

# R markdown

Jonathan de Bruin & Tessa Pronk



#### What is R Markdown?

- R Markdown is a package to turn your analyses into high quality documents, reports, presentations and dashboards.
- Developed for Reproducible Research
- R Markdown documents are fully reproducible.
- Simple syntax and easy to learn
- Intro video: https://rmarkdown.rstudio.com/lesson-1.html

R Data Cafe March 11,2018



### **Examples**

auto (managers)
managers - na.omit(managers)
managers - na.omit(managers)
managers - melt(
data.frame( index( managers ), coredata(cumprod( managers+1 )\*100 ) ),
id.vars = 1 //
colnames(managers.melt) <- c("date", "manager","val")
managers.melt[,"date"] <- format(managers.melt[,"date"],format = "%Y-%n-%d") Draw The Graph require(rChorts)
p2 <- C'hortsSnew()
p2SsetLib('libraries/widgets/nyt\_home')
p2SsetLib('libraries/widgets/nyt\_home/layouts/nyt\_home.html")
p2SsetTemplate(script = "libraries/widgets/nyt\_home/layouts/nyt\_home.html") A Pandoc Markdown Article Starter and Template\* description - "This data comes from the managers dataset included in the R package PerformanceAnalytics.", This document provides an introduction to R Markdown, argues for its benefits, and presents Steven V. Miller Clemson University Inst document provides an introduction to K Markdown, argues for its benefits, and presents a sample manuscript template intended for an academic audience. I include basic syntax to R a sample manuscript template intended for an academic audience. I include basic syntax to K

Markdown and a minimal working example of how the analysis itself can be conducted within R 96 percent more If you bought HAM1 it would be worth Price changes from Sept 2002 to Dec 2006 with the knitr package. Keywords: pandoc, r markdown, knitr Academic workflow, certainly in political science, is at a crossroads. The American Journal of Po-HAM1 Academic workflow, certainly in political science, is at a crossroads. The American Journal of Volume 18 Science (AJPS) announced a (my words) "show your work" initiative in which authors who 96 percent more intent Science (AIPS) announced a (my words) "show your work" initiative in which authors who are tentatively accepted for publication at the journal must hand over the raw code and data that produced the rescale absume is the programmer. are tentatively accepted for publication at the journal must hand over the raw code and data that produced the results shown in the manuscript. The editorial team at AIPS then reproduces the code from the manuscript. Pending successful replication, the manuscript moves toward publication. The AIPS might be at the fore of this movement. and it could be the most approach its above. code from the manuscript. Pending successful replication, the manuscript moves toward pub-lication. The APS might be at the fore of this movement, and it could be the most aggressive lication. The AJP's might be at the fore of this movement, and it could be the most aggressive among political science journals, but other journals in our field have signed the joint Data Access. among political science journals, but other journals in our field have signed the joint Data Access
& Research Transparency (DART) initiative. This, at a bare minimum, requires uploading code
to the control of the property of the interest of the property & Research Transparency (DART) initiative. This, at a bare minimum, requires uploading code from quantitatively-oriented published articles to in-house directories hosted by the journal or to Recent Downloads Trices like Dataverse.

There are workflow implications to the Lacour controversy as well. Political science, for the There are workinow implications to the Lacour controversy as well. Political science, for to foreseeable future, will struggle with the extent of the data fraud perpetrated by Michael Lacour controversy as well. Political science, for the forest property of the property of the control of th CRAN Download Monitor foreseeable future, will struggle with the extent of the data fraud perpetrased by a structure of the data fraud perpetrased by a structure of the data fraud perpetrased by the the data fraud services like Dataverse. in an article co-authored with Donald P. Green in Science, the general scientific in the United States. A failure to reproduce LaCour's results with different same in the United States. A failure to reproduce LaCour's results with different same in the United States. on the United States. A tailure to reproduce LaCour's results with different same perpendent size of the Court of "fake" data that provided results to what we be true for a provided results to what we have the court of the cou Downloads per sec (last 5 min) be true (i.e. "truthiness"). However, take data can have real consequences for be and those who want to learn from it and use it for various purposes. Even resear that the second of the consequences are not dilicent in their workflow. comprehensive either by LaCour to "take" data that provided results to what we be true (i.e. "truthiness"). However, take data can have real consequences for possible true (i.e. "truthiness"). ay suffer the same fate if researchers are not diligent in their workflow.

These recent events underscore the DART push and cast a shadow over our. The streaming data is a 1-week.

These recent events underscore the DART push and cast a shadow over our. The streaming data is a 1-week. These recent events underscore the DART push and cast a shadow over our ever, good workflow has always been an issue in our discipline. Cloud storage so delayed livestream of download are still relatively new among political scientists. Without cloud storage, practice of the proper the pressibility that work between a home computer and an office complex from cran, studio, com. The Average downloads per user box are still relatively new among political scientists. Without cloud storage, pr. delayed livestream of gowinosa left open the possibility that work between a home computer and an office complete the possibility that work between a home computer and an office complete the possibility that work between a home computer and an office complete the possibility that work between a home computer and an office complete the possibility that work between a home computer and an office complete the possibility of the possibility o Downloads per sec (last 5 min) left open the possibility that work between a home computer and an office com<sub>1</sub> logs from cran.rstuno.com, rise a function of a corrupted thumb drive, an overheated power supply, or, among, server code for that service is at wave of viruses that would particularly affect Microsoft users every summer. So, an appropriate the service is at the control of the control a function of a corrupted thumb drive, an overheated power supply, or, among , server code to H, wave of viruses that would particularly affect Microsoft users every summer. So: jchengs/cransim. like engineering, have traditionally relied on software like Microsoft Word for many many word processor reduces workflow to a series of clicks and st. Warn when rate exceeds the many word processor reduces workflow to a series of clicks and st. Warn when rate exceeds the many word processor reduces workflow to a series of clicks and st. Warn when rate exceeds the many word processor reduces workflow to a series of clicks and st. 1.6 Percent of downloads (last 5 min) wave of viruses that would particularly affect Microsoft users every summer, soy, like engineering, have traditionally relied on software like Microsoft Word for militarity through our would proceed the software like Microsoft Word for militarity through our would proceed the software like Microsoft Word for militarity through our would proceed to the software like Microsoft Word for militarity and the software like Microsoft Word ration though any word processor reduces workflow to a series of circks and sr, board. This is a terrible way to track changes or maintain version control. The advantage of the series board. This is a terrible way to track changes or maintain version control. The adorators only compounds all the aforementioned issues. The proverbial left hand package \*Replication files are available on the author's Github account (http://github.com/symiller).
February 10, 2016; Corresponding author: symille@clemson.edu. Popularity by package (last 5 min) 1 ggplot2 1.2 Performance Mile predition in the property of the performance of the p 2 matrixStats ploti tidye jablacales evaluate purrr uttis assertthing! 1.2 3 stringr purrr utf8 ssertthtringi shapaTools 4 digest sp :epadwomatgestringr grable 1.0 broorttextaeritteplotærixSt-dplyr 5 magrittr IZE, MODEL, TRANSFORM, TIDY, AND IMPORT DATA 1.0 6 RColorBrewer checkministorBjridieblorsparymbethEigen 1.0 7 viridisLite 0.9 cli mlTaHeliBzyennormaticeExtra 8 colorspace к Data Cafe crayon knitr

Q 🔥 🔘 📭 🕐 🕁 🛭 😉 🖫 🗳 🖪 😛 🕄

Garrett Grolemund & Hadley Wickh

← → C ☆ ① timelyportfolio.github.io/rCharts\_nyt\_home\_price/

supplied by the PerformanceAnalytics package. Get Data and Transform

Reusable Version in rCharts

#get the data and convert to a format that we would expect from melted xts #also original only uses a single value (val) and not other require(reshape2) require(PerformanceAnalytics)

As I mentioned above, this visualization works well with any cumulative growth time series, so let's apply it to the managers dataset



### R and Markdown: What is Markdown? (1)

- Markdown is a lightweight markup language with plain text formatting syntax.
- R and Markdown is an extended version of markdown
- Google "Markdown cheatsheet" for more info

R Data Cafe March 11, 2018



## R and Markdown: What is Markdown? (2)

21	TT4
22 - # H1	H1
23 = ## H2	H2
24 - ### H3	
25 - #### H4	H3
26 - ##### H5	H4
27 - ##### H6	
28	H5
29	H6

R Data Cafe March 11, 2018 5



### R and Markdown: What is Markdown? (2)

```
28
29 * ## Emphasis
30
31 Emphasis, aka italics, with *asterisks* or _underscores_.
32
33 Strong emphasis, aka bold, with **asterisks** or __underscores__.
34
35 Combined emphasis with **asterisks and _underscores_**.
36
37 Strikethrough uses two tildes. ~~Scratch this.~~
38
```

#### **Emphasis**

Emphasis, aka italics, with asterisks or underscores.

Strong emphasis, aka bold, with asterisks or underscores.

Combined emphasis with asterisks and *underscores*.

Strikethrough uses two tildes. Scratch this.



### R and Markdown: What is Markdown? (3)

#### Lists

- Item 1
- Item 2
- Item 3

```
38
39 * ## Lists
40
41 - Item 1
42 - Item 2
43 - Item 3
44
```

7

R Data Cafe March 11, 2018



### R and Markdown: What is Markdown? (4)

#### Code block

```
## speed dist
## Min. : 4.0 Min. : 2.00
## 1st Qu.:12.0 1st Qu.: 26.00
## Median :15.0 Median : 36.00
## Mean :15.4 Mean : 42.98
## 3rd Qu.:19.0 3rd Qu.: 56.00
## Max. :25.0 Max. :120.00
```



### R Markdown header (YAML)

```
1 ---
2 title: "Example for R Data Cafe!"
3 author: "Jonathan de Bruin"
4 date: "4/11/2018"
5 output: pdf_document
6 ---
7
```

### Example for R Data Cafe!

Jonathan de Bruin 4/11/2018

#### Introduction R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

R Data Cafe



### Installation

- 'rmarkdown' package
- For pdf output, you need Pandoc (https://pandoc.org/)

R Data Cafe March 11, 2018 10



### Demo

R Data Cafe March 11, 2018 11