

Scientific Reproducible and Provenance with



R Markdown

Plain text file with 3 types of content:

The screenshot shows the RStudio interface with an R Markdown file open. The file contains the following content:

```
1 ---  
2 title: "R Notebook"  
3 output: html_notebook  
4 ---  
5  
6 Text written in **markdown**  
7  
8 ```{r}  
9 # code written in R  
10 (x <- rnorm(7))  
11 ...  
12  
13 Text written in _markdown_  
14  
15 ```{r}  
16 # code written in R  
17 hist(x)  
18 ...  
18:4 (Top Level) :  
Console
```

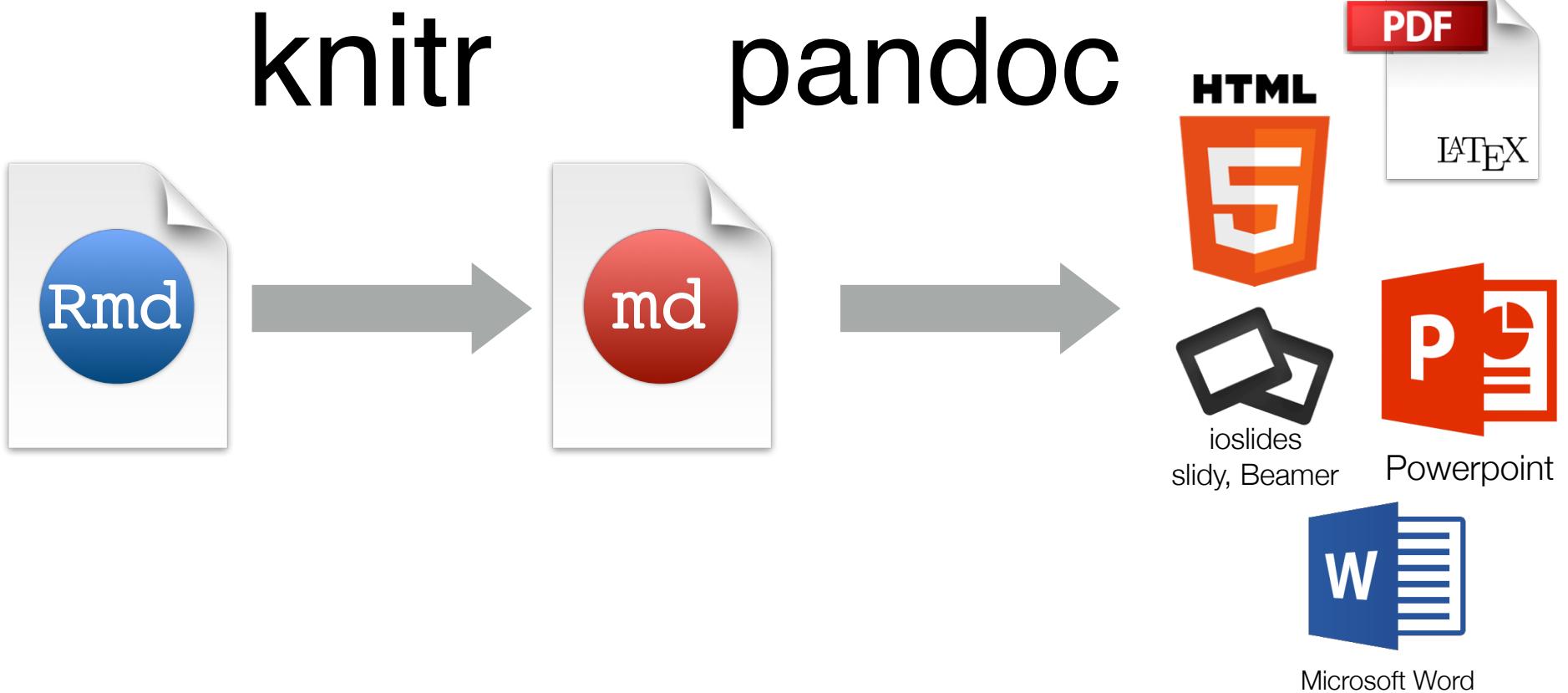
The RStudio interface includes a toolbar at the top, a sidebar on the left, and a preview pane on the right.

A YAML header surrounded by ---

Text in markdown

Code chunks surrounded by ```

How it works



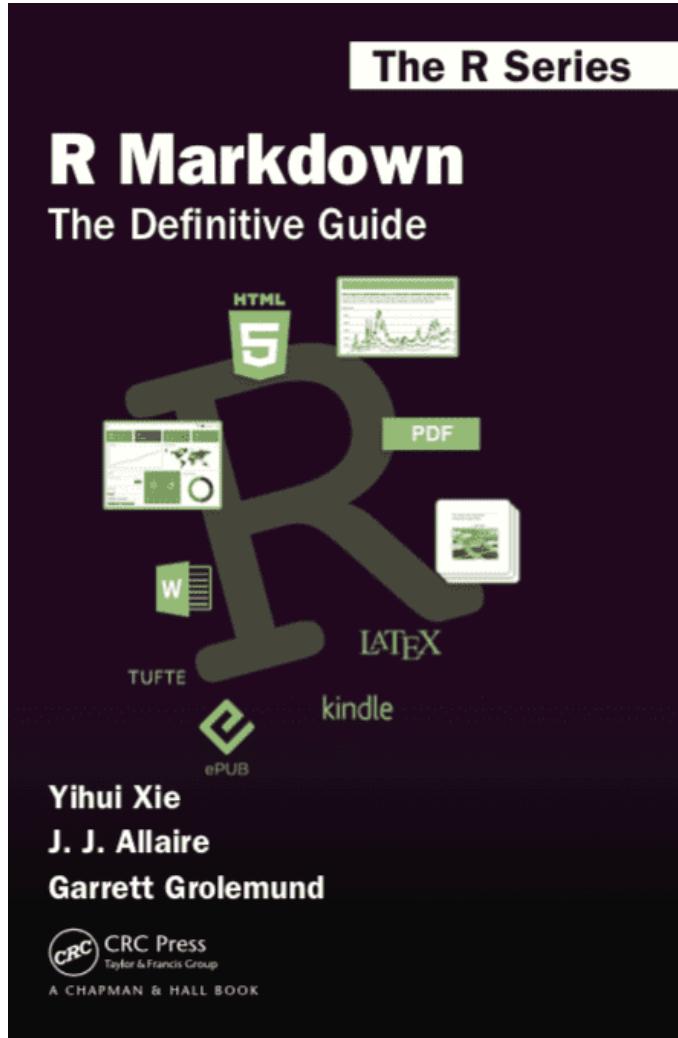
How it works

1

Knitr runs the document in a fresh R session, which means you need to load the libraries that the document uses [in the document](#)

How it works

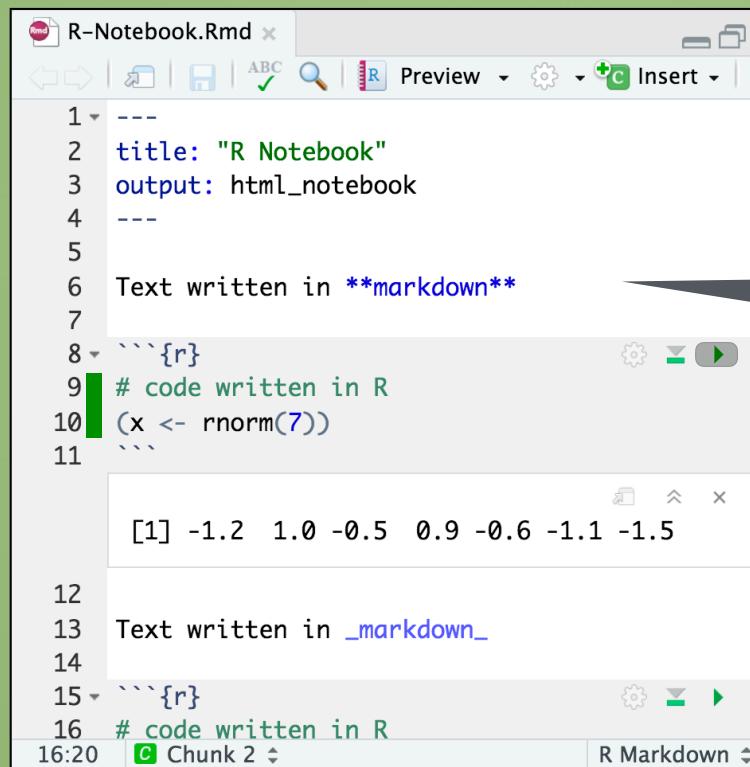
- 1 Knitr runs the document in a fresh R session, which means you need to load the libraries that the document uses [in the document](#)
- 2 Objects made in one code chunk will be available to code in later code chunks.



bookdown.org/yihui/rmarkdown/

ONLINE, FREE

Markdown



```
R-Notebook.Rmd x
1 ---  
2 title: "R Notebook"  
3 output: html_notebook  
4 ---  
5  
6 Text written in **markdown**  
7  
8 ```{r}  
9 # code written in R  
10 x <- rnorm(7)  
11 ...  
[1] -1.2 1.0 -0.5 0.9 -0.6 -1.1 -1.5  
12  
13 Text written in markdown  
14  
15 ```{r}  
16 # code written in R  
16:20 C Chunk 2 R Markdown
```

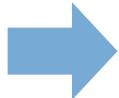
Text in
markdown

Headers

Use # to create headers.

Multiple #'s create lower level headers.

```
# Header 1  
## Header 2  
### Header 3  
#### Header 4  
##### Header 5  
##### Header 6
```



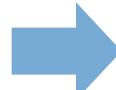
```
Header 1  
Header 2  
Header 3  
Header 4  
Header 5  
Header 6
```

Text

Add two spaces at the end of a line to start a new line

Text is rendered as plain text.
Surround text with _, ******, or ``` to format it.

Text
italics
bold
`code`



Text
italics
bold
`code`

Lists

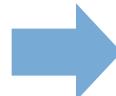
Use asterisks to make bullet points.
Use numbers to make numbered lists.

Bullets

- * bullet 1
- * bullet 2

Numbered list

1. item 1
2. item 2



Bullets

- bullet 1
- bullet 2

Numbered list

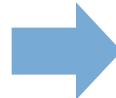
1. item 1
2. item 2

Hyperlinks

Use brackets to denote a link.

Place the URL in parentheses.

This is a
[link]
(www.git.com).



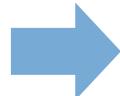
This is a **link**.

Images

Use a link preceded by an ! to insert an image.

The link text should be a URL (if the image is hosted online), or a file path (if the image is saved as a file)

The RStudio
logo.



The RStudio logo.

Equations

Write equations with latex math commands and surround them in \$'s.

According
to
Einstein,
 $E=mc^2$



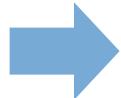
According to
Einstein, $E = mc^2$

Equation blocks

Use two \$'s to make
centered equation blocks.

According to
Einstein,

$\$\$E=mc^2\$\$$

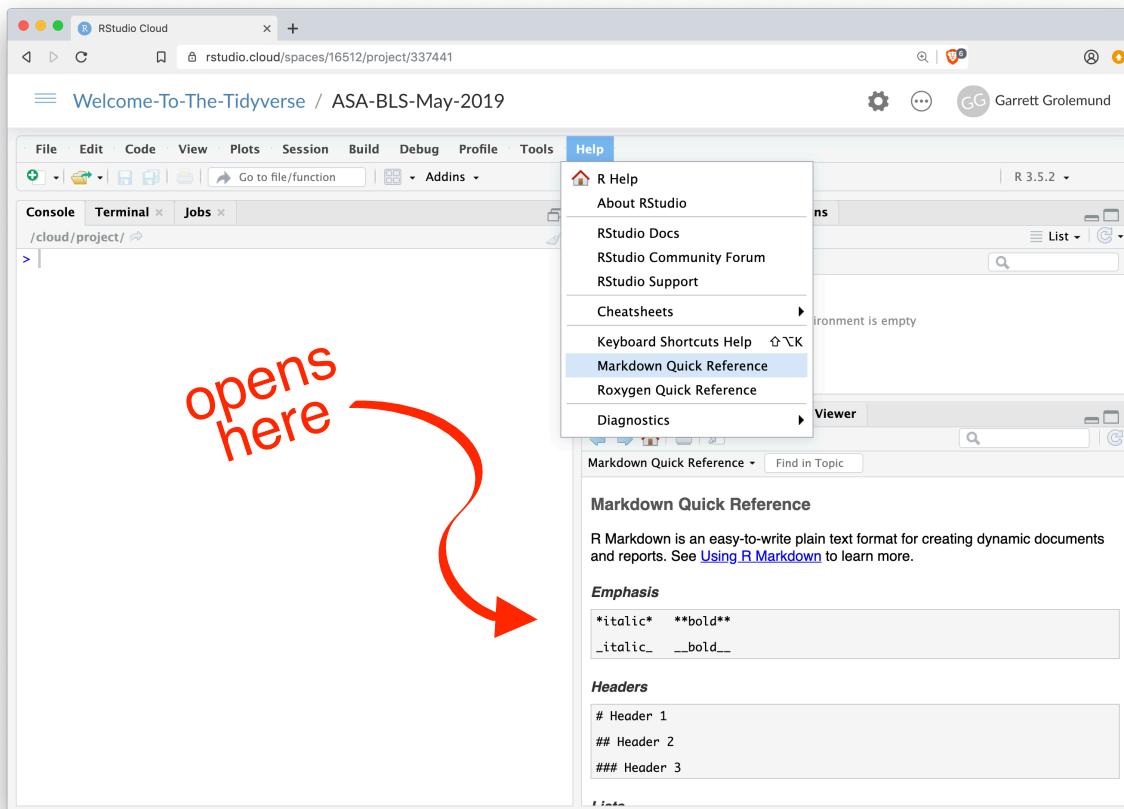


According to
Einstein,

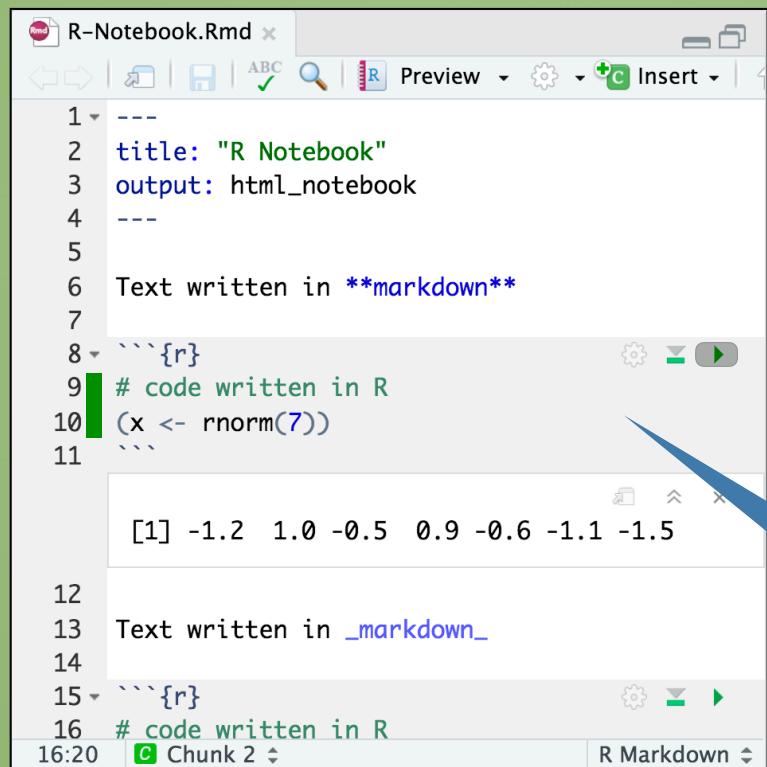
$E = mc^2$

IDE Reference

Go to Help > Markdown Quick Reference



Code



A screenshot of an R-Notebook window titled "R-Notebook.Rmd". The code editor shows the following R Markdown code:

```
1 ---  
2 title: "R Notebook"  
3 output: html_notebook  
4 ---  
5  
6 Text written in **markdown**  
7  
8 ```{r}  
9 # code written in R  
10 x <- rnorm(7)  
11 ...  
[1] -1.2 1.0 -0.5 0.9 -0.6 -1.1 -1.5  
12  
13 Text written in _markdown_  
14  
15 ```{r}  
16 # code written in R
```

The notebook interface includes a toolbar with icons for file operations, search, and preview. A status bar at the bottom shows the time as 16:20 and the chunk number as "Chunk 2".

Code chunks
surrounded by ````

Insert a chunk of R code with

```
```{r}
some code
```
```

When you render the report, R Markdown will run the code and include its results. R Markdown will also remove the ```{r} and ```.

Insert a chunk of R code with

```
```{r}  
some code
```
```

⌘

+

Opt

+

i

(Mac)

Ctrl

+

Alt

+

i

(PC)

chunk options

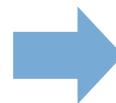
By default, R Markdown includes both the code and its results

Here's some code

```
```{r}
```

```
dim(iris)
```

```
```
```



Here's some code

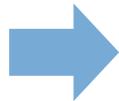
```
dim(iris)
```

```
## [1] 150 5
```

echo

Add options in the brackets after r.
echo = FALSE hides the code.

```
Here's some  
code  
```{r  
echo=FALSE}
dim(iris)
```
```



```
Here's some code  
## [1] 150    5
```

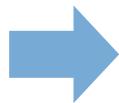
Useful
for plots!

eval

`eval = FALSE` prevents the code from being run. As a result, no results will be displayed with the code.

Here's some
code

```
```{r  
eval=FALSE}
dim(iris)
```
```



Here's some code

```
dim(iris)
```

include

`include = FALSE` runs the code, but prevents both the code and the results from appearing (e.g. to setup).

```
Here's some  
code  
```{r  
include=FALSE}
dim(iris)
```
```

Here's some code

fig.height, fig.width

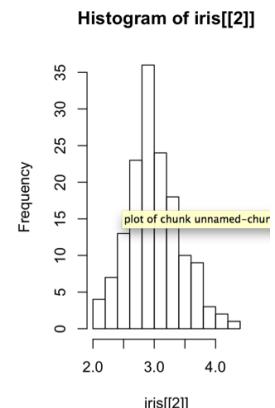
Specify the dimension of plots (in inches) with `fig.width` and `fig.height`. Separate multiple arguments with commas.

Here's a plot

```
```{r echo=FALSE, fig.width=3, fig.height=5}  
hist(iris[[2]])
```
```

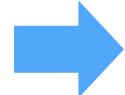


Here's a plot



Place code in a sentence with `r <code>`. R Markdown will replace the code with its results.

```
Today is  
`r  
Sys.Date()`  
.
```



```
Today is 2015-04-16.
```

Code whose results are inserted

```
Today is `r Sys.Date()`.
```

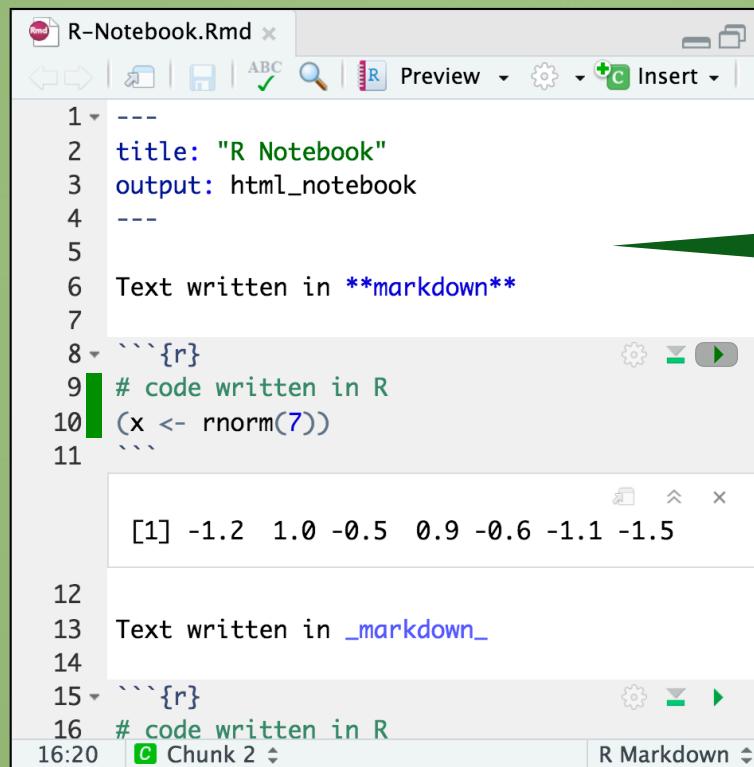
Surround
with `r`

Code to run. Only the result
will be included.

1. Open 05_exercise_report.Rmd
2. Notice the long set-up chunk! Towards the bottom, you'll see some code that stores objects that you will need to call.
3. Knit the document and see the results
 - Why was the {setup} chunk omitted from the report?
4. Replace every ****BOLD**** with inline R code
5. Remove the {students} chunk so it is not included with the output
6. Finally, change the {graph} chunk so that only the output of the plot is shown, but not the code.
7. Re-Knit the document



YAML



A screenshot of an R-Notebook window titled "R-Notebook.Rmd". The code editor shows the following content:

```
1 ---  
2 title: "R Notebook"  
3 output: html_notebook  
4 ---  
5  
6 Text written in **markdown**  
7  
8 ```{r}  
9 # code written in R  
10 x <- rnorm(7)  
11 ...  
[1] -1.2 1.0 -0.5 0.9 -0.6 -1.1 -1.5  
12  
13 Text written in _markdown_  
14  
15 ```{r}  
16 # code written in R  
16:20 [C] Chunk 2 ▾ R Markdown ▾
```

The code editor has syntax highlighting for YAML and R. A callout bubble points to the YAML header with the text "A YAML header surrounded by ---".

A YAML header
surrounded by

YAML

A section of key:value pairs
separated by dashed lines ——

```
---  
title: "Untitled"  
author: "RStudio"  
date: "February 4, 2015"  
output: html_document  
---  
Text of document
```



Untitled

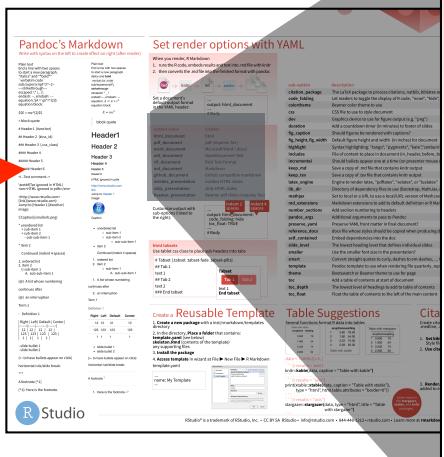
RStudio

February 4, 2015

Text of document

output

The output: field sets the format of



| output value | creates |
|-----------------------|----------------------------------|
| html_document | html |
| pdf_document | pdf (requires Tex) |
| word_document | Microsoft Word (.docx) |
| odt_document | OpenDocument Text |
| rtf_document | Rich Text Format |
| md_document | Markdown |
| github_document | Github compatible markdown |
| ioslides_presentation | ioslides HTML slides |
| slidy_presentation | slidy HTML slides |
| beamer_presentation | Beamer pdf slides (requires Tex) |

More at rmarkdown.rstudio.com/formats.html

Parameters

A faint watermark of the R logo is located in the bottom right corner of the slide. The logo consists of a dark green circle containing a light green letter 'R'.

Open 05_params_exercise.Rmd.

Click the dropdown menu next to Knit and use Knit with Parameters to render the document.

What happens if you type in a different name?



Parameters

A list of values that you can call in R code chunks

params list
elements
and values



```
---
```

```
title: "Untitled"
```

```
output: html_document
```

```
params:
```

```
  filename: "data.csv"
```

```
  symbol: "FB"
```

```
---
```

colon

New line.
Indented two
spaces

Using Parameters

Call parameter values as elements
of the params list, `params$num`

```
---
```

```
params:
```

```
  num: 42
```

```
---
```

```
The value of the
```

```
parameter is
```

```
`r params$num` , e.g.
```

```
```{r}
```

```
params$num
```

```

```



The value of the  
parameter is 42, e.g.

```
params$num
```

```
[1] 42
```

Re-inspect  
05\_params\_exercise.Rmd.

How many times was the  
parameter called in the code?



# Homework

Open 05\_homework.Rmd.

Summarize each of the datasets

Insert the title “Tidybiology Dataset  
Summary”

Knit the final dataset summary report



# Acknowledgements

*Teaching Assistants*

*Allie Mills, Ph.D.*

*Akshay Bareja, D.Phil.*

*Inspiration, ideas, packages, code*

*#R4DS (Garrett Grolemund and Hadley Wickham)*

*Mine Çetinkaya-Rundel (datasciencebox.org)*

*Chester Ismay and Albert Y. Kim (Modern Dive)*

*Garrett Grolemund (Remastering the Tidyverse)*

*Tidyverse devs and community*

*Rstudio*

 @matthewhirshey

 www.hirscheylab.org

 github.com/hirscheylab