

## Introduction to R Markdown

<https://github.com/psboonstra/markdown-workshop>

BDSI 2021; University of Michigan



The basics

## When to use

- ▶ Reports
- ▶ Slides
- ▶ Manuscripts / books
- ▶ Websites

## Why to use

- ▶ R code and interpretations integrated into a single document
- ▶ Separates task of *reporting* the results from *formatting* the results:
  - ▶ decreases risk of copy-paste errors
  - ▶ decreases workload
- ▶ Quickly create the same document in different formats, e.g. slides to show and handouts for the audience

# How it works



When you run `render`, R Markdown feeds the .Rmd  
(.md) document which includes the code and it's o

# How it works



When you run `render`, R Markdown feeds the .Rmd  
(.md) document which includes the code and it's o

# How it works



When you run `render`, R Markdown feeds the .Rmd  
(.md) document which includes the code and it's o

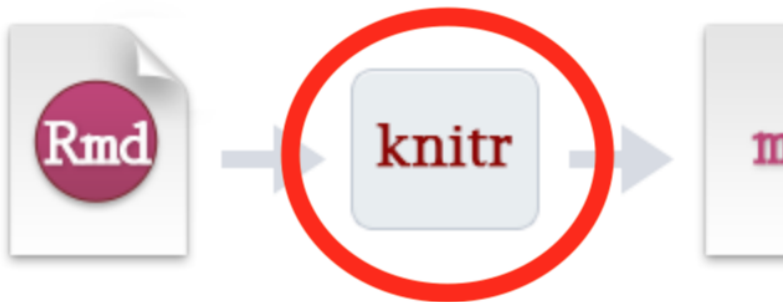


# How it works



When you run `render`, R Markdown feeds the .Rmd  
(.md) document which includes the code and it's o

# How it works



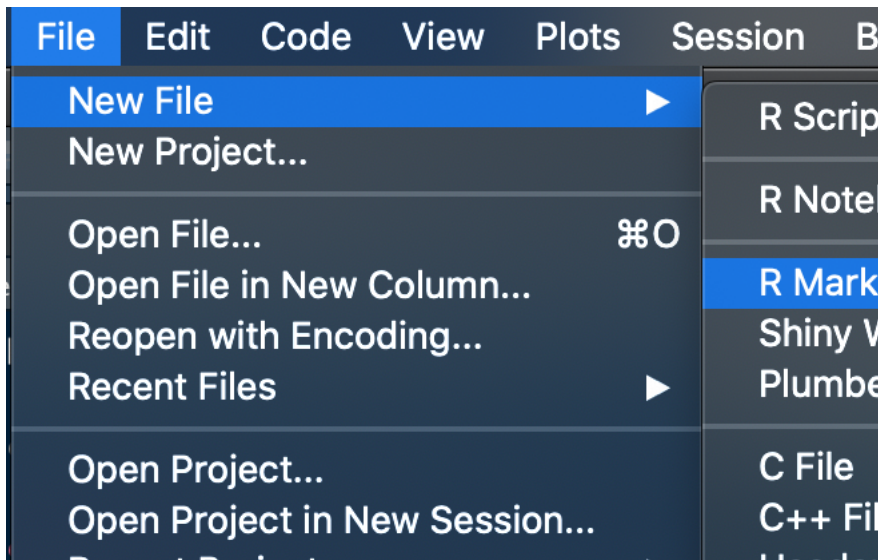
When you run `render`, R Markdown feeds the .Rmd  
(.md) document which includes the code and it's o

# How it works




When you run `render`, R Markdown feeds the .Rmd  
(.md) document which includes the code and it's o


From RStudio, go to File > New File > R Markdown...





## Choose your document type

New R Markdown

 Document

 Presentation

 Shiny

 From Template

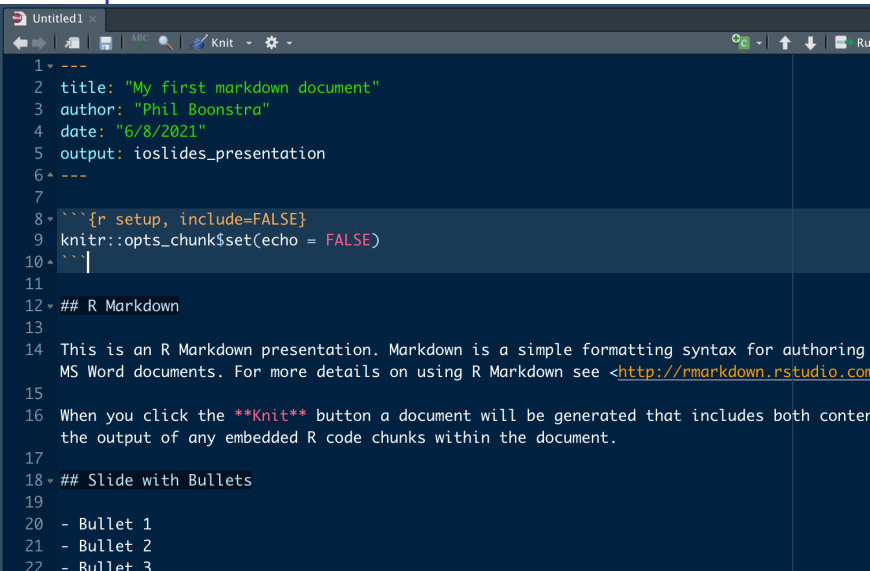
**Title:**

**Author:**

**Default Output Format:**

- ☒ **HTML (ioslides)**  
HTML presentation viewable with any browser (you can also print ioslides to PDF with Chrome).
- ☐ **HTML (Slidy)**  
HTML presentation viewable with any browser (you can also print Slidy to PDF with Chrome).
- ☐ **PDF (Beamer)**  
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).
- ☐ **PowerPoint**  
PowerPoint previewing requires an installation of PowerPoint or OpenOffice.

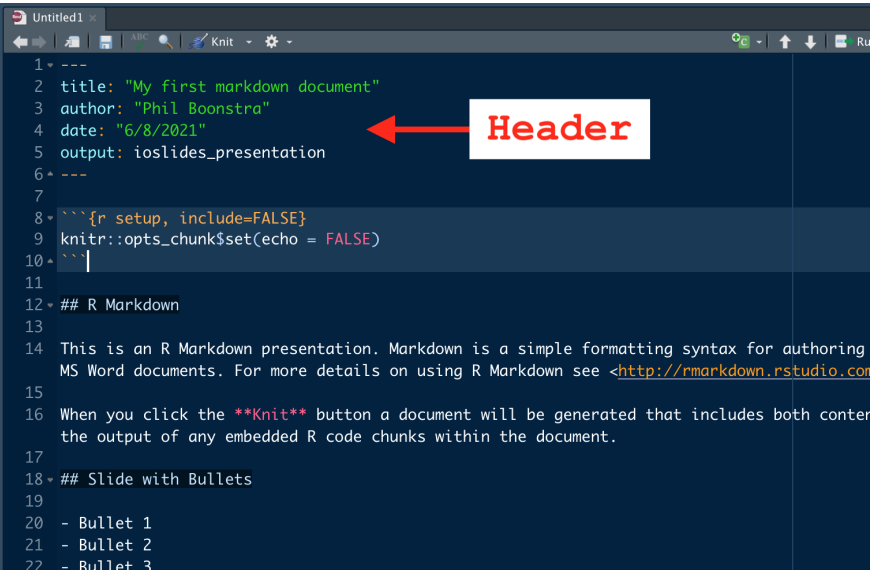
# Get a template



The image shows a screenshot of the RStudio IDE interface. The title bar at the top reads "Untitled1". Below the title bar is a toolbar with icons for navigation, editing, and running code. The main editor area displays a template R Markdown document with the following content:

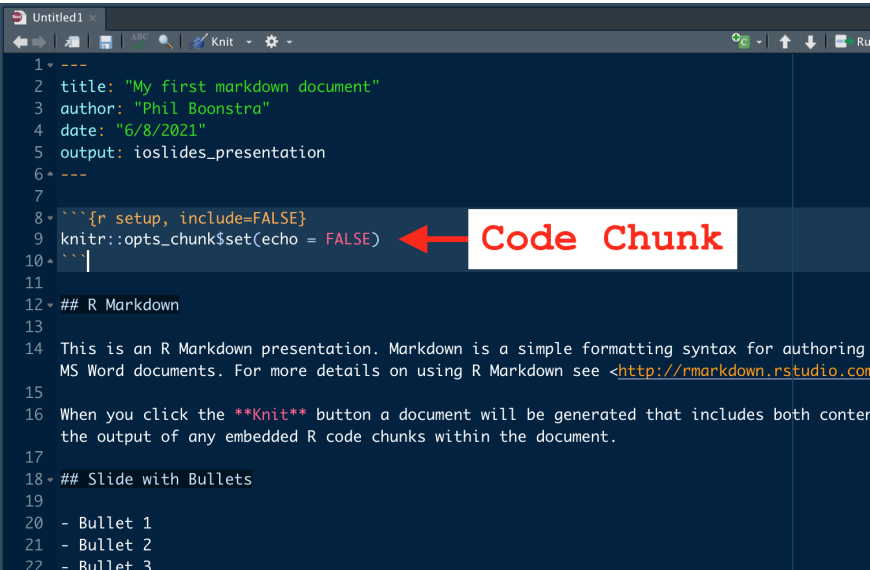
```
1 ▾ ---
2 title: "My first markdown document"
3 author: "Phil Boonstra"
4 date: "6/8/2021"
5 output: ioslides_presentation
6 ▲ ---
7
8 ▾ ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = FALSE)
10 ▲ ```
11
12 ▾ ## R Markdown
13
14 This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring
15 MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com
16
17 When you click the **Knit** button a document will be generated that includes both content
18 the output of any embedded R code chunks within the document.
19
20 ▾ ## Slide with Bullets
21
22 - Bullet 1
23 - Bullet 2
24 - Bullet 3
```

# “YAML” Header



```
1 ---
2 title: "My first markdown document"
3 author: "Phil Boonstra"
4 date: "6/8/2021"
5 output: ioslides_presentation
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = FALSE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring
15 MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>
16
17 When you click the Knit button a document will be generated that includes both content
18 and the output of any embedded R code chunks within the document.
19
20 ## Slide with Bullets
21
22 - Bullet 1
23 - Bullet 2
24 - Bullet 3
```

# Write R code in chunks



```
1 ---
2 title: "My first markdown document"
3 author: "Phil Boonstra"
4 date: "6/8/2021"
5 output: ioslides_presentation
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = FALSE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring
15 MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>
16
17 When you click the Knit button a document will be generated that includes both content
18 and the output of any embedded R code chunks within the document.
19
20 ## Slide with Bullets
21
22 - Bullet 1
23 - Bullet 2
24 - Bullet 3
```

Code Chunk

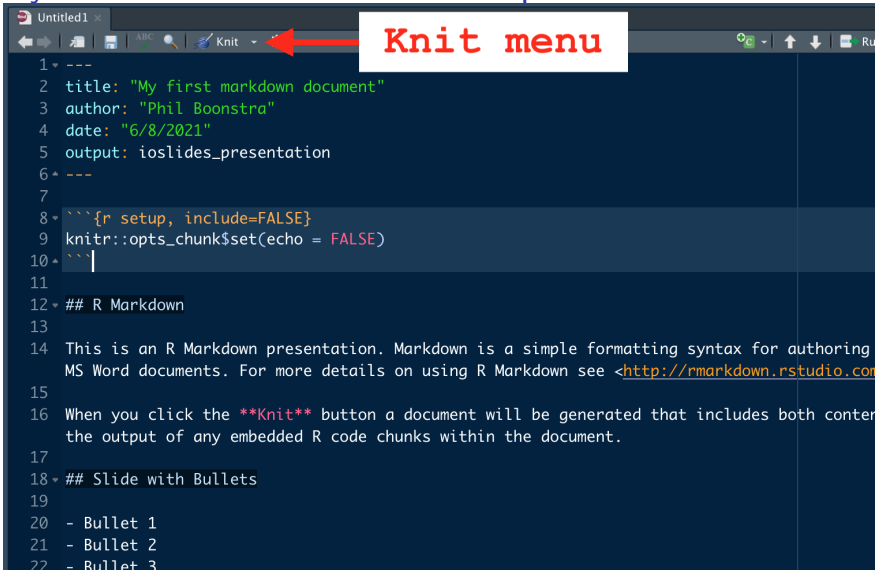


## Write plain text

```
Untitled1 x
← → | 📄 | 🔍 | 🛠 Knit | ⚙
1 ▾ ---
2 title: "My first markdown document"
3 author: "Phil Boonstra"
4 date: "6/8/2021"
5 output: ioslides_presentation
6 ▸ ---
7
8 ```${r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = FALSE)
10 ```
11
12 ▾ ## R Markdown
13
14 This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring
  MS Word documents. For more details on using R Markdown see https://yihui.org/en/articles/2014/02/06-r-markdown.html
15
16 When you click the Knit button a document will be generated containing both content and
  the output of any embedded R code chunks within the document.
17
18 ▾ ## Slide with Bullets
19
20 - Bullet 1
21 - Bullet 2
22 - Bullet 3
```

Plain text

## Knit your document to see the final product



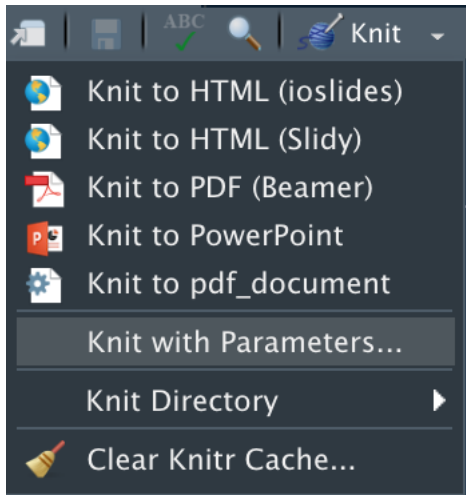
```
1 ---
2 title: "My first markdown document"
3 author: "Phil Boonstra"
4 date: "6/8/2021"
5 output: ioslides_presentation
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = FALSE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring
15 MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>
16
17 When you click the Knit button a document will be generated that includes both content
18 and the output of any embedded R code chunks within the document.
19
20 ## Slide with Bullets
21
22 - Bullet 1
23 - Bullet 2
24 - Bullet 3
```

Knit your document to see the final product

## R Markdown

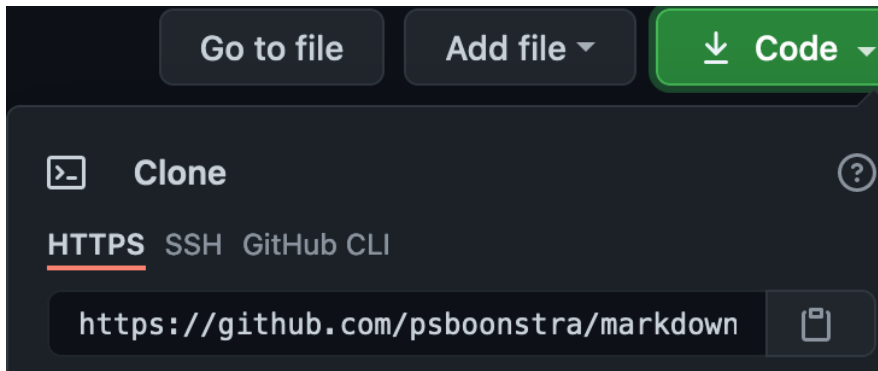
This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

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.



## Try it out: Option 1

- Download R (<https://cran.r-project.org/>)
- Download RStudio to interface with R (<https://www.rstudio.com/>)
- Go to <https://github.com/psboonstra/markdown-workshop>, then 'Code', then 'Download ZIP'



## Try it out: Option 2

- a. Go to <https://rstudio.cloud/> > Get Started
- b. Create an account
- c. Click the dropdown menu *next to* the New Project button, and enter the workshop URL of the workshop repository:  
<https://github.com/psboonstra/markdown-workshop>
- d. Click on 'Files' at the bottom, and pull up  
`01-exercise.Rmd`

Your turn

## Takeaways

- ▶ Chunk options control how the chunk is evaluated and used
- ▶ You can knit the same document to different formats (sometimes easy to do, sometimes requires a bit of finagling)
- ▶ Consider using in-line chunks instead of hard-coding results



## Use Markdown to tell your story

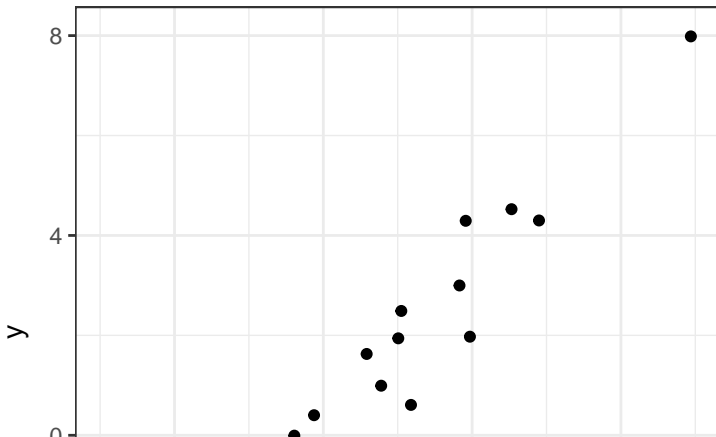
If you name a variable in an earlier code chunk, you can use it again in a later chunk.

early code chunk

```
x <- rnorm(20);  
y <- 3 * x + rnorm(length(x));  
foo = tibble(x = x, y = y);
```

## later code chunk

```
library(ggplot2)
ggplot(data = foo) +
  geom_point(aes(x, y));
```



## Tables

```
foo;
```

```
## # A tibble: 20 x 2
##           x           y
##       <dbl>      <dbl>
##  1  1.45      4.30
##  2  0.589     0.607
##  3  0.0883   -0.257
##  4  2.47      7.98
##  5 -0.990    -3.84
##  6 -0.0622   0.402
##  7  0.985     1.97
##  8 -1.02     -3.44
##  9 -0.195    -0.00781
## 10  0.957     4.29
## 11 -0.0498  -0.806
```

## Tables using 'kable'

x	y
1.44956	4.29944
0.58912	0.60740
0.08832	-0.25688
2.47001	7.98460
-0.98951	-3.83577
-0.06223	0.40188
0.98532	1.97302
-1.01918	-3.44103
-0.19476	-0.00781
0.95700	4.29272
-0.04975	-0.80609
1.26492	4.52509
0.29139	1.62786
0.91519	2.99829

## Other Markdown basics

- ▶ Use #, ##, ###, etc to indicate deeper layers of a header
- ▶ Use \*, + for bulleted (unordered) lists
- ▶ Use (i), (a), or 1. for ordered lists
- ▶ Use `{text}` for *italics*, `{text}` for **bold**

Random lessons I've learned

Markdown can be really, really finicky about horizontal and vertical spacing

If something (a new header option, a code chunk, etc) is not working as you expect, try adding an additional linebreak

If experimenting with a new feature, re-knit frequently



## Caching

If, like me, you become a compulsive re-knitter, the code chunk option `cache = TRUE` is both useful and dangerous.

```
```{r, cache = TRUE}
```

```
(some intensive task)
```

```
```
```

As long as you don't change *anything* in the chunk, you won't need to re-run the intensive task upon re-knitting. However, things can go awry...

- ▶ Open the file `caching_mishap.Rmd` and make sure you understand the intended behavior (should be trivial!)
- ▶ Knit the document
- ▶ Now edit your first chunk, changing to `x <- rnorm(n = 1, mean = 100)` and leaving the second chunk alone
- ▶ Re-knit your document

That's how we get results like this:

```
x <- rnorm(n = 1, mean = 100);
```

```
x;
```

```
## [1] 1.214
```

## What happened

We triggered a recache of the first chunk without triggering a recache of the second

## Possible solutions

- ▶ Cache with caution and only cache costly chunks
- ▶ Think about when and where you want to split your chunk
- ▶ For chunks that may be susceptible, trigger a re-cache by adding a comment character (#) at the end of a line, or making some other innocuous change to your chunk. Even extra white space will trigger a re-cache
- ▶ Go to Knit > Clear Knitr Cache... or delete directly the folder ending in `[filename]_cache` in your working directory

## knitr can run code in other languages

Including

- ▶ Python
- ▶ SQL
- ▶ Julia
- ▶ Stan
- ▶ Javascript

Use ````{python}` to start a python code chunk, ````{julia}` to start a julia code chunk, ````{bash}` to start a Shell script, etc.

You may need external language engines to successfully call other languages. I have not used this functionality before.

More practice

## You can knit R scripts!

You are not limited to using Markdown in Rmd files – you can knit R scripts using the same shortcut: *Cmd+Shift+K* / *Ctrl+Shift+K*

- ▶ Use #' to indicate a switch to markdown
- ▶ Use #+ to start a new chunk

Your turn again

Open 02-exercise.R and complete the tasks. Indicate when you are done.



## Data analyses in R

## readr package

Part of the tidyverse (along with dplyr and ggplot2):



readr gives you tools to read in data from files outside R,  
wrangled and manipulated, and then written to files outside R:

# read data into R



readr::read\_csv  
readr::read\_txt

The workhorse of the readr package is `read_csv`, which reads a comma-separated value (csv) file into R as a `data.frame`. From the help page:

```
read_csv(file, col_names = TRUE, col_types = NULL, locale =  
  na = c("", "NA"), quoted_na = TRUE, quote = "\"", comment = "#",  
  skip = 0, n_max = Inf, guess_max = min(1000, n_max), progress = FALSE,  
  skip_empty_rows = TRUE)
```

Typical use is `my_data <- read_csv("my_files_path.csv")`

## Mouse xenograft study

- ▶  $n = 37$  mice implanted with human tumor
- ▶ Randomized to one of three treatment groups (radiation only; drug only; or both drug and radiation) or no treatment
- ▶ Each tumor on each mouse measured daily for up to 4 weeks
- ▶ Available at American Statistical Association's Section on Teaching of Statistics in the Health Sciences (TSHS) data portal
- ▶ File is called `tumor_growth.csv`

Varna M, Bertheau P, Legres LG. Tumor Microenvironment in Human Tumor Xenografted Mouse Models. *Journal of Analytical Oncology* 2014; 3(3): 159-166.

```
(tumor_growth <- read_csv("tumor_growth.csv"))
```

```
## # A tibble: 574 x 5
```

```
##   Grp   Group   ID   Day   Size
##   <chr> <dbl> <dbl> <dbl> <dbl>
## 1 1.CTR     1   101     0  41.8
## 2 1.CTR     1   101     3   85
## 3 1.CTR     1   101     4  114
## 4 1.CTR     1   101     5 162.
## 5 1.CTR     1   101     6 178.
## 6 1.CTR     1   101     7 325
## 7 1.CTR     1   101    10 624.
## 8 1.CTR     1   101    11 648.
## 9 1.CTR     1   101    12 836.
## 10 1.CTR     1   101    13 1030.
```

```
## # ... with 564 more rows
```

## Digression: testing your dplyr knowledge

```
tumor_growth %>%  
  filter(Day %in% c(0, 14)) %>%  
  group_by(Grp, Day) %>%  
  summarize(mean_Size = mean(Size))
```

## Digression: testing your dplyr knowledge

```
tumor_growth %>%  
  filter(Day %in% c(0, 7, 14)) %>%  
  group_by(Grp, Day) %>%  
  summarize(mean_Size = mean(Size),  
            sd_Size = sd(Size))
```



## Digression: testing your dplyr knowledge

```
tumor_growth %>%  
  filter(Grp == "1.CTR") %>%  
  group_by(ID) %>%  
  summarize(n = n()) %>%  
  summarize(n = mean(n)) %>%  
  pull(n) # pull
```

What to do next

<https://rmarkdown.rstudio.com/>

R Markdown: The definitive guide

- ▶ Free, online version of a book written by the Rstudio experts

R Markdown cheatsheet

- ▶ Helpful quick reference

Mastering markdown

- ▶ Reference site for markdown

## References