



Week 14: *Reproducible Reporting*

EMSE 4574: Intro to Programming for Analytics

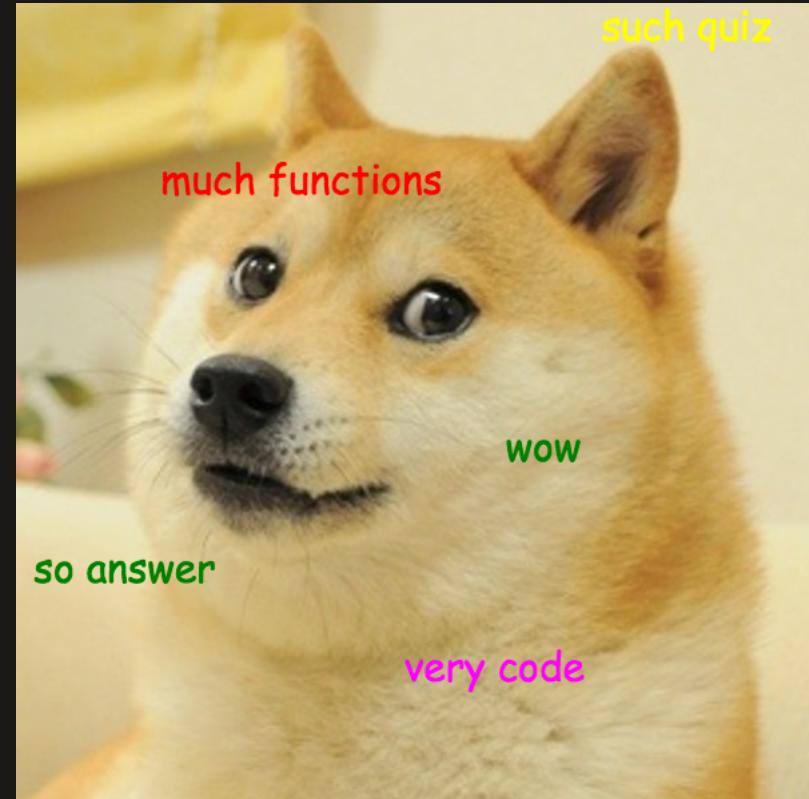
John Paul Helveston

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

Rmarkdown

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



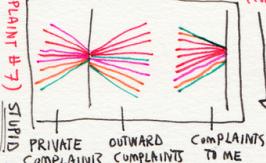
Art by [Allison Horst](#)



DEAR DATA - WEEK 07

A WEEK OF COMPLAINTS*

AND GENERAL GRUMPINESS
HOW TO READ IT: (I ~~THREW DOWN MY PENS WHEN I FINISHED~~)
(COMPLAINT #7) WHAT IS WRONG W/MG??



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	

(#1) THIS PEN (COMPLAINT #7) LEAKED + SMUDGED
MY DRAWING + GOT ALL OVER MY HANDS!
(My stupid LEFT HAND!)

FROM:
S POSAVEL

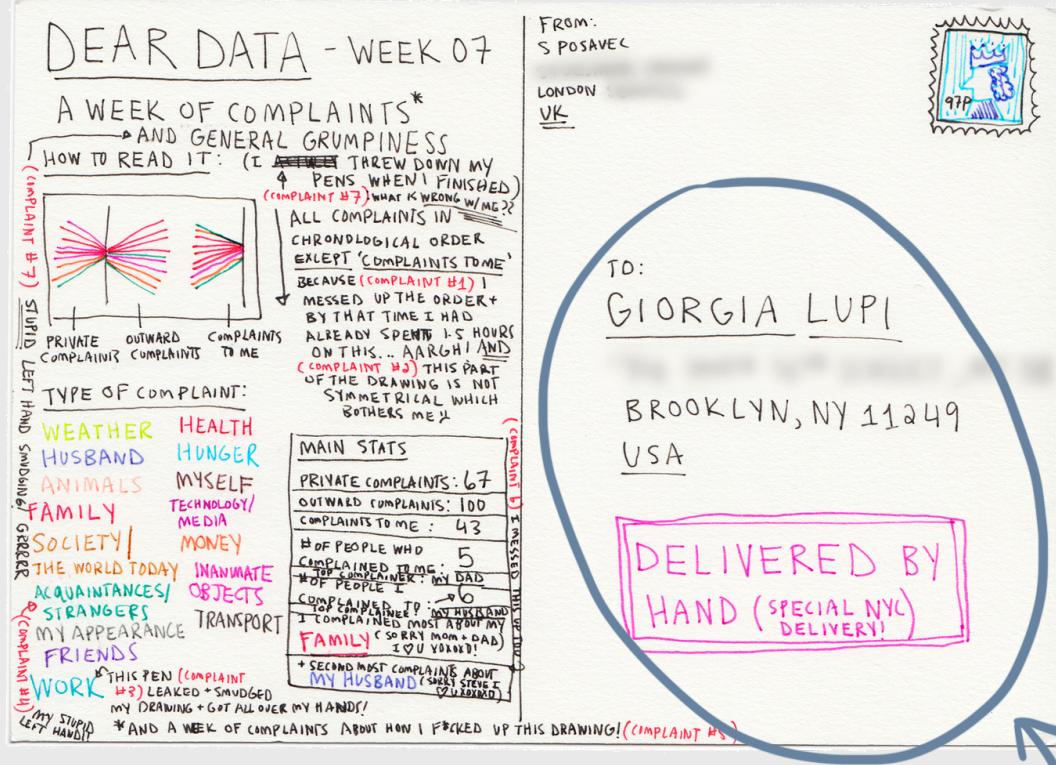
LONDON
UK



TO:
GIORGIA LUPI

BROOKLYN, NY 11249
USA

DELIVERED BY
HAND (SPECIAL NYC
DELIVERY!)



metadata

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

metadata: YAML

YAML
Ain't
Markup
Language

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



Output options

Save output options in your YAML

```
---
```

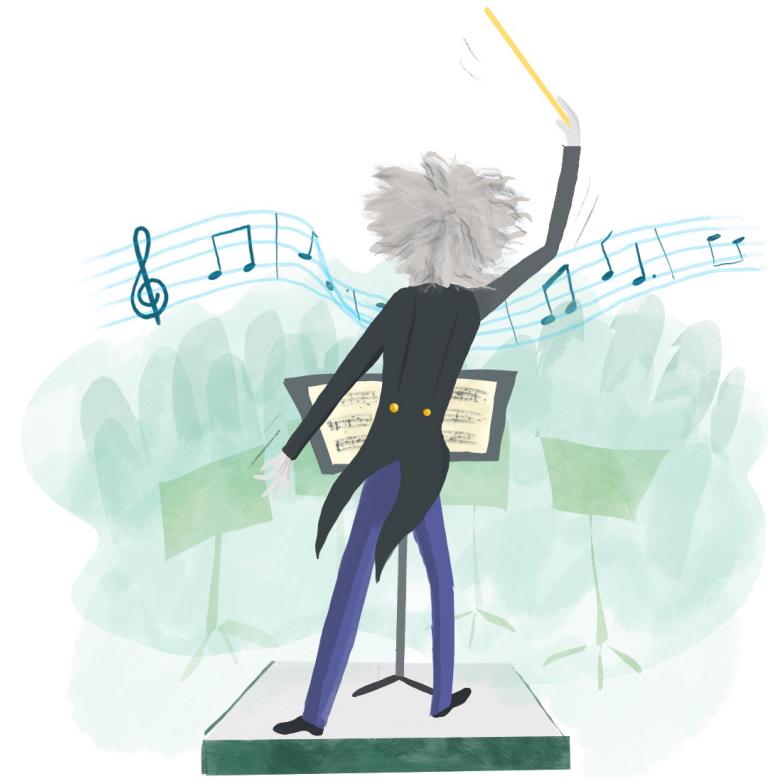
```
title: Your title here
author: Your name here
output: html_document
```

```
---
```

```
---
```

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

```
---
```



demo - RStudio

demo.Rmd

ABC Knit

Insert Run

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)(http://rmarkdown.rstudio.com).|
29
30 When you click the **Knit** button a document will be generated that includes

Console Jobs

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 R Markdown

Environment History Connections Tutorial

demo - RStudio

demo.Rmd

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)(http://rmarkdown.rstudio.com).|
29
30 When you click the **Knit** button a document will be generated that includes

17 / 87

metadat = YAML

Console Jobs ~/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, 10:36 AM
demo.Rmd	1.5 KB	Nov 30, 2020, 11:00 AM
notes.Rproj	205 B	Nov 30, 2020, 10:59 AM

demo - RStudio

demo.Rmd

ABC Knit

Insert Run

Console Jobs

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 R Markdown

Environment History Connections Tutorial

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)(http://rmarkdown.rstudio.com).|
29
30 When you click the **Knit** button a document will be generated that includes

metadat = YAML

demo - RStudio

demo.Rmd

ABC Knit

Insert Run

Console Jobs

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 R Markdown

Environment History Connections Tutorial

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)(http://rmarkdown.rstudio.com).|
29
30 When you click the **Knit** button a document will be generated that includes

metadat = YAML



QuickTime Player File Edit View Window Help

demo - RStudio

demo.Rmd x Addins x

ARC Knit Insert Run

```
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  
▶ 10:00 y/o:18 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

Environment History Connections Tutorial

A screenshot of a web browser window titled "A quick RMarkdown demo". The address bar shows the file path: "/Users/jhelvy/Downloads/demo/demo.html". The browser interface includes standard navigation buttons (back, forward, search) and a toolbar with various icons.

The main content area displays an R Markdown document:

- R Markdown** (highlighted in a dark blue bar)
- 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

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

R Markdown

Plots

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() +
```

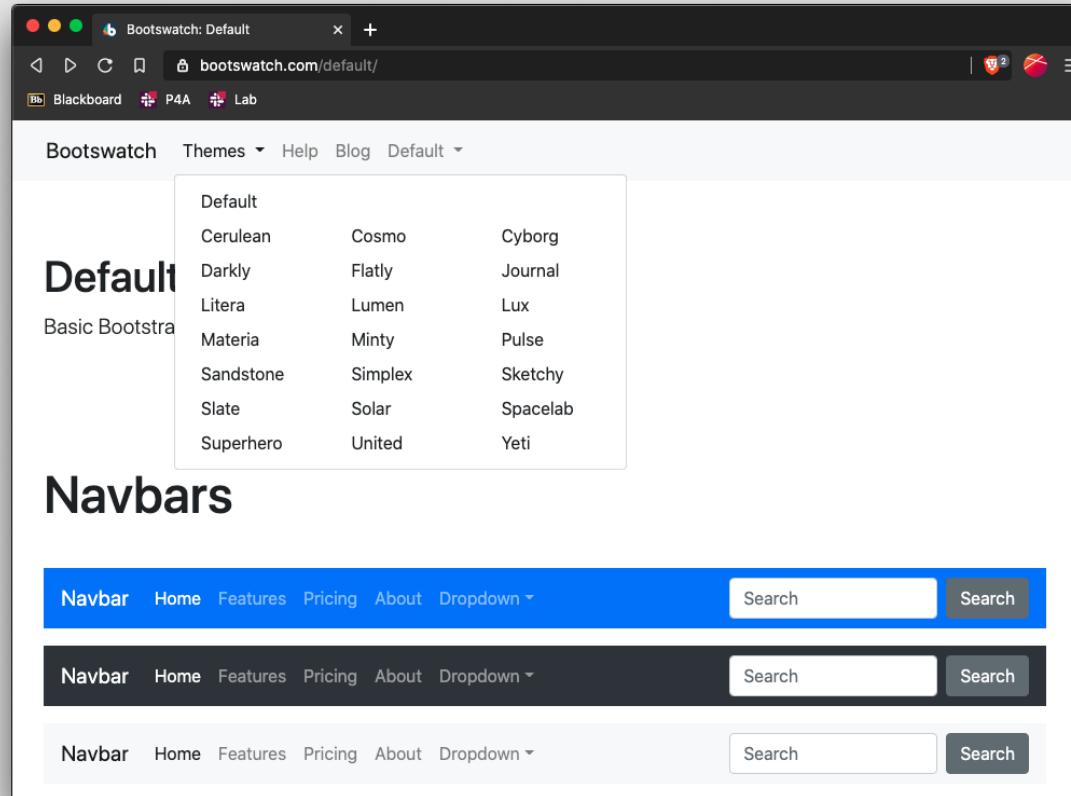
toc:
toc_float:

↑

font
+ = theme
color

color

Preview bootswatch themes



<https://bootswatch.com/>

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

The screenshot shows a web browser window displaying the rmarkdown documentation at rmarkdown.rstudio.com/docs/reference/index.html. The page title is "Function reference • rmarkdown". The navigation bar includes links for "Blackboard", "P4A", "Lab", "rmarkdown 2.5.3", "Get started", "Articles", "Reference", "Changelog", and "Home". The main content area is titled "Reference" and contains a section titled "Output formats". A paragraph explains that output formats can be specified in YAML frontmatter and lists various functions and their descriptions:

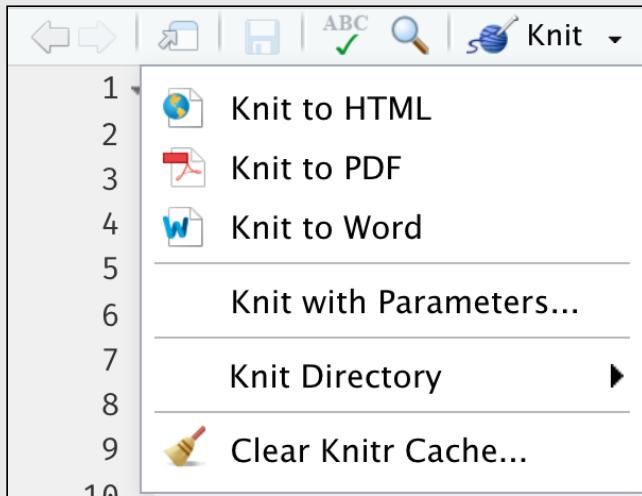
Function	Description
<code>beamer_presentation()</code>	Convert to a Beamer presentation
<code>ioslides_presentation()</code>	Convert to an ioslides Presentation
<code>powerpoint_presentation()</code>	Convert to a PowerPoint presentation
<code>slidy_presentation()</code>	Convert to a slidy presentation
<code>context_document()</code>	Convert to a ConTeXt document
<code>github_document()</code>	Convert to GitHub Flavored Markdown
<code>html_document()</code>	Convert to an HTML document
<code>md_document()</code>	Convert to a markdown document
<code>odt_document()</code>	Convert to an OpenDocument Text (ODT) document
<code>pdf_document() latex_document()</code>	Convert to a PDF/LaTeX document
<code>latex_fragment()</code>	
<code>rtf_document()</code>	Convert to an RTF document
<code>word_document()</code>	Convert to an MS Word document
<code>html_vignette()</code>	Convert to an HTML vignette
<code>tufte_handout()</code>	Tufte handout format (PDF)
<code>html_notebook()</code>	Convert to an HTML notebook
<code>html_fragment()</code>	Convert to an HTML fragment.

<https://rmarkdown.rstudio.com/docs/reference/index.html#section-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)

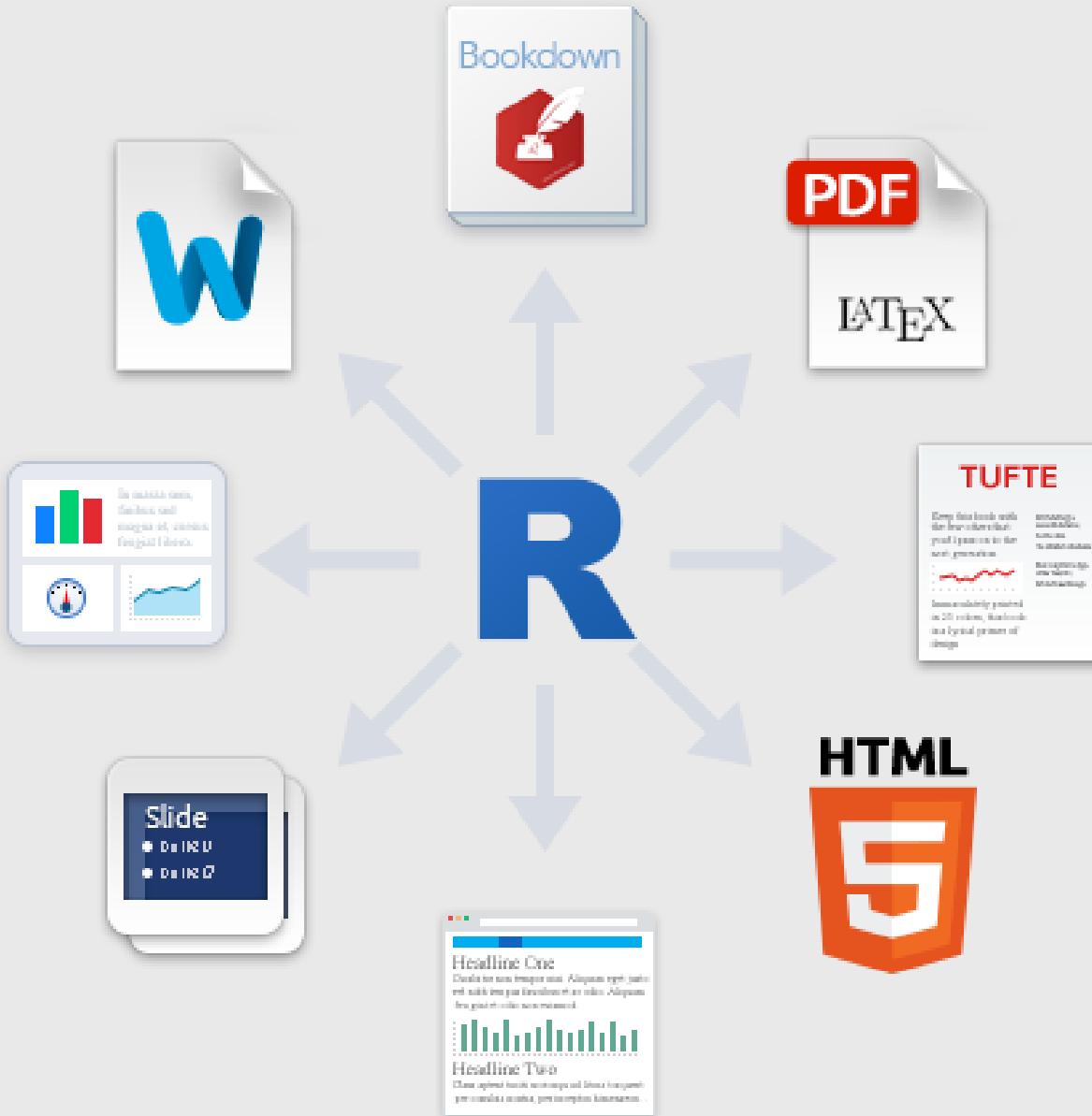
```
---  
output:  
  html_document:  
    toc: true  
  word_document: default  
---
```

(b)

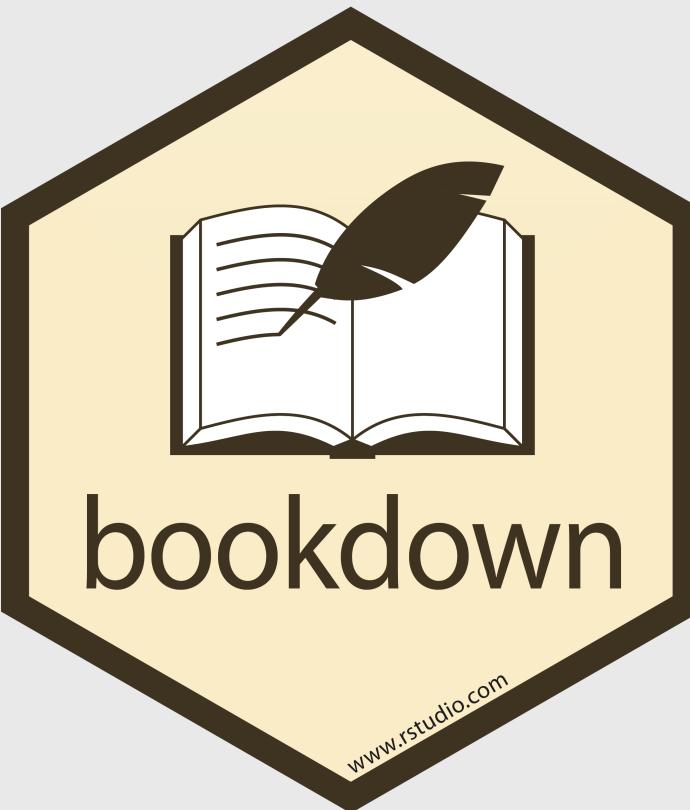
```
---  
output:  
  html_document(toc=true)  
  word_document(default)  
---
```

Built-in output formats





Extension output formats



The screenshot shows a web browser window with the title "Distill for R Markdown" and the URL "rstudio.github.io/distill/" in the address bar. The browser's toolbar includes icons for back, forward, search, and various extensions. Below the toolbar, the main content area displays the "Distill for R Markdown" page. The page features a large title "Distill for R Markdown" and a subtitle "Scientific and technical writing, native to the web". Below the subtitle, there is a section for "AUTHORS" listing JJ Allaire, Rich Iannone, Alison Presmanes Hill, and Yihui Xie, each with a green GitHub icon. To the right, under "AFFILIATIONS", it lists RStudio for all authors. Further down, under "PUBLISHED", it shows "Sept. 10, 2018" and under "CITATION", it lists "Allaire, et al., 2018". On the left side, there is a "Contents" sidebar with links to "Installation", "Creating an article", "Figures", "Tables", "Equations", and "Citations". The main content area also describes Distill as a web publishing format optimized for scientific and technical communication, mentioning its reader-friendly typography, LaTeX support, and flexible figure layout options.

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

<https://rstudio.github.io/distill/>

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`

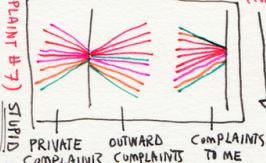
Week 14: *Reproducible Reporting*

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 DOWN MY PENS WHEN I FINISHED~~)
(COMPLAINT #7) WHAT IS WRONG W/MG??



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	

(#1) THIS PEN (COMPLAINT #7) LEAKED + SMUDGED
MY DRAWING + GOT ALL OVER MY HANDS!
(My stupid LEFT HAND!)

FROM:
S POSAVEL

LONDON
UK

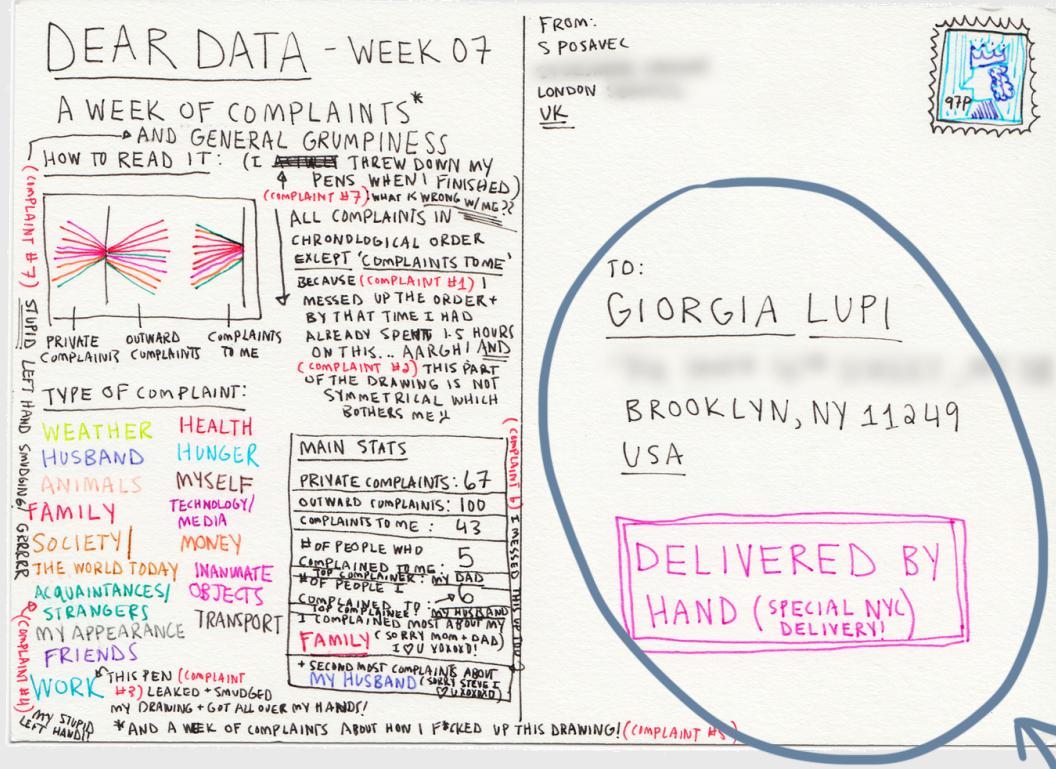


TO:
GIORGIA LUPI

BROOKLYN, NY 11249
USA

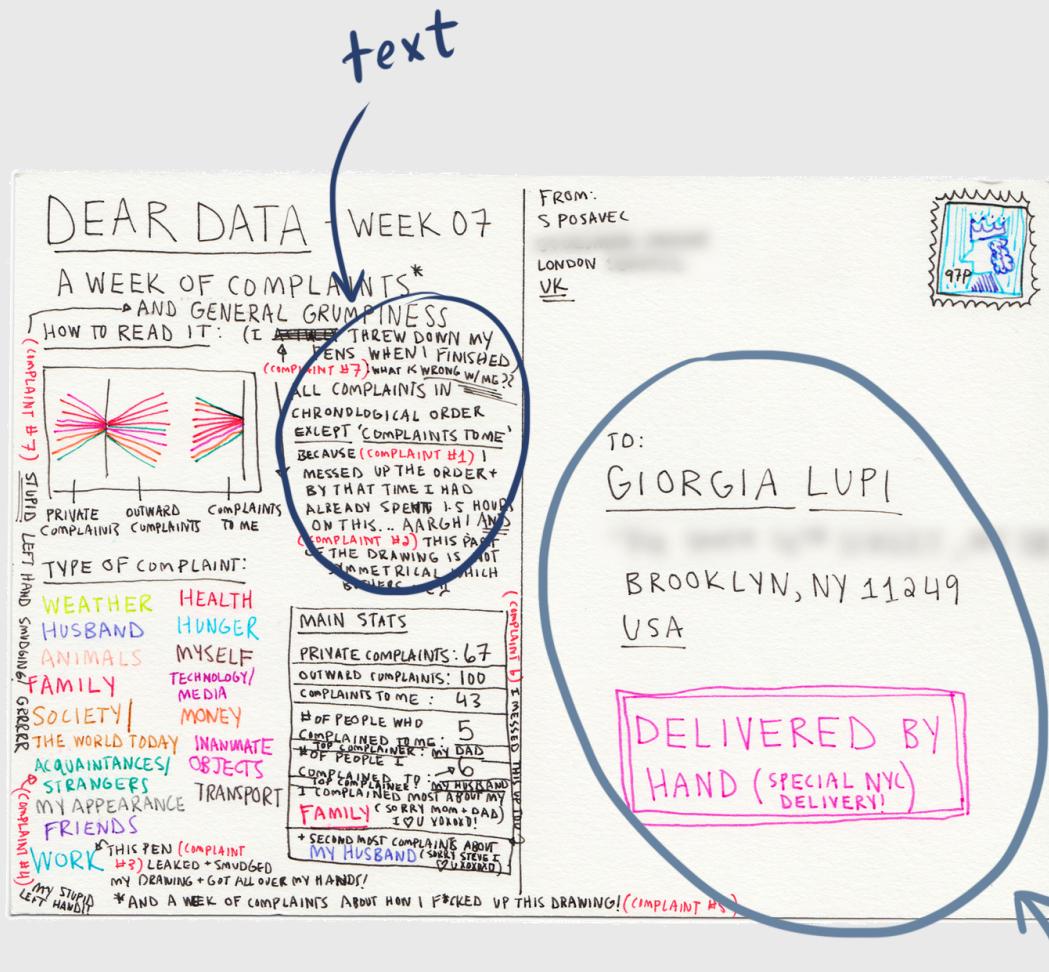
DELIVERED BY
HAND (SPECIAL NYC
DELIVERY!)

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



metadata

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



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:

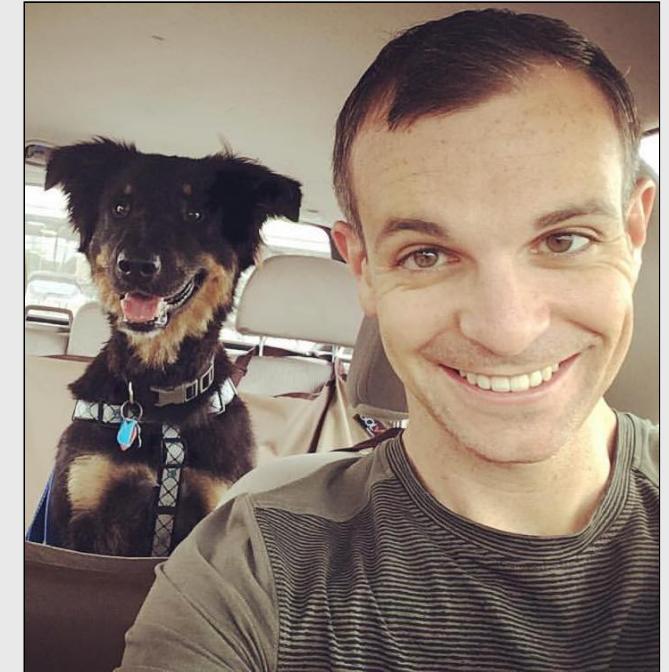
[Download R] (<http://www.r-project.org/>)

[Download R](#)

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?

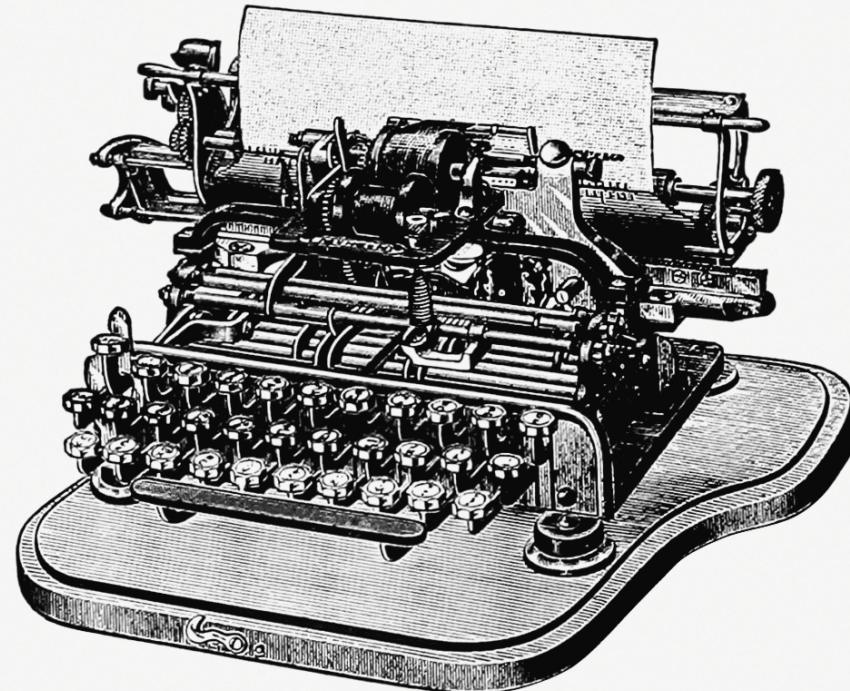
a. ! Item 1

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> 01:00 - 01:50	<i>center</i> 1	<i>right</i> 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

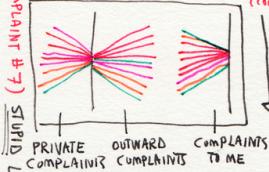
Week 14: *Reproducible Reporting*

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 DOWN MY PENS WHEN I FINISHED~~)
(COMPLAINT #7) WHAT IS WRONG W/MG??



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	

(#1) THIS PEN (COMPLAINT #7) LEAKED + SMUDGED
MY DRAWING + GOT ALL OVER MY HANDS!
(My stupid LEFT HAND!)

FROM:
S POSAVEL

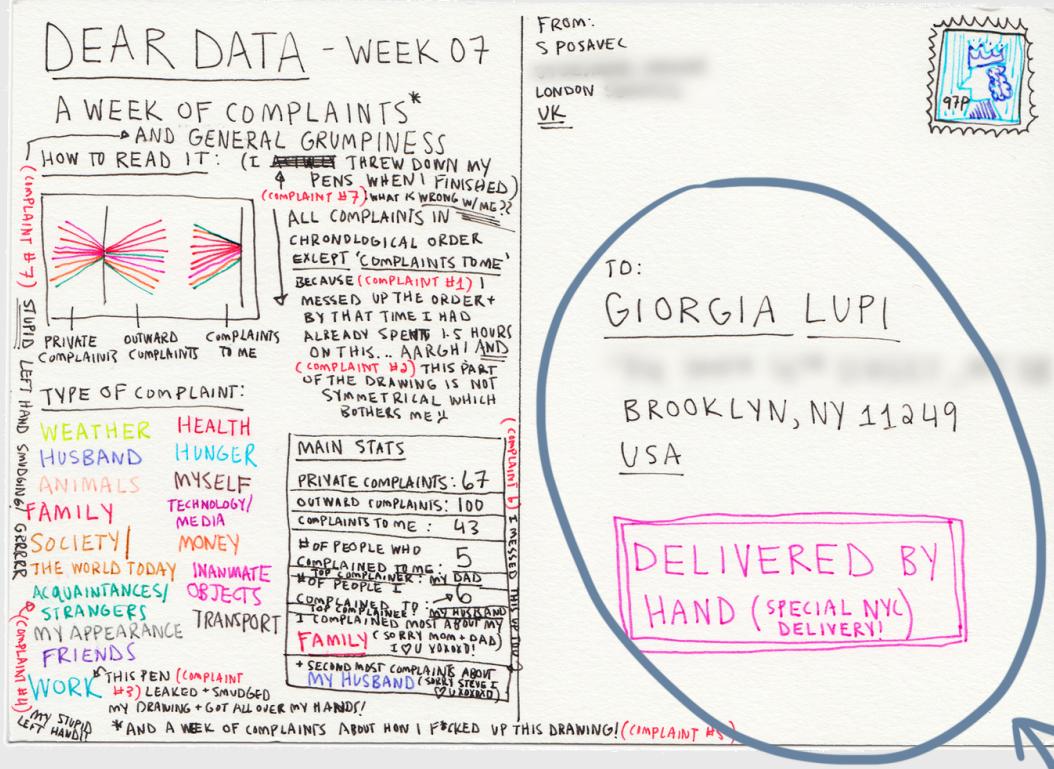
LONDON
UK



TO:
GIORGIA LUPI

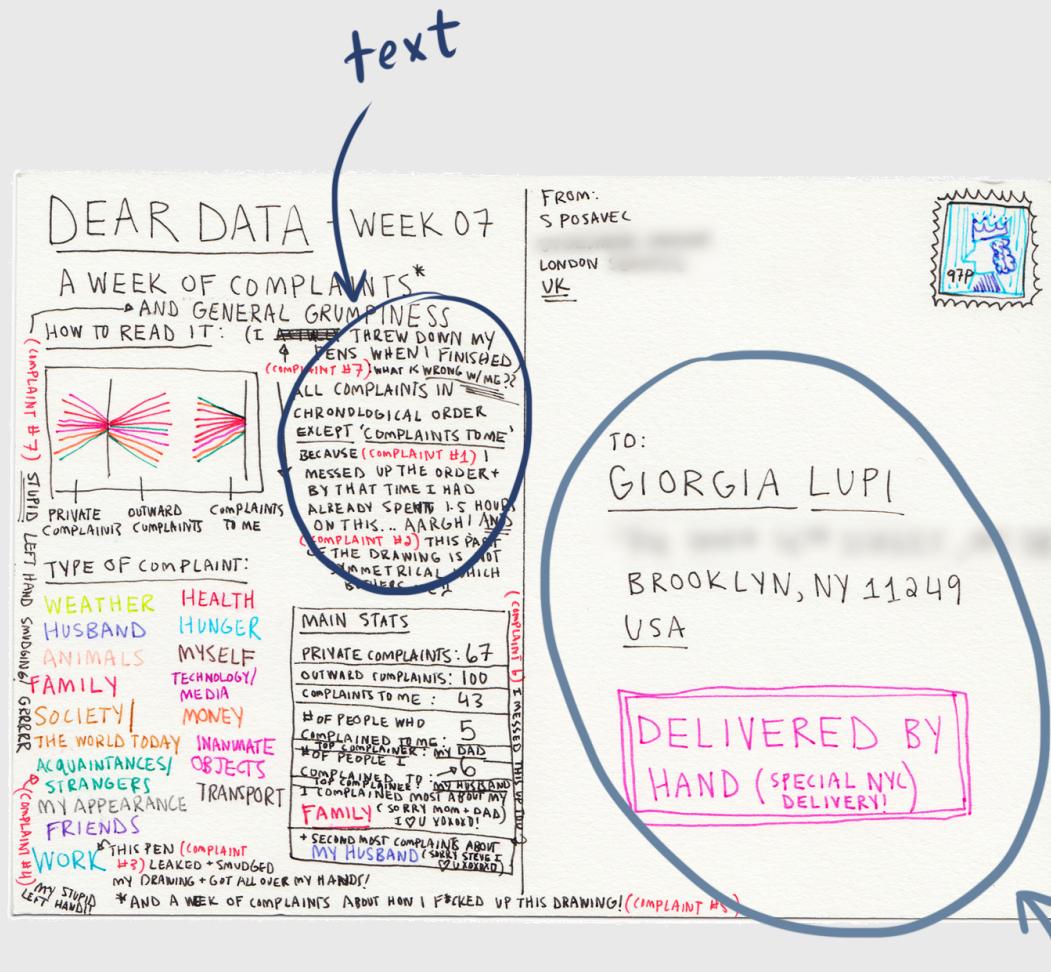
BROOKLYN, NY 11249
USA

DELIVERED BY
HAND (SPECIAL NYC
DELIVERY!)



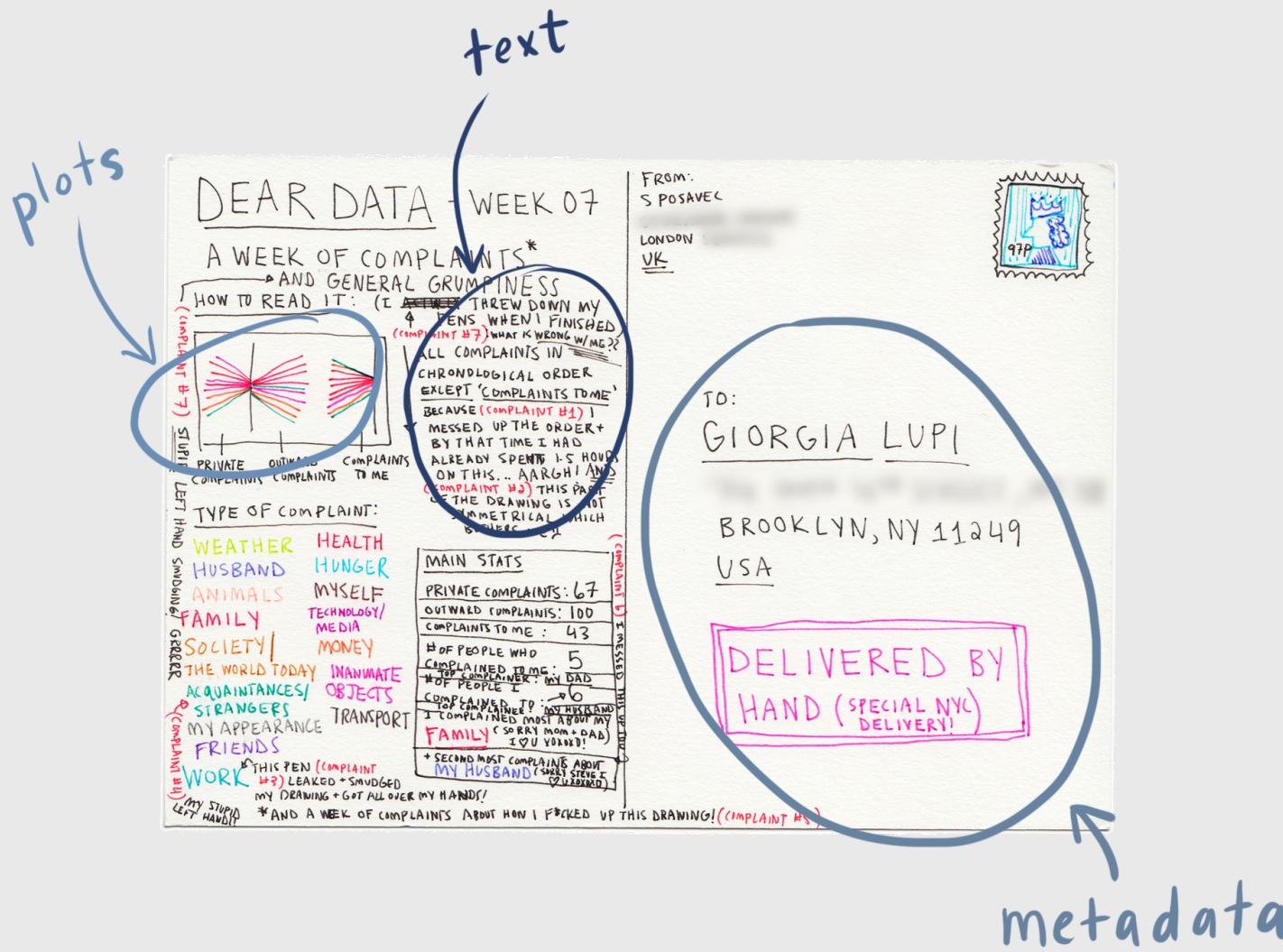
metadata

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



metadata

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



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

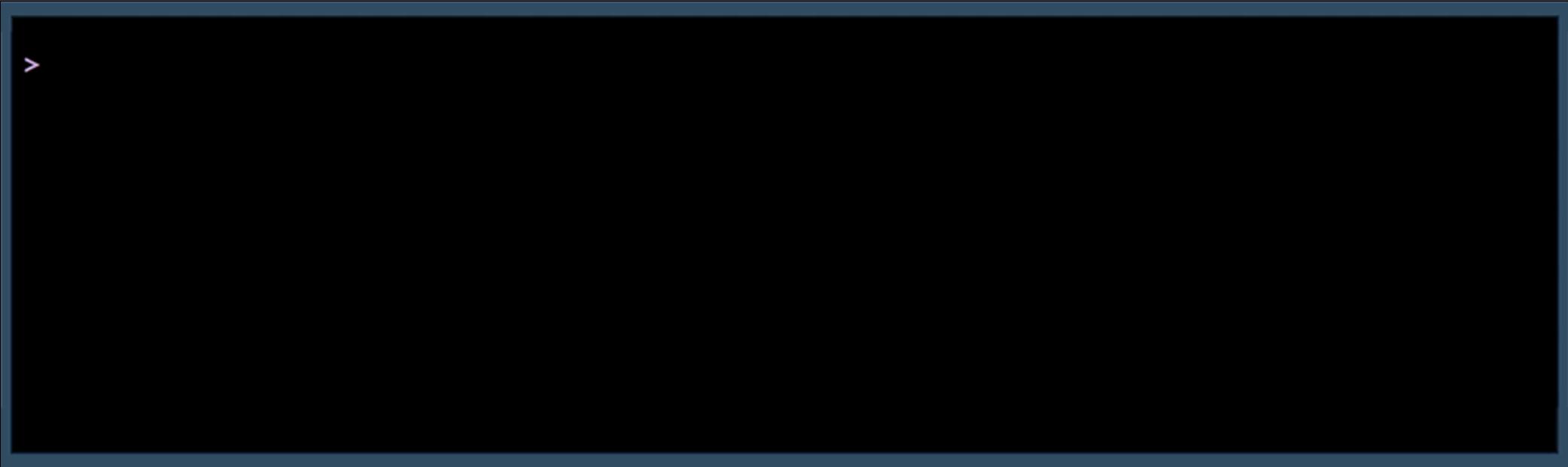
```
cat('hello world!')
```

```
```{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

15:00

Think pair share 1: Birds & Bears

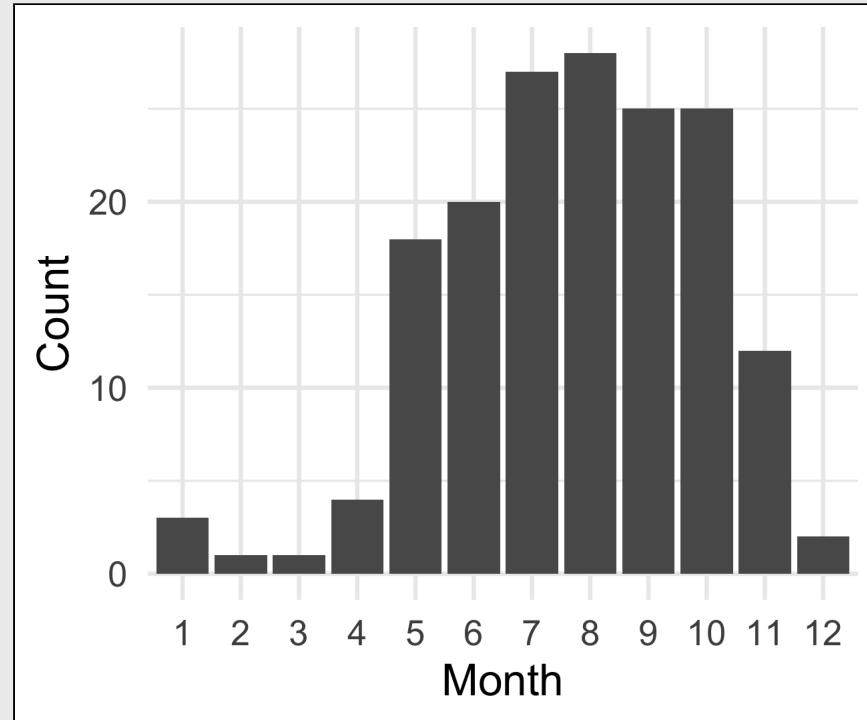
- 1) Create a new R Markdown file (**.Rmd**) in RStudio - title it "*Birds and Bears Analysis*"
- 2) Create a "setup" code chunk to load the **tidyverse** library and the **birds.csv** and **bears.csv** files.
- 3) Use text and code to find answers each of the following questions - show your code and 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

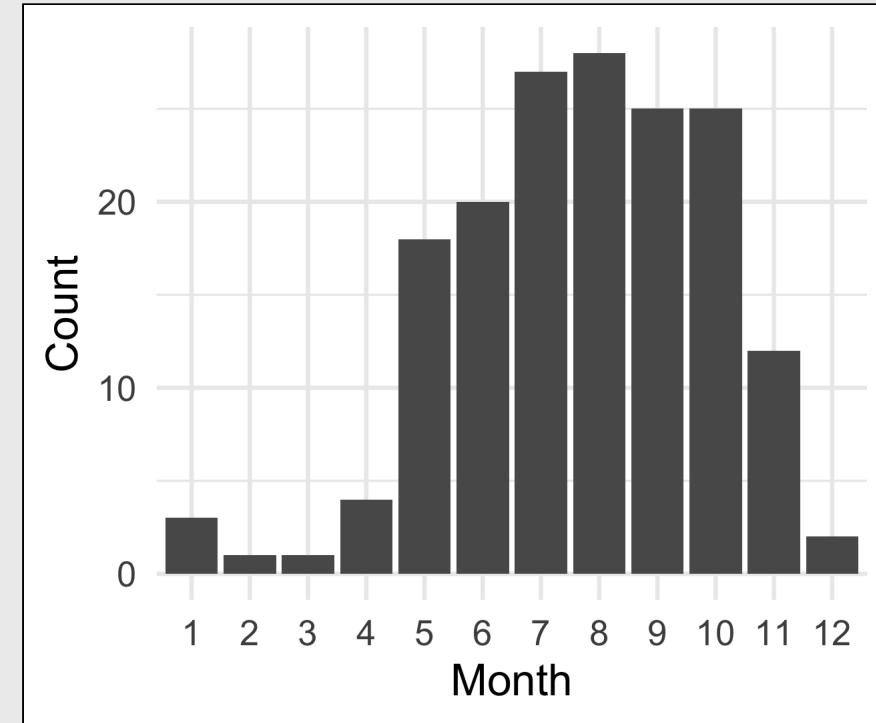
```
bearMonthPlot <- bears %>%  
  count(month) %>%  
  ggplot() +  
  geom_col(  
    aes(x = as.factor(month), y=n)) +  
  theme_minimal(base_size = 22) +  
  labs(x = 'Month', y = 'Count')
```

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?



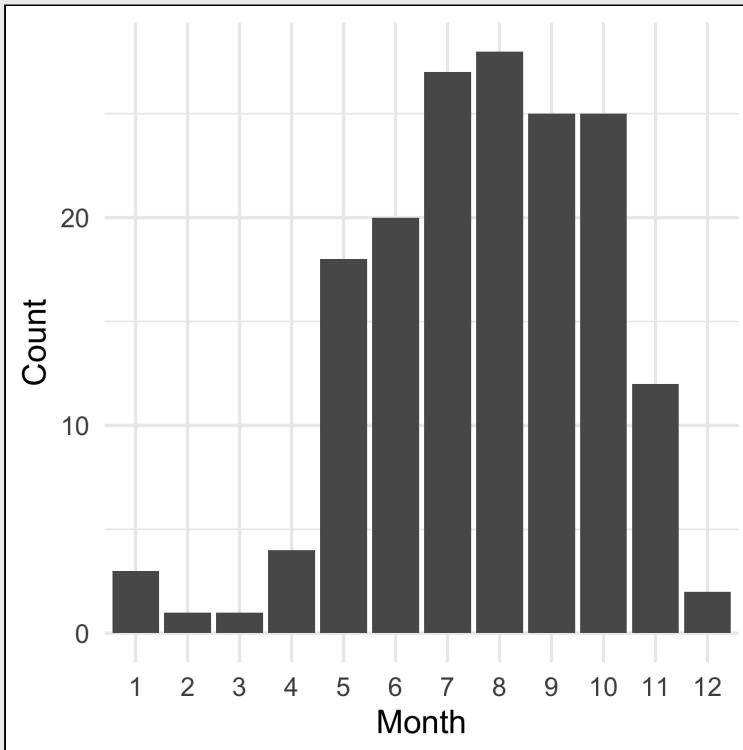
Chunk options for plots

- fig size
- fig resolution

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

out.width

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



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

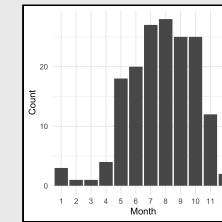
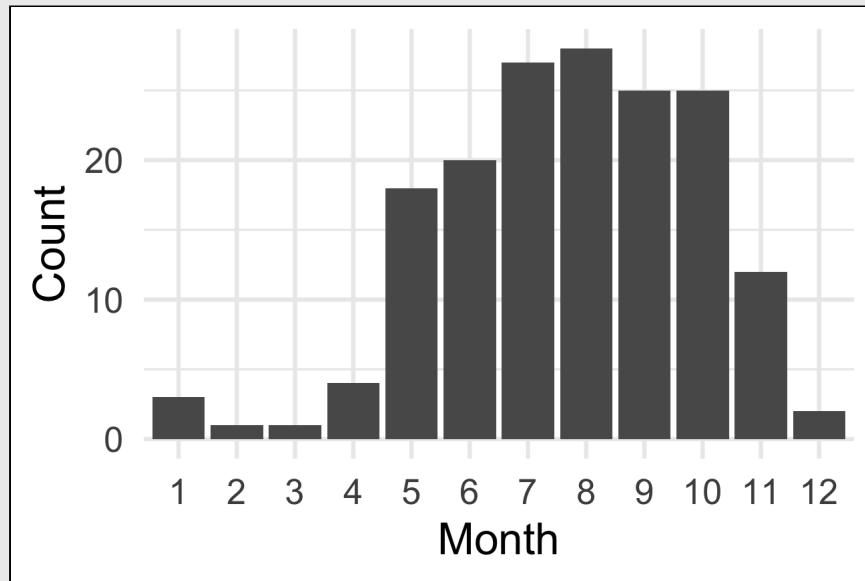


fig.width & fig.height

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



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

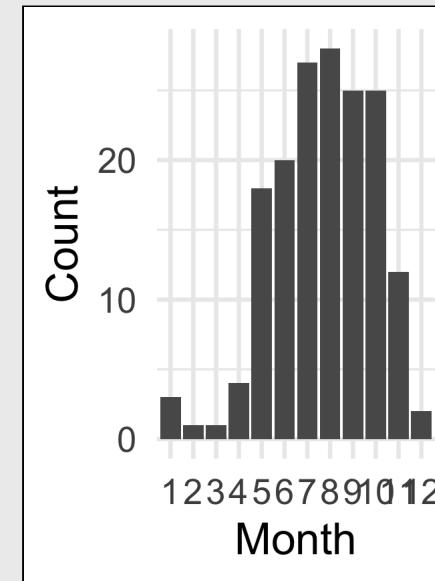


fig.path

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

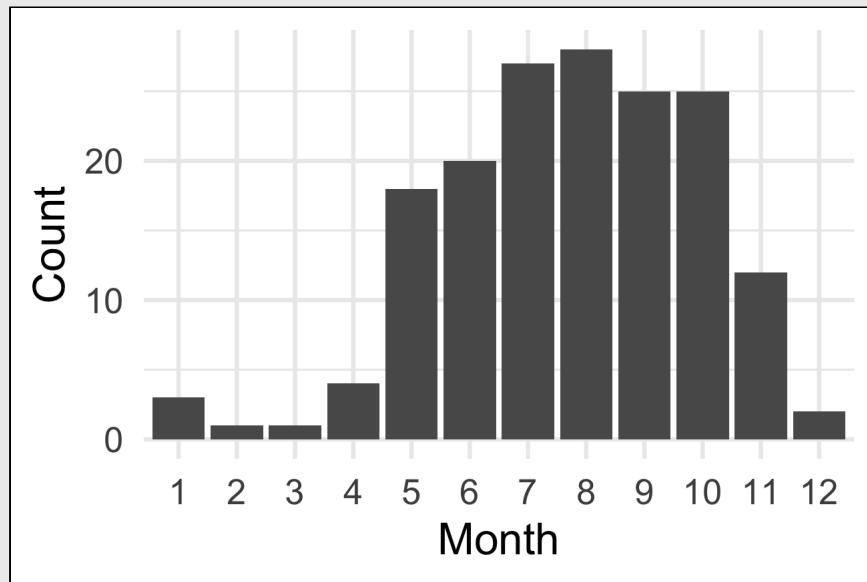
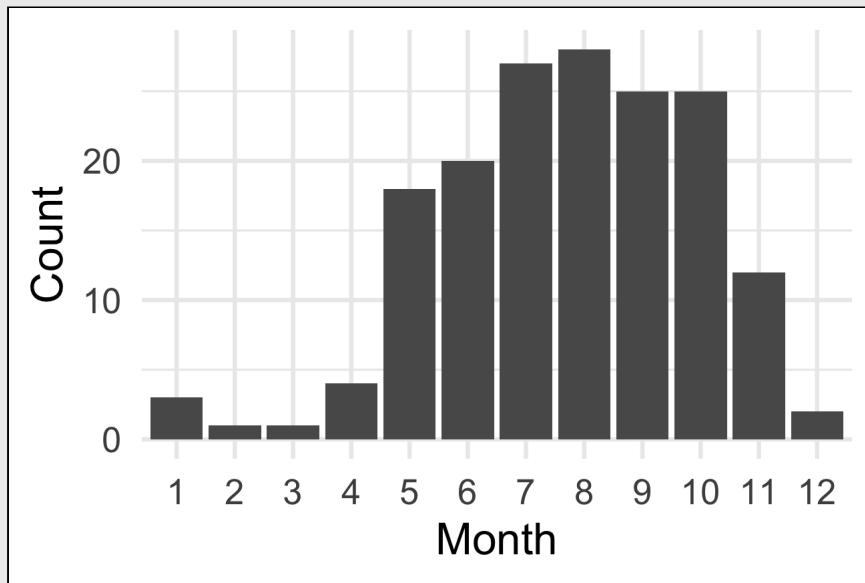


fig.path

```
```{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()
```

bearType	wildOrCaptive	n
Black	Captive	16
Black	Wild	60
Brown	Captive	8
Brown	Wild	72
Polar	Captive	4
Polar	Wild	6

15:00

Think pair share 1: College Majors

- 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 **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 categories of majors make more money than others?
 - What categories of majors have the best employment rate?
 - What are the highest earning majors?
 - Within the engineering majors, which ones have a better gender balance?