



Week 14: *Reproducible Reporting*

EMSE 4574: Intro to Programming for Analytics

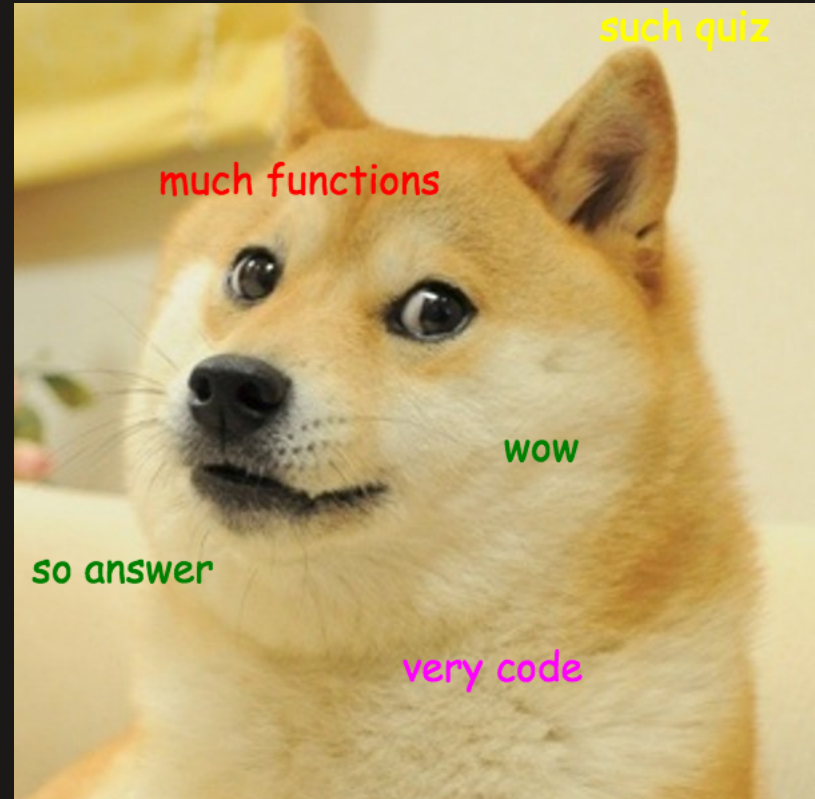
John Paul Helveston

December 01, 2020

Quiz 7 (last one!)

- Go to `#classroom` channel in Slack for link
- Open up RStudio before you start - you'll probably want to use it.

05:00



(You should always cite your sources)

A bunch of today's slides are adapted from the brilliant [Alison Presmanes Hill](#)

 [@apreshill](#)

Check out her [RMarkdown slide deck](#)



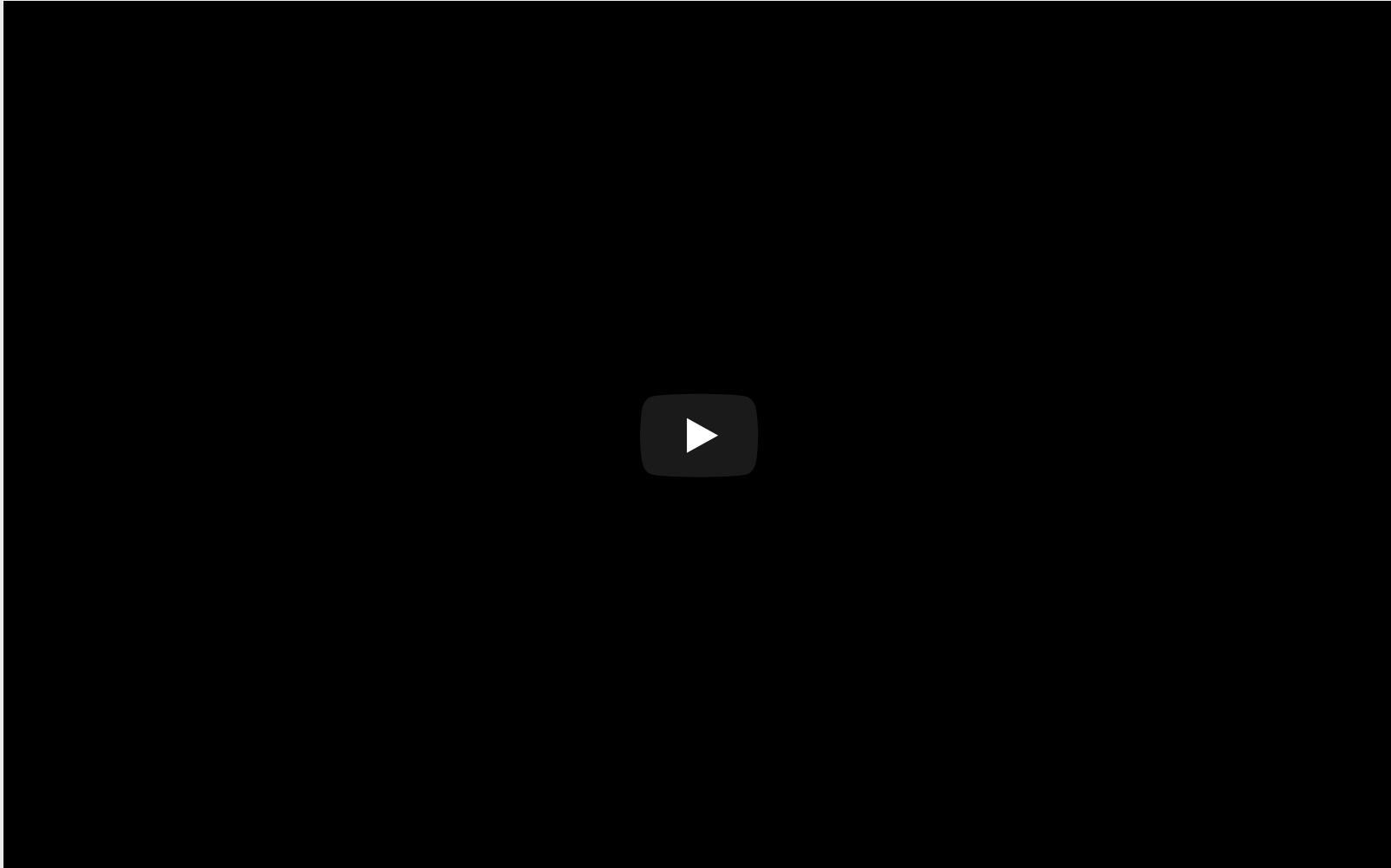
Week 14: *Reproducible Reporting*

1. Why RMarkdown?
2. Metadata and output formats
3. Text
4. Code chunks

Week 14: *Reproducible Reporting*

1. Why RMarkdown?
2. Metadata and output formats
3. Text
4. Code chunks

The horrors of a non-reproducible workflow



RMarkdown to the rescue!



"Literate programming"

Treat programs as a
"literature" understandable
to **human beings**



Donald E. Knuth

Week 14: *Reproducible Reporting*

1. Why RMarkdown?

2. **Metadata and output formats**

3. Text

4. Code chunks

R markdown

TEXT. CODE. OUTPUT.
(GET IT TOGETHER, PEOPLE.)



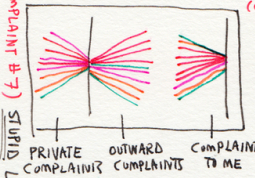
Art by [Allison Horst](#)

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS*

AND GENERAL GRUMPINESS

HOW TO READ IT: (I ~~THREW~~ TREW DOWN MY PENS WHEN I FINISHED)



STUPID LEFT HAND SHOWERING

TYPE OF COMPLAINT:

- WEATHER
- HUSBAND
- ANIMALS
- FAMILY
- SOCIETY!
- THE WORLD TODAY
- ACQUAINTANCES!
- STRANGERS
- MY APPEARANCE
- FRIENDS
- WORK
- HEALTH
- HUNGER
- MYSELF
- TECHNOLOGY/MEDIA
- MONEY
- INANIMATE OBJECTS
- TRANSPORT

ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... ARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH BOTHERS ME!

MAIN STATS	
PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE I COMPLAINED TO:	10
* SECOND MOST COMPLAINTS ABOUT MY HUSBAND (SMART STEVE I LOVE YOU!)	

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!

* AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

FROM:
S POSAVEC
LONDON
UK



TO:
GIORGIA LUPI

BROOKLYN, NY 11249
USA

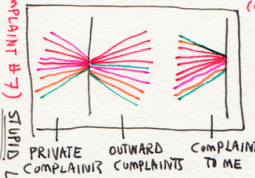
DELIVERED BY
HAND (SPECIAL NYC DELIVERY!)

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS*

AND GENERAL GRUMPINESS

HOW TO READ IT: (I ~~WAS~~ TREW DOWN MY PENS WHEN FINISHED)



- TYPE OF COMPLAINT:
- WEATHER
 - HUSBAND
 - ANIMALS
 - FAMILY
 - SOCIETY!
 - THE WORLD TODAY
 - ACQUAINTANCES!
 - STRANGERS
 - MY APPEARANCE
 - FRIENDS
 - WORK
 - HEALTH
 - HUNGER
 - MYSELF
 - TECHNOLOGY/MEDIA
 - MONEY
 - INANIMATE OBJECTS
 - TRANSPORT

ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... ARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH BOTHERS ME!

MAIN STATS	
PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE WHO COMPLAINED:	MY DAD
COMPLAINED TO:	0
TOP COMPLAINER:	NOT THIS BRAND
I COMPLAINED MOST ABOUT MY:	FAMILY (SORRY MOM + DAD I LOVE YOU!!!)
* SECOND MOST COMPLAINED ABOUT:	MY HUSBAND (SMART STEVE I LOVE YOU!!!)

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!
* AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

FROM:
S POSAVEC
LONDON
UK



TO:
GEORGIA LUPI
BROOKLYN, NY 11249
USA

DELIVERED BY
HAND (SPECIAL NYC
DELIVERY!)

↑
metadata

metadata: YAML

YAML
Ain't
Markup
Language

```
---  
key: value  
---
```



Output options

Save output options in your YAML

```
----  
title: Your title here  
author: Your name here  
output: html_document  
----
```

```
----  
title: Your title here  
author: Your name here  
output:  
  html_document:  
    toc: true  
    toc_float: true  
    theme: flatly  
----
```



The screenshot displays the RStudio environment with a demo.Rmd file open in the editor and its rendered HTML output in the viewer. The editor shows a mix of R code and Markdown text, including a code chunk for setting knitr options and loading data. The viewer shows the rendered HTML, which includes a license notice, a list of features, and introductory text about R and R Markdown.

```
1 ---
2 title: A quick RMarkdown demo
3 subtitle: A summary of just a few features
4 author: John Helveston
5 output:
6   html_document:
7     toc: true
8     toc_float: true
9     theme: flatly
10 ---
11
12 ```{r setup, include=FALSE}
13 knitr::opts_chunk$set(
14   warning = FALSE,
15   message = FALSE,
16   fig.retina = 3)
17
18 # Import libraries
19 library(tidyverse)
20 library(here)
21 # This is an example of loading data within a code chunk
22 birds <- read_csv(here('data', 'wildlife_impacts.csv'))
23 bears <- read_csv(here('data', 'bear_killings.csv'))
24 ```
25
26 # R Markdown
27
28 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 [rmarkdown.rstudio.com](http://rmarkdown.rstudio.com).|
29
30 When you click the Knit button a document will be generated that includes
```

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Name	Size	Modified
..		
.Rhistory	356 B	Nov 30, 2020, 10:38 AM
data		
demo.html	1.2 MB	Nov 30, 2020, 10:36 AM
demo.Rmd	1.5 KB	Nov 30, 2020, 11:00 AM
notes.Rproj	205 B	Nov 30, 2020, 10:59 AM

```
1 ---
2 title: A quick RMarkdown demo
3 subtitle: A summary of just a few features
4 author: John Helveston
5 output:
6   html_document:
7     toc: true
8     toc_float: true
9     theme: flatly
10 ---
11
12 ```{r setup, include=FALSE}
13 knitr::opts_chunk$set(
14   warning = FALSE,
15   message = FALSE,
16   fig.retina = 3)
17
18 # Import libraries
19 library(tidyverse)
20 library(here)
21 # This is an example of loading data within a code chunk
22 birds <- read_csv(here('data', 'wildlife_impacts.csv'))
23 bears <- read_csv(here('data', 'bear_killings.csv'))
24 ```
25
26 # R Markdown
27
28 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 [rmarkdown.rstudio.com](http://rmarkdown.rstudio.com).|
29
30 When you click the Knit button a document will be generated that includes
```

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Name	Size	Modified
..		
.Rhistory	356 B	Nov 30, 2020, 10:38 AM
data		
demo.html	1.2 MB	Nov 30, 2020, 10:36 AM
demo.Rmd	1.5 KB	Nov 30, 2020, 11:00 AM
notes.Rproj	205 B	Nov 30, 2020, 10:59 AM

The image shows the RStudio interface with an R Markdown document on the left and its rendered HTML output on the right. A red bracket on the left side groups the metadata fields (output, html_document, toc, toc_float, theme) and is labeled "metadata = YAML".

```
1 ---
2 title: A quick RMarkdown demo
3 subtitle: A summary of just a few features
4 author: John Helveston
5 output:
6   html_document:
7     toc: true
8     toc_float: true
9     theme: flatly
10 ---
11
12 ```{r setup, include=FALSE}
13 knitr::opts_chunk$set(
14   warning = FALSE,
15   message = FALSE,
16   fig.retina = 3)
17
18 # Import libraries
19 library(tidyverse)
20 library(here)
21 # This is an example of loading data within a code chunk
22 birds <- read_csv(here('data', 'wildlife_impacts.csv'))
23 bears <- read_csv(here('data', 'bear_killings.csv'))
24 ```
25
26 # R Markdown
27
28 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 [rmarkdown.rstudio.com](http://rmarkdown.rstudio.com).|
29
30 When you click the Knit button a document will be generated that includes
```

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Name	Size	Modified
..		
.Rhistory	356 B	Nov 30, 2020, 10:38 AM
data		
demo.html	1.2 MB	Nov 30, 2020, 10:36 AM
demo.Rmd	1.5 KB	Nov 30, 2020, 11:00 AM
notes.Rproj	205 B	Nov 30, 2020, 10:59 AM

The image shows the RStudio interface with an R Markdown document on the left and its rendered HTML output on the right. The document contains metadata, a knitr chunk, and a main text section. Handwritten annotations highlight the metadata section.

```
1 ---
2 title: A quick RMarkdown demo
3 subtitle: A summary of just a few features
4 author: John Helveston
5 output:
6   html_document:
7     toc: true
8     toc_float: true
9     theme: flatly
10 ---
11
12 ```{r setup, include=FALSE}
13 knitr::opts_chunk$set(
14   warning = FALSE,
15   message = FALSE,
16   fig.retina = 3)
17
18 # Import libraries
19 library(tidyverse)
20 library(here)
21 # This is an example of loading data within a code chunk
22 birds <- read_csv(here('data', 'wildlife_impacts.csv'))
23 bears <- read_csv(here('data', 'bear_killings.csv'))
24 ```
25
26 # R Markdown
27
28 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 [rmarkdown.rstudio.com](http://rmarkdown.rstudio.com).|
29
30 When you click the Knit button a document will be generated that includes
```

Handwritten annotation: metadata = YAML

Handwritten annotation: output:

Handwritten annotation: html_document:

Handwritten annotation: toc: true

Handwritten annotation: toc_float: true

Handwritten annotation: theme: flatly

Rendered output:

R is free software and comes with ABSOLUTELY NO WARRANTY. You are welcome to redistribute it under certain conditions. Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'q()' to quit R.

> |

Files | Plots | Packages | Help | Viewer

New Folder | Delete | Rename | More

Home > Downloads > demo

Name	Size	Modified
..		
.Rhistory	356 B	Nov 30, 2020, 10:38 AM
data		
demo.html	1.2 MB	Nov 30, 2020, 10:36 AM
demo.Rmd	1.5 KB	Nov 30, 2020, 11:00 AM
notes.Rproj	205 B	Nov 30, 2020, 10:59 AM

28:215 R Markdown

Environment | History | Connections | Tutorial

QuickTime Player File Edit View Window Help

demo - RStudio

demo.Rmd x

```
1 ---
2 title: A quick RMarkdown demo
3 subtitle: A summary of just a few features
4 author: John Helveston
5 output:
6   html_document:
7     toc: true
8     toc_float: true
9     theme: flatly
10 ---
11
12 ```{r setup, include=FALSE}
13 knitr::opts_chunk$set(
14   warning = FALSE,
15   message = FALSE,
16   fig.retina = 3)
17
18 # Import libraries
19 library(tidyverse)
20 library(here)
21 # This is an example of loading data within a code chunk
22 birds <- read_csv(here('data', 'wildlife_impacts.csv'))
23 bears <- read_csv(here('data', 'bear_killings.csv'))
24 ```
25
26 # R Markdown
27
28 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 [rmarkdown.rstudio.com](http://rmarkdown.rstudio.com).
29
30 Click the Knit button a document will be generated that includes
```

Console R Markdown x Jobs x

~/Downloads/demo/

R is free software and comes with ABSOLUTELY NO WARRANTY. You are welcome to redistribute it under certain conditions. Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'q()' to quit R.

> |

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home > Downloads > demo

Name	Size	Modified
..		
.Rhistory	356 B	Nov 30, 2020, 10:38 AM
data		
demo.html	1.2 MB	Nov 30, 2020, 11:04 AM
demo.Rmd	1.5 KB	Nov 30, 2020, 11:00 AM
notes.Rproj	205 B	Nov 30, 2020, 10:59 AM

5:8 A quick RMarkdown demo - R Markdown

Environment History Connections Tutorial

A quick RMarkdown demo

File | /Users/jhelvy/Downloads/demo/demo.html

Blackboard P4A Lab

R Markdown
Plots

A quick RMarkdown demo

A summary of just a few features

John Helveston

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 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.

Useful tools

- Quick markdown reference guide: <https://commonmark.org/help/>
- Quick demo guide: <https://markdown-it.github.io/>.
- Online table converter: <http://www.tablesgenerator.com>

Link to data source of final practice today

<https://github.com/rfordatascience/tidytuesday/tree/master/data/2018/2018-10-16>

Plots

Using ggplot

Here is an example of how to make a plot with ggplot2:

```
bears %>%  
  count(month) %>%  
  ggplot() +
```

A quick RMarkdown demo

File | /Users/jhelvy/Downloads/demo/demo.html

Blackboard P4A Lab

R Markdown
Plots

A quick RMarkdown demo

A summary of just a few features

John Helveston

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 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.

Useful tools

- Quick markdown reference guide: <https://commonmark.org/help/>
- Quick demo guide: <https://markdown-it.github.io/>
- Online table converter: <http://www.tablesgenerator.com>

Link to data source of final practice today

<https://github.com/rfordatascience/tidytuesday/tree/master/data/2018/2018-10-16>

Plots

Using ggplot

Here is an example of how to make a plot with ggplot2:

```
bears %>%  
  count(month) %>%  
  ggplot() +
```

font
+ = theme
color

A quick RMarkdown demo

File | /Users/jhelvy/Downloads/demo/demo.html

Blackboard P4A Lab

R Markdown

Plots

A quick RMarkdown demo

A summary of just a few features

John Helveston

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 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.

Useful tools

- Quick markdown reference guide: <https://commonmark.org/help/>
- Quick demo guide: <https://markdown-it.github.io/>
- Online table converter: <http://www.tablesgenerator.com>

Link to data source of final practice today

<https://github.com/rfordatascience/tidytuesday/tree/master/data/2018/2018-10-16>

Plots

Using ggplot

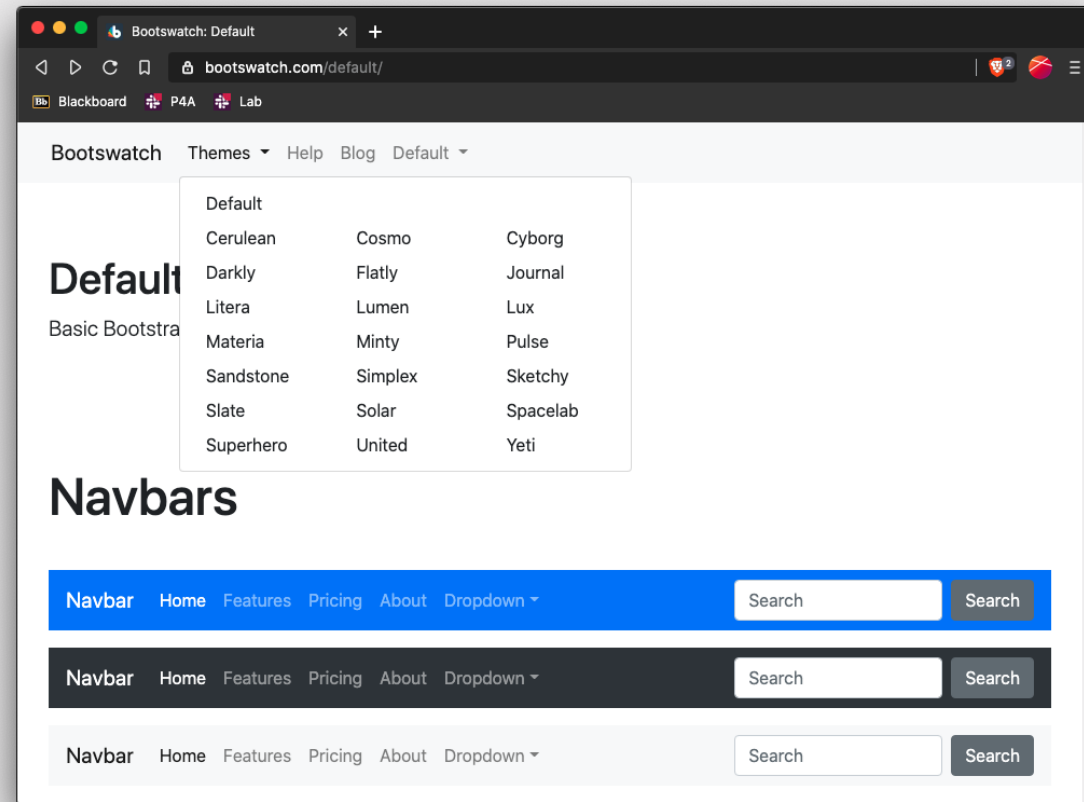
Here is an example of how to make a plot with ggplot2:

```
bears %>%
  count(month) %>%
  ggplot() +
```

toc:
toc_float:

font
+ = theme
color

Preview bootswatch themes



Quick check-in

How do you use only the **default** output options?

(a)

```
----  
output: html_document()  
----
```

(b)

```
----  
output: html_document  
----
```

How do you add an output **option** to a format in your YAML?

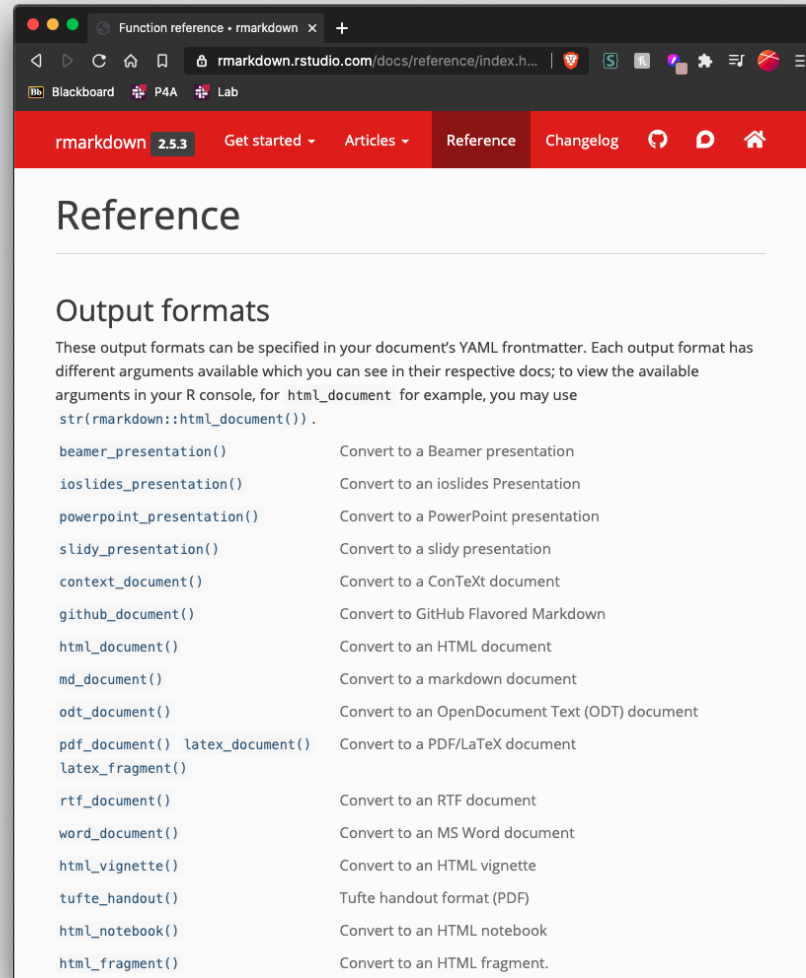
(a)

```
----  
output:  
  html_document:  
    toc: true  
----
```

(b)

```
----  
output:  
  html_document(toc=true)  
----
```

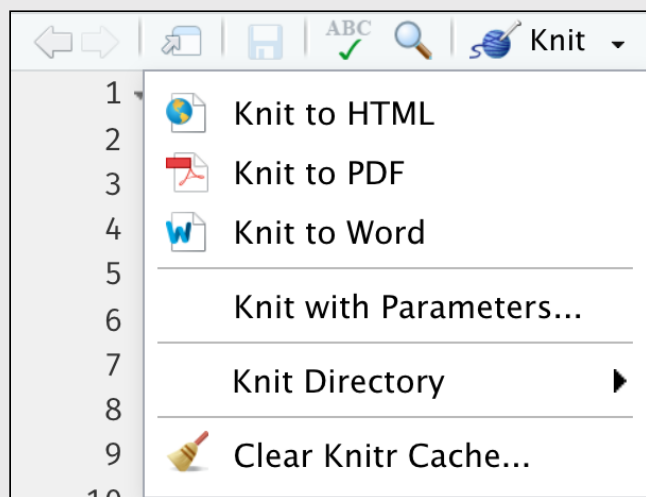
Output formats





Knit to multiple outputs

```
rmarkdown::render("notes.Rmd", output_format = "all")
```



```
---  
title: Your title here  
author: Your name here  
output:  
  html_document:  
    toc: true  
    toc_float: true  
    theme: flatly  
  word_document: default  
  pdf_document: default  
---
```


Quick check-in

How do you add another output **format** to your YAML?

(a)

```
----  
output:  
  html_document: default  
  word_document: default  
----
```

(b)

```
----  
output:  
  html_document()  
  word_document()  
----
```

How do you now add output **options** to your YAML?

(a)

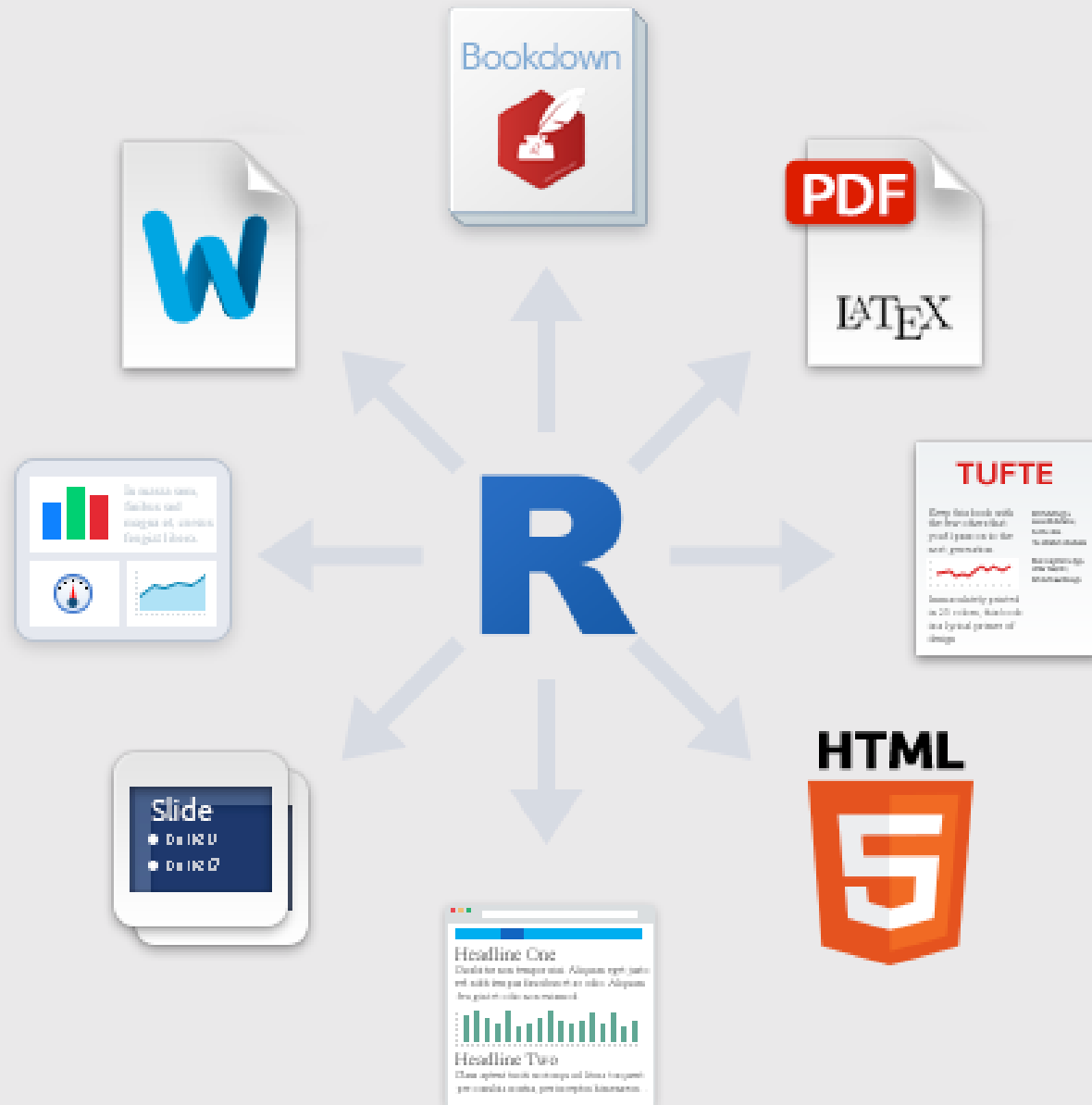
```
----  
output:  
  html_document:  
    toc: true  
  word_document: default  
----
```

(b)

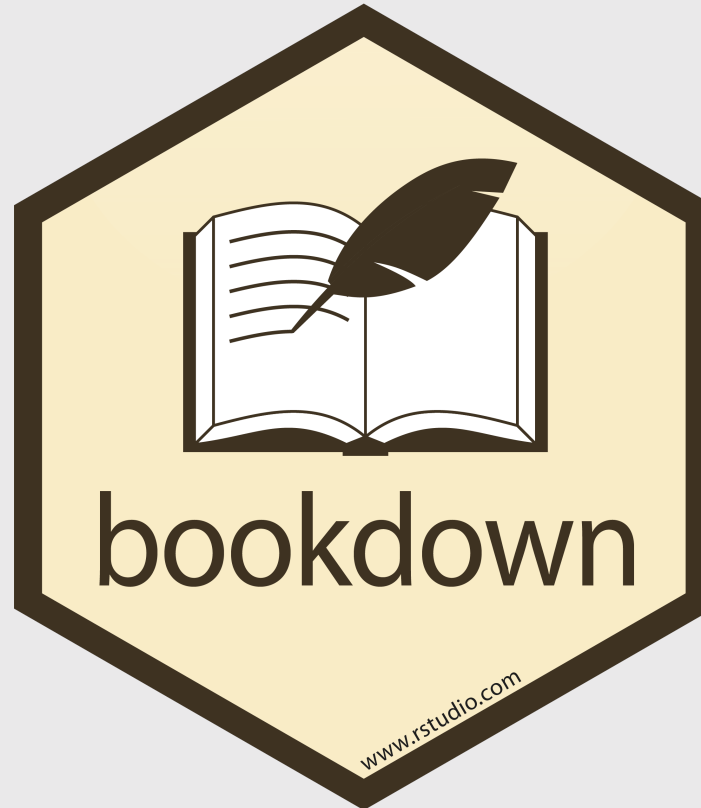
```
----  
output:  
  html_document(toc=true)  
  word_document(default)  
----
```

Built-in output formats





Extension output formats



Distill for R Markdown

Scientific and technical writing, native to the web

AUTHORS

- JJ Allaire
- Rich Iannone
- Alison Presmanes Hill
- Yihui Xie

AFFILIATIONS

- RStudio
- RStudio
- RStudio
- RStudio

PUBLISHED

Sept. 10, 2018

CITATION

Allaire, et al., 2018

Contents

- Installation
- Creating an article
- Figures
- Tables
- Equations
- Citations

Distill for R Markdown is a web publishing format optimized for scientific and technical communication. Distill articles include:

- Reader-friendly typography that adapts well to mobile devices.
- Features essential to technical writing like [LaTeX math](#), [citations](#), and [footnotes](#).
- Flexible figure layout options (e.g. displaying figures at a larger width than the

Use an extension package

```
----  
author: Your name here  
title: Your title here  
output: distill::distill_article  
----
```

```
----  
author: Your name here  
title: Your title here  
output:  
  distill::distill_article:  
    toc: true  
----
```

02:00

Quick practice

Go to your `notes.Rmd` file and knit it to the following outputs:

- `html_document` with a table of contents
- `distill_article` with a table of contents
- `word_document`
- `pdf_document`

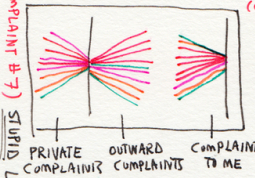
Week 14: *Reproducible Reporting*

1. Why RMarkdown?
2. Metadata and output formats
3. Text
4. Code chunks

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS* AND GENERAL GRUMPINESS

HOW TO READ IT: (I ~~THREW~~ TREW DOWN MY PENS WHEN I FINISHED)



(COMPLAINT #7) STUPID LEFT HAND

(COMPLAINT #7) SHADOWING

GENERAL

(COMPLAINT #11)

(MY STUPID LEFT HAND!)

ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... ARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH BOTHERS ME X

- TYPE OF COMPLAINT:
- WEATHER
 - HUSBAND
 - ANIMALS
 - FAMILY
 - SOCIETY!
 - THE WORLD TODAY
 - ACQUAINTANCES!
 - STRANGERS
 - MY APPEARANCE
 - FRIENDS
 - WORK
 - HEALTH
 - HUNGER
 - MYSELF
 - TECHNOLOGY/MEDIA
 - MONEY
 - INANIMATE OBJECTS
 - TRANSPORT

MAIN STATS	
PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE WHO COMPLAINED TO ME:	MY DAD
COMPLAINED TO:	0
TOP COMPLAINER:	NOT THIS BRAND
I COMPLAINED MOST ABOUT MY:	(SORRY MOM + DAD) I LOVE YOU!!!
* SECOND MOST COMPLAINED ABOUT:	MY HUSBAND (SHEET STEVE I LOVE YOU!!!)

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!

* AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

FROM: S POSAVEC

LONDON UK



TO: GEORGIA LUPI

BROOKLYN, NY 11249
USA

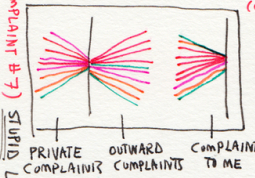
DELIVERED BY
HAND (SPECIAL NYC DELIVERY!)

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS*

AND GENERAL GRUMPINESS

HOW TO READ IT: (I ~~THREW~~ TREW DOWN MY PENS WHEN FINISHED)



- TYPE OF COMPLAINT:
- WEATHER
 - HUSBAND
 - ANIMALS
 - FAMILY
 - SOCIETY!
 - THE WORLD TODAY
 - ACQUAINTANCES!
 - STRANGERS
 - MY APPEARANCE
 - FRIENDS
 - WORK
 - HEALTH
 - HUNGER
 - MYSELF
 - TECHNOLOGY/MEDIA
 - MONEY
 - INANIMATE OBJECTS
 - TRANSPORT

ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... AARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH BOTHERS ME!

MAIN STATS

PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE WHO COMPLAINED TO ME:	MY DAD
COMPLAINED TO:	0
TOP COMPLAINER:	NOT THIS GRAND I COMPLAINED MOST ABOUT MY FAMILY (SORRY MOM + DAD I LOVE YOU!!!)
* SECOND MOST COMPLAINED ABOUT:	MY HUSBAND (SMART STEVE I LOVE YOU!!!)

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!
* AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

FROM: S POSAVEC
LONDON
UK



TO:
GEORGIA LUPI
BROOKLYN, NY 11249
USA

DELIVERED BY
HAND (SPECIAL NYC
DELIVERY!)

metadata

text

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS* AND GENERAL GRUMPINESS

HOW TO READ IT: (I TREW DOWN MY PENS WHEN I FINISHED ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... AARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH IS A BITE

TYPE OF COMPLAINT:

- WEATHER
- HUSBAND
- ANIMALS
- FAMILY
- SOCIETY!
- THE WORLD TODAY
- ACQUAINTANCES!
- STRANGERS
- MY APPEARANCE
- FRIENDS
- WORK
- HEALTH
- HUNGER
- MYSELF
- TECHNOLOGY/MEDIA
- MONEY
- INANIMATE OBJECTS
- TRANSPORT

MAIN STATS

- PRIVATE COMPLAINTS: 67
- OUTWARD COMPLAINTS: 100
- COMPLAINTS TO ME: 43
- # OF PEOPLE WHO COMPLAINED TO ME: 5
- # OF PEOPLE WHO COMPLAINED: MY DAD
- COMPLAINED TO: 10
- * SECOND MOST COMPLAINTS ABOUT MY HUSBAND (SMART STEVE I LOVE YOU)

FROM: S POSAVEC

LONDON UK

97P

TO: GEORGIA LUPI

BROOKLYN, NY 11249

USA

DELIVERED BY HAND (SPECIAL NYC DELIVERY!)

metadata

<https://www.dear-data.com/>

Right now, bookmark this! 👉

<https://commonmark.org/help/>

(When you have 10 minutes, do this! 👉)

<https://commonmark.org/help/tutorial/>

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

Childhood ****vaccines****
are one of the
great triumphs
of modern medicine.

Childhood **vaccines**
are one of the
great triumphs
of modern medicine.

Text

Type this...

- normal text
- *italic text*
- **bold text**
- ***bold italic text***
- ~~strikethrough~~
- `code text`

..to get this

- normal text
- *italic text*
- **bold text**
- ***bold italic text***
- ~~strikethrough~~
- code text

Lists

Bullet list:

```
- first item  
- second item  
- third item
```

- first item
- second item
- third item

Numbered list:

```
1. first item  
2. second item  
3. third item
```

1. first item
2. second item
3. third item

Links

Simple **url link** to another site:

```
[Download R] (http://www.r-project.org/)
```

[Download R](http://www.r-project.org/)

Images



Local images

```

```



Quick check-in

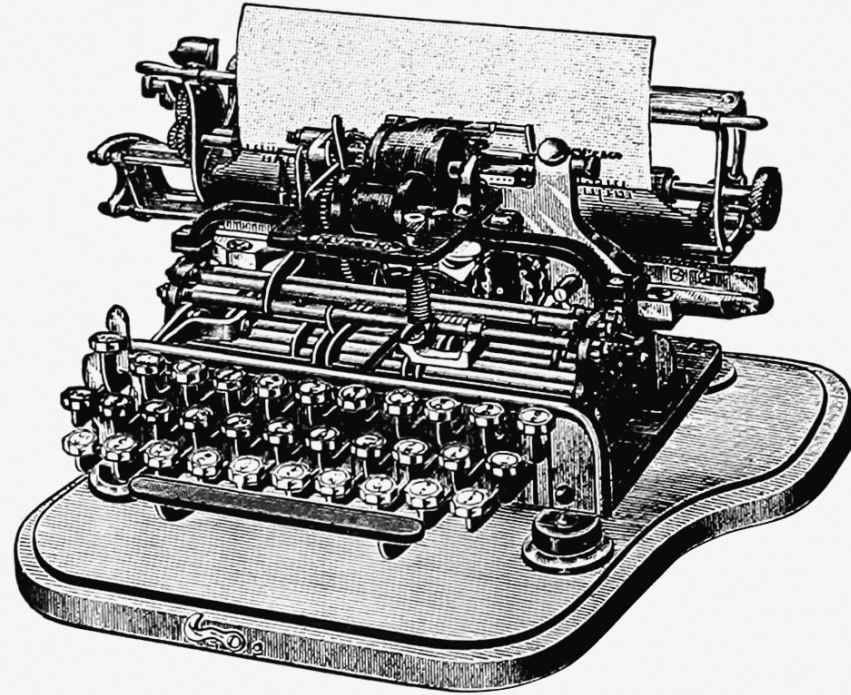
How do you add headers in Markdown?

- a. ! Header
- b. - Header
- c. # Header
- d. 1. Header

What about lists? Bulleted? Numbered?

- a. ! Item 1
- b. - Item 1
- c. # Item 1
- d. 1. Item 1

Markdown tables



Markdown tables

This...

Column 1	Column 2
Cell 1, 1	Cell 2, 1
Cell 1, 2	Cell 2, 2

...produces this

Column 1	Column 2
Cell 1, 1	Cell 2, 1
Cell 1, 2	Cell 2, 2

Markdown tables

This...

Time	Session	Topic
<i>_left_</i>	<i>_center_</i>	<i>_right_</i>
01:00 - 01:50	1	Practice
01:50 - 02:00		**Break**
02:00 - 02:45	2	Class
02:45 - 03:00		**Break**

...produces this

Time	Session	Topic
<i>left</i>	<i>center</i>	<i>right</i>
01:00 - 01:50	1	Practice
01:50 - 02:00		Break
02:00 - 02:45	2	Class
02:45 - 03:00		Break

Break

05:00

Week 14: *Reproducible Reporting*

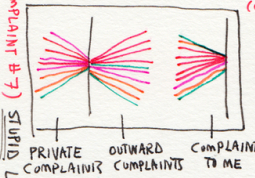
1. Why RMarkdown?
2. Metadata and output formats
3. Text
4. Code chunks

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS*

AND GENERAL GRUMPINESS

HOW TO READ IT: (I ~~THREW~~ TREW DOWN MY PENS WHEN I FINISHED)



TYPE OF COMPLAINT:

- WEATHER
- HUSBAND
- ANIMALS
- FAMILY
- SOCIETY!
- THE WORLD TODAY
- ACQUAINTANCES!
- STRANGERS
- MY APPEARANCE
- FRIENDS
- WORK
- HEALTH
- HUNGER
- MYSELF
- TECHNOLOGY/MEDIA
- MONEY
- INANIMATE OBJECTS
- TRANSPORT

ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... ARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH BOTHERS ME!

MAIN STATS	
PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE WHO COMPLAINED TO:	MY DAD
COMPLAINED TO:	0
* SECOND MOST COMPLAINTS ABOUT MY HUSBAND (SMART STEVE I LOVE YOU!)	

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!
 * AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

FROM:
S POSAVEC
LONDON
UK



TO:
GIORGIA LUPI

BROOKLYN, NY 11249
USA

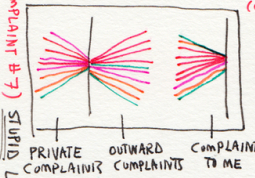
DELIVERED BY
HAND (SPECIAL NYC DELIVERY!)

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS*

AND GENERAL GRUMPINESS

HOW TO READ IT: (I ~~THREW~~ TREW DOWN MY PENS WHEN FINISHED)



(complaint #7) STUPID LEFT HAND SHOWER GEL (complaint #1) MESSING UP THE ORDER

ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... ARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT SYMMETRICAL WHICH BOTHERS ME!

TYPE OF COMPLAINT:

- WEATHER
- HUSBAND
- ANIMALS
- FAMILY
- SOCIETY!
- THE WORLD TODAY
- ACQUAINTANCES!
- STRANGERS
- MY APPEARANCE
- FRIENDS
- WORK
- HEALTH
- HUNGER
- MYSELF
- TECHNOLOGY/MEDIA
- MONEY
- INANIMATE OBJECTS
- TRANSPORT

MAIN STATS

PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE I COMPLAINED TO:	10
* SECOND MOST COMPLAINTS ABOUT:	MY HUSBAND

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!
 * AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

FROM: S POSAVEC
 LONDON
 UK



TO:
GEORGIA LUPI
 BROOKLYN, NY 11249
 USA

DELIVERED BY
 HAND (SPECIAL NYC DELIVERY!)

↑ metadata

text

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS* AND GENERAL GRUMPINESS

HOW TO READ IT: (I TREW DOWN MY PENS WHEN I FINISHED ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... AARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT METRICAL WHICH IS SHITTY

TYPE OF COMPLAINT:

WEATHER	HEALTH
HUSBAND	HUNGER
ANIMALS	MYSELF
FAMILY	TECHNOLOGY/MEDIA
SOCIETY!	MONEY
THE WORLD TODAY	INANIMATE OBJECTS
ACQUAINTANCES/STRANGERS	TRANSPORT
MY APPEARANCE	
FRIENDS	
WORK	

MAIN STATS

PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE I COMPLAINED TO:	10
* SECOND MOST COMPLAINTS ABOUT:	MY HUSBAND

FROM: S POSAVEC
LONDON
UK

TO: GEORGIA LUPI
BROOKLYN, NY 11249
USA

DELIVERED BY HAND (SPECIAL NYC DELIVERY!)

97P
UK

metadata

<https://www.dear-data.com/>

plots

text

DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS* AND GENERAL GRUMPINESS

HOW TO READ IT: (I TREW DOWN MY PENS WHEN I FINISHED ALL COMPLAINTS IN CHRONOLOGICAL ORDER EXCEPT 'COMPLAINTS TO ME' BECAUSE (COMPLAINT #1) I MESSED UP THE ORDER + BY THAT TIME I HAD ALREADY SPENT 1.5 HOURS ON THIS... AARGH! AND (COMPLAINT #2) THIS PART OF THE DRAWING IS NOT METRICAL WHICH BEEC...

TYPE OF COMPLAINT:

WEATHER	HEALTH
HUSBAND	HUNGER
ANIMALS	MYSELF
FAMILY	TECHNOLOGY/MEDIA
SOCIETY!	MONEY
THE WORLD TODAY	INANIMATE OBJECTS
ACQUAINTANCES/STRANGERS	TRANSPORT
MY APPEARANCE	
FRIENDS	
WORK	

MAIN STATS

PRIVATE COMPLAINTS:	67
OUTWARD COMPLAINTS:	100
COMPLAINTS TO ME:	43
# OF PEOPLE WHO COMPLAINED TO ME:	5
# OF PEOPLE I COMPLAINED TO:	10
* SECOND MOST COMPLAINTS ABOUT:	MY HUSBAND

FROM: S POSAVEC
LONDON
UK

TO: GEORGIA LUPI
BROOKLYN, NY 11249
USA

DELIVERED BY HAND (SPECIAL NYC DELIVERY!)

* THIS PEN (COMPLAINT #3) LEAKED + SMUDGED MY DRAWING + GOT ALL OVER MY HANDS!
* AND A WEEK OF COMPLAINTS ABOUT HOW I F*CKED UP THIS DRAWING! (COMPLAINT #5)

metadata

R Code

Inline code

```
`r insert code here`
```

Code chunks

```
```${r}  
insert code here
insert more code here
```
```

Inline R code

Embed R code directly in a markdown

```
`r <insert code here>`
```

For example:

```
The sum of 3 and 4 is `r 3 + 4`
```

Produces this:

The sum of 3 and 4 is 7

R Code chunks

```
```{r}  
bears %>%
 count(month)
```
```

What is the fate of this chunk?

```
bears %>%  
  count(month)
```

```
#> # A tibble: 12 x 2  
#>   month     n  
#>   <dbl> <int>  
#> 1     1     3  
#> 2     2     1  
#> 3     3     1  
#> 4     4     4  
#> 5     5    18  
#> 6     6    20  
#> 7     7    27  
#> 8     8    28  
#> 9     9    25  
#> 10    10    25  
#> 11    11    12  
#> 12    12     2
```


Code chunks

```
```{r}  
monthlyCount <- bears %>%
 count(month)
```
```

```
monthlyCount <- bears %>%  
  count(month)
```

What fate do you predict here?

Code chunks

```
```{r}
monthlyCount <- bears %>%
 count(month)
monthlyCount
```
```

```
monthlyCount <- bears %>%
  count(month)
monthlyCount
```

```
#> # A tibble: 12 x 2
#>   month     n
#>   <dbl> <int>
#> 1     1     3
#> 2     2     1
#> 3     3     1
#> 4     4     4
#> 5     5    18
#> 6     6    20
#> 7     7    27
#> 8     8    28
#> 9     9    25
#> 10    10    25
#> 11    11    12
#> 12    12     2
```

Chunk options

Control what chunks output using options inside `{r}`:

Example: `{r, echo=FALSE, message=FALSE}`

| option | default | effect |
|-------------------------|----------|---|
| <code>eval</code> | TRUE | Whether to evaluate the code and include its results |
| <code>echo</code> | TRUE | Whether to display code along with its results |
| <code>warning</code> | TRUE | Whether to display warnings |
| <code>error</code> | FALSE | Whether to display errors |
| <code>message</code> | TRUE | Whether to display messages |
| <code>tidy</code> | FALSE | Whether to reformat code in a tidy way when displaying it |
| <code>results</code> | "markup" | "markup", "asis", "hold", or "hide" |
| <code>cache</code> | FALSE | Whether to cache results for future renders |
| <code>comment</code> | "###" | Comment character to preface results with |
| <code>fig.width</code> | 7 | Width in inches for plots created in chunk |
| <code>fig.height</code> | 7 | Height in inches for plots created in chunk |

Chunk options

By default, code chunks print **code** + **output**:

This...

```
```{r}
cat('hello world!')
```
```

...produces this

```
cat('hello world!')
```

```
#> hello world!
```

Chunk output options

```
```{r, echo=FALSE}  
cat('hello world!')
```
```

Prints only **output**
(doesn't show code)

Output:

```
#> hello world!
```

```
```{r, eval=FALSE}  
cat('hello world!')
```
```

Prints only **code**
(doesn't run the code)

Output:

```
cat('hello world!')
```

```
```{r, include=FALSE}  
cat('hello world!')
```
```

Runs, but doesn't print
anything

Output:

message / warning

```
>
```

message / warning

```
```{r, message=FALSE, warning=FALSE}  
library(tidyverse)
```
```

```
library(tidyverse)
```

Using chunk options

```
`` `{r, message=FALSE, warning=FALSE}  
library(tidyverse)  
```
```

- Place between curly braces  
`{r option=value}`
- Multiple options separated by commas  
`{r option1=value, option2=value}`
- Careful! The `r` part is the **code engine** (other engines possible)



# Inserting a Python code chunk

Change `{r}` to `{python}` in the code chunk. Example:

```
'In Python, you can concatenate strings' + ' like this!'
```

```
#> 'In Python, you can concatenate strings like this!'
```

# A global `setup` chunk

One chunk to rule them all!

```
```{r setup, include = FALSE}
knitr::opts_chunk$set(
  warning = FALSE,
  message = FALSE,
  comment = "#>",
  fig.retina = 3,
  fig.path = "figs/"
)
```
```

- A special chunk label: `setup`
- Typically the first chunk
- All following chunks will use these options (i.e., sets global chunk options)
- **Tip:** set `include=FALSE`
- You can (and should) use individual chunk options too

# Think pair share 1: Birds & Bears

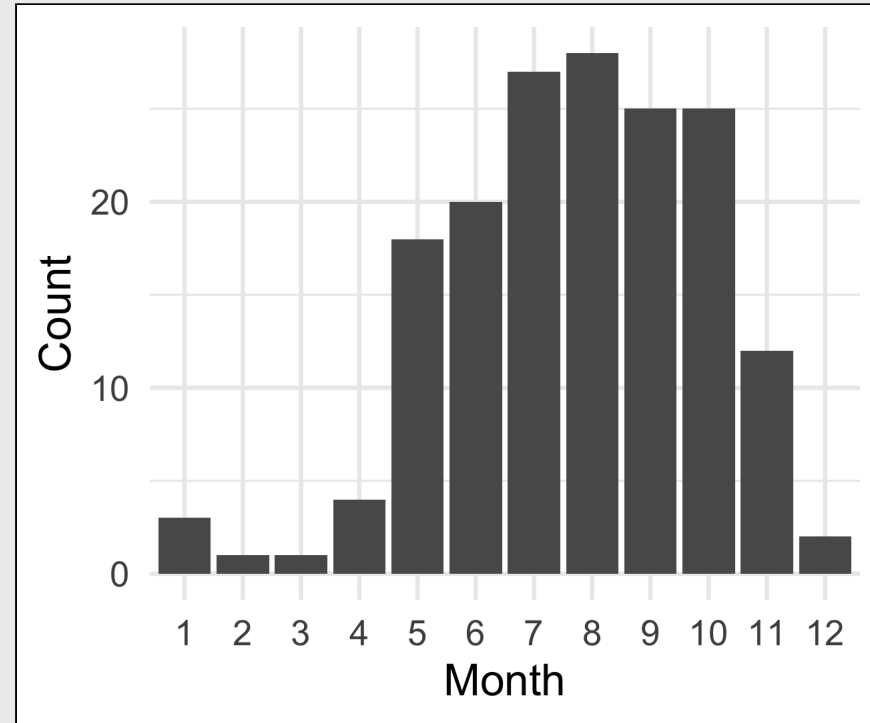
- 1) Create a new R Markdown file (`.Rmd`) in RStudio - title it *"Birds and Bears Analysis"*
- 2) Create a "setup" code chunk to load the `tidyverse` library and the `birds.csv` and `bears.csv` files.
- 3) Use text and code to find answers each of the following questions - show your code and results to justify each answer:
  - Which months have the highest and lowest number of bird impacts with aircraft?
  - Does the annual number of bird impacts appear to be changing over time?
  - Which months have the highest frequency of bear killings?
  - Who has been killed more often by bears: hunters or hikers?
  - How do the the number of bear attacks on men vs women compare?

Including plots

# Including plots

```
bears %>%
 count(month) %>%
 ggplot() +
 geom_col(
 aes(x = as.factor(month), y=n)) +
 theme_minimal(base_size = 22) +
 labs(x = 'Month', y = 'Count')
```

Will this print?



# Including plots

```
bearMonthPlot <- bears %>%
 count(month) %>%
 ggplot() +
 geom_col(
 aes(x = as.factor(month), y=n)) +
 theme_minimal(base_size = 22) +
 labs(x = 'Month', y = 'Count')
```

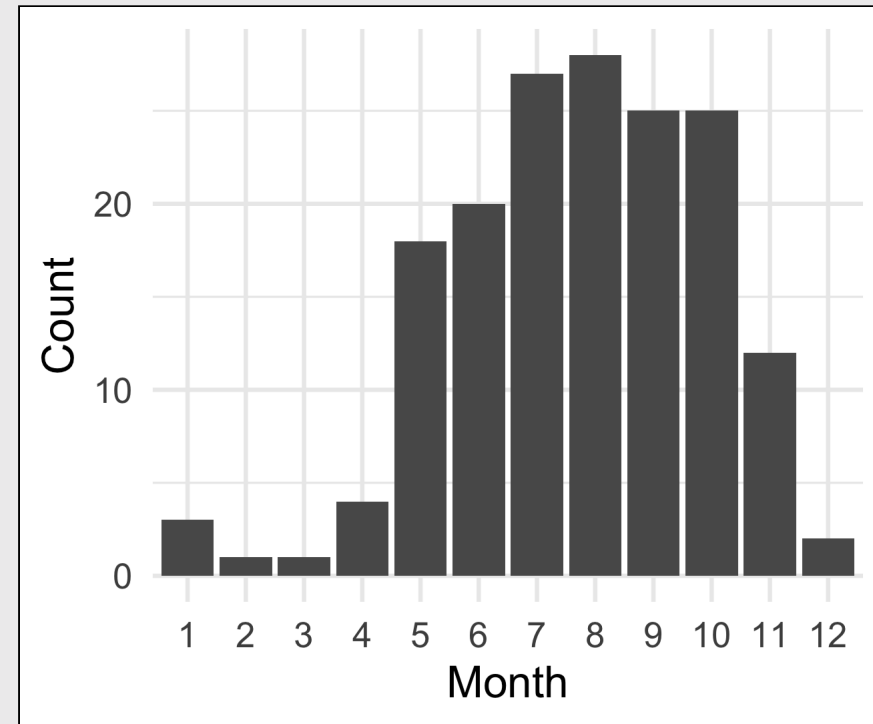
What about this?

# Including plots

```
bearMonthPlot <- bears %>%
 count(month) %>%
 ggplot() +
 geom_col(
 aes(x = as.factor(month), y=n)) +
 theme_minimal(base_size = 22) +
 labs(x = 'Month', y = 'Count')
```

bearMonthPlot

What about this?



# Chunk options for plots

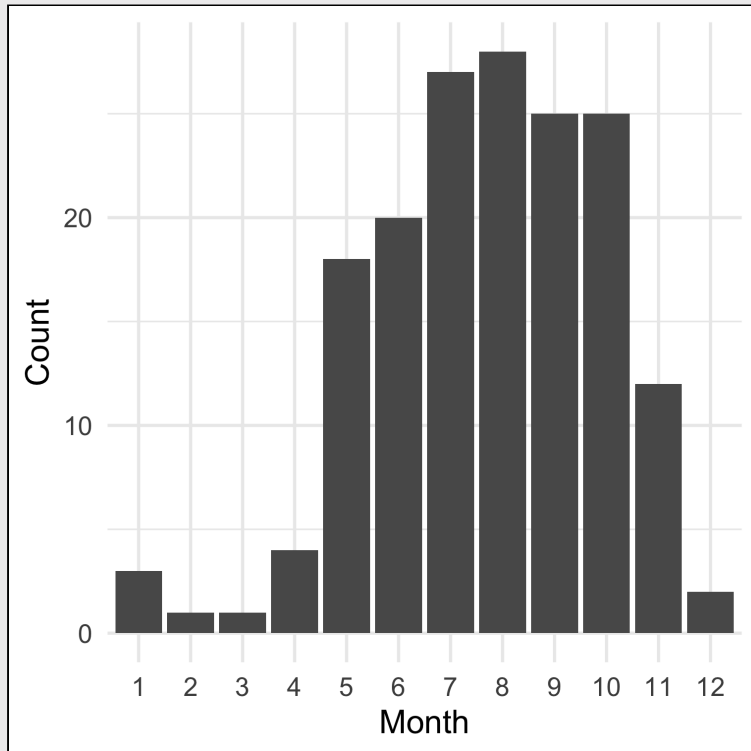
- fig size
- fig resolution

<https://yihui.name/knitr/options/#plots>

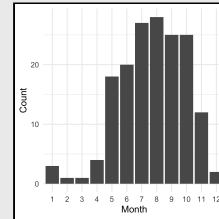


# out.width

```
```{r, out.width="70%"}  
bearMonthPlot  
```
```

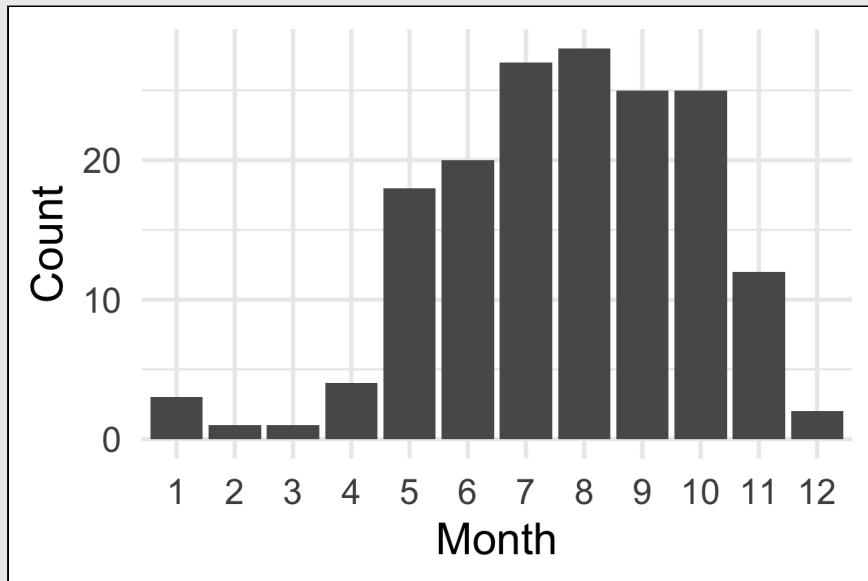


```
```{r, out.width="20%"}  
bearMonthPlot  
```
```

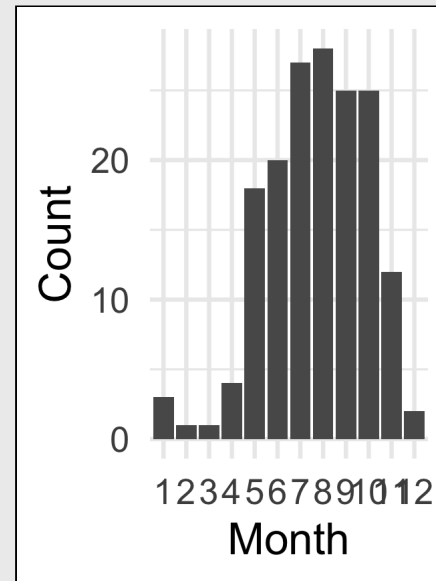


# fig.width & fig.height

```
```{r, fig.width=6, fig.height=4}  
bearMonthPlot  
```
```

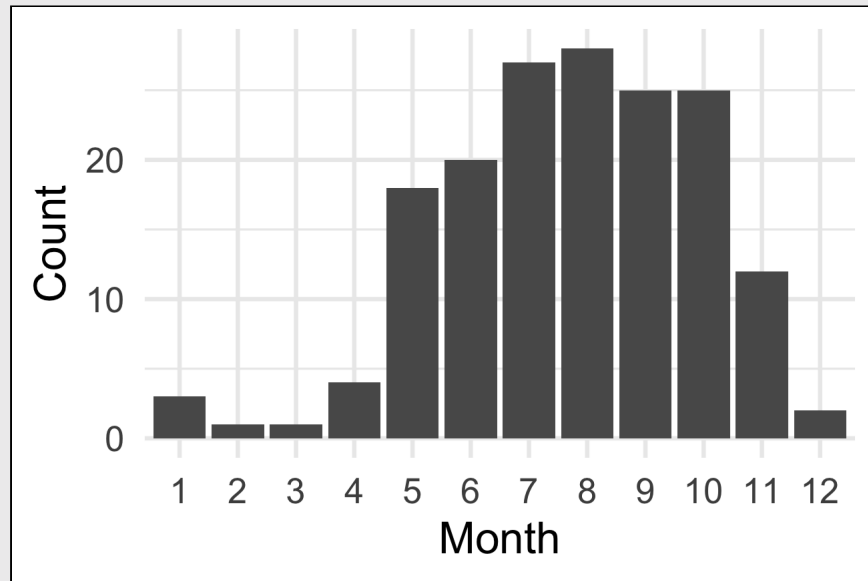


```
```{r, fig.width=3, fig.height=4}  
bearMonthPlot  
```
```



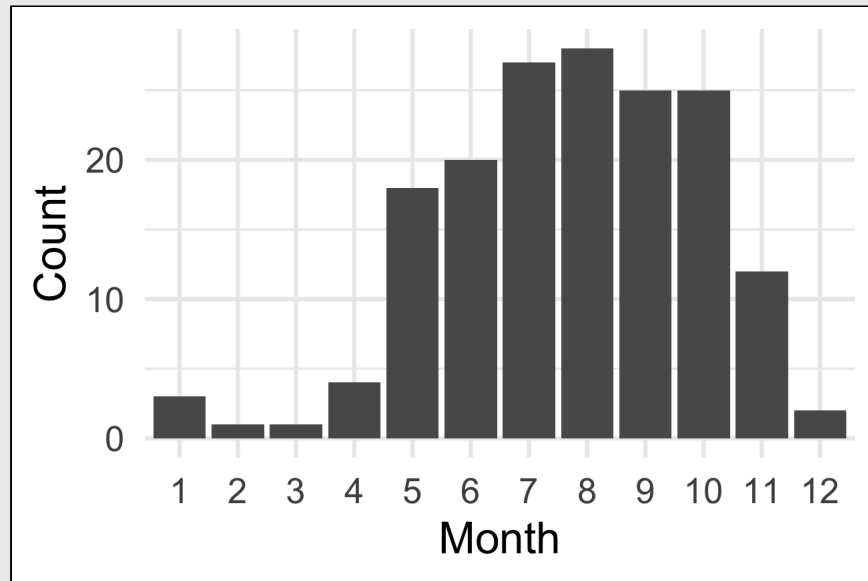
# fig.path

```
```{r, fig.path="figs/", echo=FALSE}  
bearMonthPlot  
```
```



# fig.path

```
```{r bear-month-plot, fig.path="figs/", echo=FALSE}  
bearMonthPlot  
```
```



# A good chunk label

**Think: kebabs, not snakes**

Good

`my-plot`

`myplot`

`myplot1`

`myplot-1`

`MY-PL0T`

Bad

`my_plot`

`my plot`

everything else!

# View default options

```
str(knitr::opts_chunk$get())
```

```
#> List of 53
#> $ eval : logi TRUE
#> $ echo : logi TRUE
#> $ results : chr "markup"
#> $ tidy : logi FALSE
#> $ tidy.opts : NULL
#> $ collapse : logi FALSE
#> $ prompt : logi FALSE
#> $ comment : chr "#>"
#> $ highlight : logi TRUE
#> $ strip.white : logi TRUE
#> $ size : chr "normalsize"
#> $ background : chr "#F7F7F7"
#> $ cache : logi FALSE
#> $ cache.path : chr "slides-14-reproducible-reporting_cache/html/"
#> $ cache.vars : NULL
#> $ cache.lazy : logi TRUE
#> $ dependson : NULL
#> $ autodep : logi FALSE
#> $ cache.rebuild: logi FALSE
#> $ fig.keep : chr "high"
#> $ fig.show : chr "asis"
#> $ fig.align : chr "default"
#> $ fig.path : chr "slides-14-reproducible-reporting_files/figure-html/"
#> $ dev : chr "png"
```

## Two more important chunks:

- Images
- Tables

# Image chunks

Insert images with markdown

```

```

Insert images with chunks (so you can resize it)

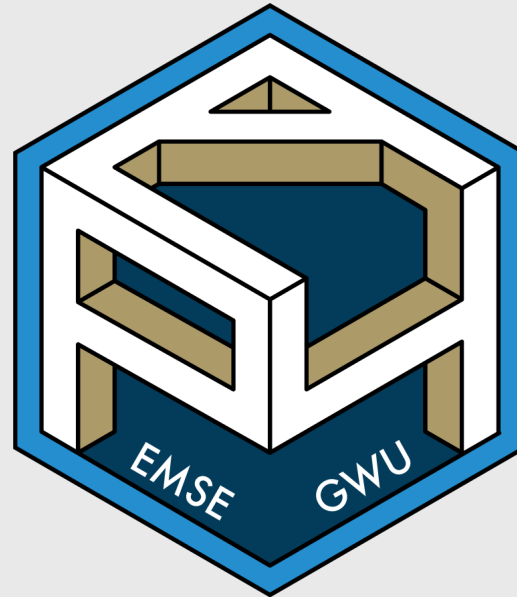
```
```{r, echo=FALSE, out.width="20%"}  
knitr::include_graphics("images/p4a_hex_sticker.png")  
```
```



# Image chunks

```
```{r, echo=FALSE, out.width="20%"}  
knitr::include_graphics("images/p4a_hex")  
```
```

```
```{r, echo=FALSE, out.width="50%"}  
knitr::include_graphics("images/p4a_hex")  
```
```



# Convert a data frame to a table with `kable()`

```
bears %>%
 count(bearType, wildOrCaptive)
```

```
#> # A tibble: 6 x 3
#> bearType wildOrCaptive n
#> <chr> <chr> <int>
#> 1 Black Captive 16
#> 2 Black Wild 60
#> 3 Brown Captive 8
#> 4 Brown Wild 72
#> 5 Polar Captive 4
#> 6 Polar Wild 6
```

```
bears %>%
 count(bearType, wildOrCaptive) %>%
 kable()
```

| <b>bearType</b> | <b>wildOrCaptive</b> | <b>n</b> |
|-----------------|----------------------|----------|
| Black           | Captive              | 16       |
| Black           | Wild                 | 60       |
| Brown           | Captive              | 8        |
| Brown           | Wild                 | 72       |
| Polar           | Captive              | 4        |
| Polar           | Wild                 | 6        |

# Think pair share 2: College Majors

- 1) Create a new R Markdown file (`.Rmd`) in RStudio - title it *"College Majors Analysis"*
- 2) Create a "setup" code chunk to load the `tidyverse` library and the `recent_grads.csv` file.
- 3) Use text, code, and plots to find answers each of the following questions - show your code and results to justify each answer:
  - What are the highest earning engineering majors?
  - Within the engineering majors, which ones have better employment rates?
  - Within the engineering majors, which ones have a better gender balance?

(Use good code chunk names for your figures!)