Advanced Data Analysis in R

Advanced `rmarkdown`

Michael DeWitt

2018-02-09 (Updated 2019-01-28)

Why is `rmarkdown` important?

It's all about communication and documentation!

1. We used have notebooks to document our work:

2. Our reports change change with our data

Reproducibility

- Tying our analysis to our output documentation

- No more of this https://youtu.be/s3JldKoA0zw

Sold? So how do I do it?

- Rmarkdown integrates R and the Markdown language into a single method

- Rmarkdown documents end in `.Rmd` extension

- Can be created from within the R Studio Integrated Development Environment (IDE)

Some technical details. . .

- rmarkdown (and bookdown) compile their outputs to pandoc
- Depending on the output specified
    - $\LaTeX$for pdf style outputs (and beamer)
    - html
    - html + javascript
    - epub
    - . . .
- Specific commands can be issued depending on the output used ($\LaTeX$and/or html tags)

Building Your Documents

Three Components to an Rmarkdown Documents

1. yml header than contains metadata and build instructions

2. Markdown mark-up conventions

3. Code chunks with language and output instructions

_Y_et Another _M_arkup _L_anguage

Parts of An R Markdown Documents

yml header instructs to pandoc engine how to build the documents

```
---
title:
subtitle:
author:
abstract:
date:
output:
---
```

You can access R code from within the yml

Utilising backticks and the letter "r" you can include R code into your 'yml'

```
---
title: "This is a quick example"
subtitle: "Just to illustrate a point"
author: "William Gosset"
date: "2018-01-23 Updated(2019-01-28)"
abstract: "Just a little exloration of things. We looked at `nro
output:
  pdf_document
---
```

A Little More About Outputs...

Rmarkdown Outputs
- **pdf_document**
- **word_document**
- **html_document**

Presentations
- **io_presentation**
- **beamer_presentation**

My Prefered
- **bookdown::pdf_document2**[1]

---

[1]Almost identical to pdf_output but provides additional control over code chuck references

Markdown

Rmarkdown = R + Markdown

- Markdown developed as an easy way to implment `html` style formatting

- Keyboard symbols to genrate basic `html` outputs

Markdown Examples

italics or *italics*

__bold__ or **bold**

sub~script~

super^script^

⤳

Markdown Examples

italics or *italics*

bold or **bold**

sub$_{script}$

super^{script}

Markdown Examples

Code Chunk

Writing R Code

Use Chunk Options to Control the Outputs

Accessing *Other* Languages

```r
names(knitr::knit_engines$get())
## [1] "awk"       "bash"      "coffee"    "gawk"
## [5] "groovy"    "haskell"   "lein"      "mysql"
## [9] "node"      "octave"    "perl"      "psql"
## [13] "Rscript"  "ruby"      "sas"       "scala"
## [17] "sed"      "sh"        "stata"     "zsh"
## [21] "highlight" "Rcpp"     "tikz"      "dot"
## [25] "c"        "fortran"   "fortran95" "asy"
## [29] "cat"      "asis"      "stan"      "block"
## [33] "block2"   "js"        "css"       "sql"
## [37] "go"       "python"    "julia"
```

Setting Engine

```
knitr::opts_chunk$set(engine.path = list(
  python = '~/anaconda/bin/python',
  ruby = '/usr/local/bin/ruby'
))
```

A Motivating Example

```
parameters {
  real y[2];
}
model {
  y[1] ~ normal(0, 1);
  y[2] ~ double_exponential(0, 2);
}

library(rstan)
fit <- sampling(stan_example, cores = 2, iter = 50, re
print(fit)
## Inference for Stan model: 6c400a2ac89dae0e85da9f76
## 4 chains, each with iter=50; warmup=25; thin=1;
## post-warmup draws per chain=25, total post-warmup
##
## mean se_mean sd 2.5% 25% 50% 75% 97
```

Extending RMarkdown

Table/ Figure Generation

- kableExtra
- gt

Templates

- rticles
- papaja
- markdowntemplates

Even More

- flexdashboard
- shinydashboard

Other Resources

Rmarkdown
R Studio Cheat Sheets for the basic commands Introduction to Rmarkdown for the basic ideas and getting started R Markdown Definitive Guide for the details for how Rmarkdown works

Bookdown
For Writing Books

Blogdown
For writing blogs/ websites

Test Slide

Stuff   Other Stuff