:
  - https://github.com/Pakillo/rmdTemplates
- rticles package:
  - https://github.com/rstudio/rticles

Slides

- These very slides are made in rmarkdown!
- RStudio makes it easy
- Highly recommend Xaringan

Notebooks

- http://rmarkdown.rstudio.com/r_notebooks.html

Other formats for rmarkdown

- bookdown: https://bookdown.org/yihui/bookdown/
  - or your thesis!:
    https://github.com/cpsievert/phd-thesis
- blogdown: https://github.com/rstudio/blogdown
- Read about them all here:
  http://rmarkdown.rstudio.com/formats.html

Advanced topics

- citation styles
  - add specific citation styles from
  - http://rmarkdown.rstudio.com/authoring_bibliographies
  - example in
- parameters
  - programmatically control the inputs for your rmarkdown document
  - http://rmarkdown.rstudio.com/developer_parameterized
- flexdashboard
  - http://rmarkdown.rstudio.com/flexdashboard/index.html
- interactive components
  - add a shiny application into your documents
  - http://rmarkdown.rstudio.com/authoring_shiny.html
- How to go from an RScript to Rmd
  - using spin: http://deanattali.com/2015/03/24/knitrs-best-hidden-gem-spin/

Learning more

- The book by Yihui! https://yihui.name/knitr/
- Highly recommend the rmarkdown chapter in R for data science:
    - http://r4ds.had.co.nz/r-markdown.html
- knitr in a nutshell: http://kbroman.org/knitr_knutshell/
- Maëlle Salmons course: https://github.com/maelle/rmd_course_isglobal

References

This tutorial was heavily inspired by the excellent teaching material that is made freely available on the internet:

- R for Data Science
- happy git with r
- Maëlle Salmon's Rmarkdown course: https://github.com/maelle/rmd_course_isglobal

Plan

- Break
- GitHub session next