




Getting started with R Markdown

Mine Çetinkaya-Rundel



@minebocek 
mine-cetinkaya-rundel 
mine@stat.duke.edu 

Logistics

- ▶ Using RStudio Cloud: Go to <http://bit.ly/duke-rmd-cloud> and join the space.
- ▶ Using local install of RStudio cloud, either on your own machine or the lab computer: Go to <http://bit.ly/duke-rmd-materials> to download materials.



What is Markdown?

- ▶ Lightweight markup language with plain text formatting syntax
- ▶ Designed so that it can be converted to HTML (and many other formats)

README.md

```
1 # rstudioconf-2018-intro-shiny-rmarkdown
2
3 Materials for the "Intro to Shiny and R Markdown" 2-day workshop at \[rstudio::conf 2018\](https://www.rstudio.com/conference/).
4
5 ---
6
7 ## Outline
8
9 ### Day 1: January 31
10
11 #### :computer: 09:00 - 10:30: Session 1 - Introduction to R Markdown
12 - Welcome and motivation
13 - Demo: R, Markdown, knitr, Pandoc/YAML, Parameters, Templates, Tables, htmlwidgets, rpubs (and maybe shiny/shinyapps.io)
14 - Composing your prose (markdown)
15   - Text, headers, lists, hyperlinks
16   - Math text
17   - Images
18   - Tables
19   - Bibliographies and citations
20   - The Markdown Quick Reference
21
```



rstudioconf-2018-intro-shiny-rmarkdown

README

Materials for the "Intro to Shiny and R Markdown" 2-day workshop at [rstudio::conf 2018](https://www.rstudio.com/conference/).

Outline

Day 1: January 31

09:00 - 10:30: Session 1 - Introduction to R Markdown

- Welcome and motivation
- Demo: R, Markdown, knitr, Pandoc/YAML, Parameters, Templates, Tables, htmlwidgets, rpubs (and maybe shiny/shinyapps.io)
- Composing your prose (markdown)
 - Text, headers, lists, hyperlinks
 - Math text
 - Images
 - Tables
 - Bibliographies and citations
 - The Markdown Quick Reference

What is R Markdown?

- ▶ Markdown + R
- ▶ Text + R code (in chunks) gets converted to text + R code + R output in HTML (and many other formats)

my-first-rmd.rmd

```
1 ---
2 title: "My First R Markdown doc"
3 author: "Mine Cetinkaya-Rundel"
4 date: "1/23/2018"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF,
15 and 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 as
18 well as the output of any embedded R code chunks within the document. You can embed an R code
19 chunk like this:
20
21 ```{r cars}
22 summary(cars)
23 ```
```

my-first-rmd.html

My First R Markdown doc

Mine Cetinkaya-Rundel

1/23/2018

R Markdown

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

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

```
summary(cars)
```

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

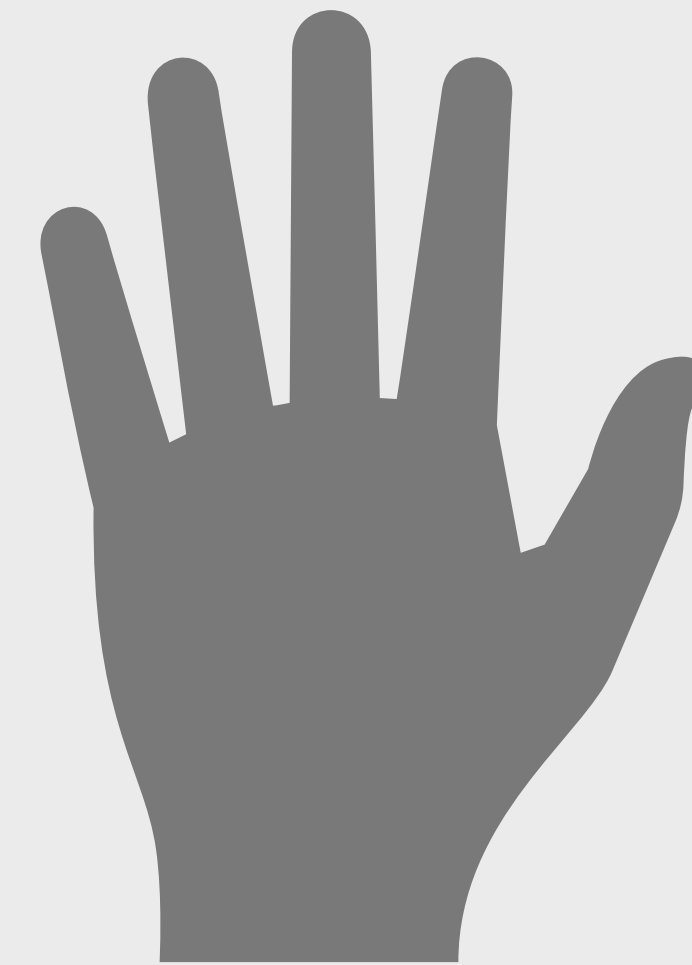


- ▶ In RStudio Cloud, make a copy of the project called “First R Markdown”
- ▶ In RStudio, File -> New File -> R Markdown...
- ▶ Title: My first R Markdown doc
- ▶ Author: [INSERT NAME]
- ▶ OK
- ▶ Knit the document



EXERCISE

- ▶ When you create an Rmd document, it will first ask you to give it a title, that is the **title** of the document. Use whatever capitalization you prefer in titles, and you can edit it later too.
- ▶ When you go to knit the file, it will ask for another name. This is the **filename**. You might want to keep it no-spaces-no-capitals.



TIP

Composing your prose

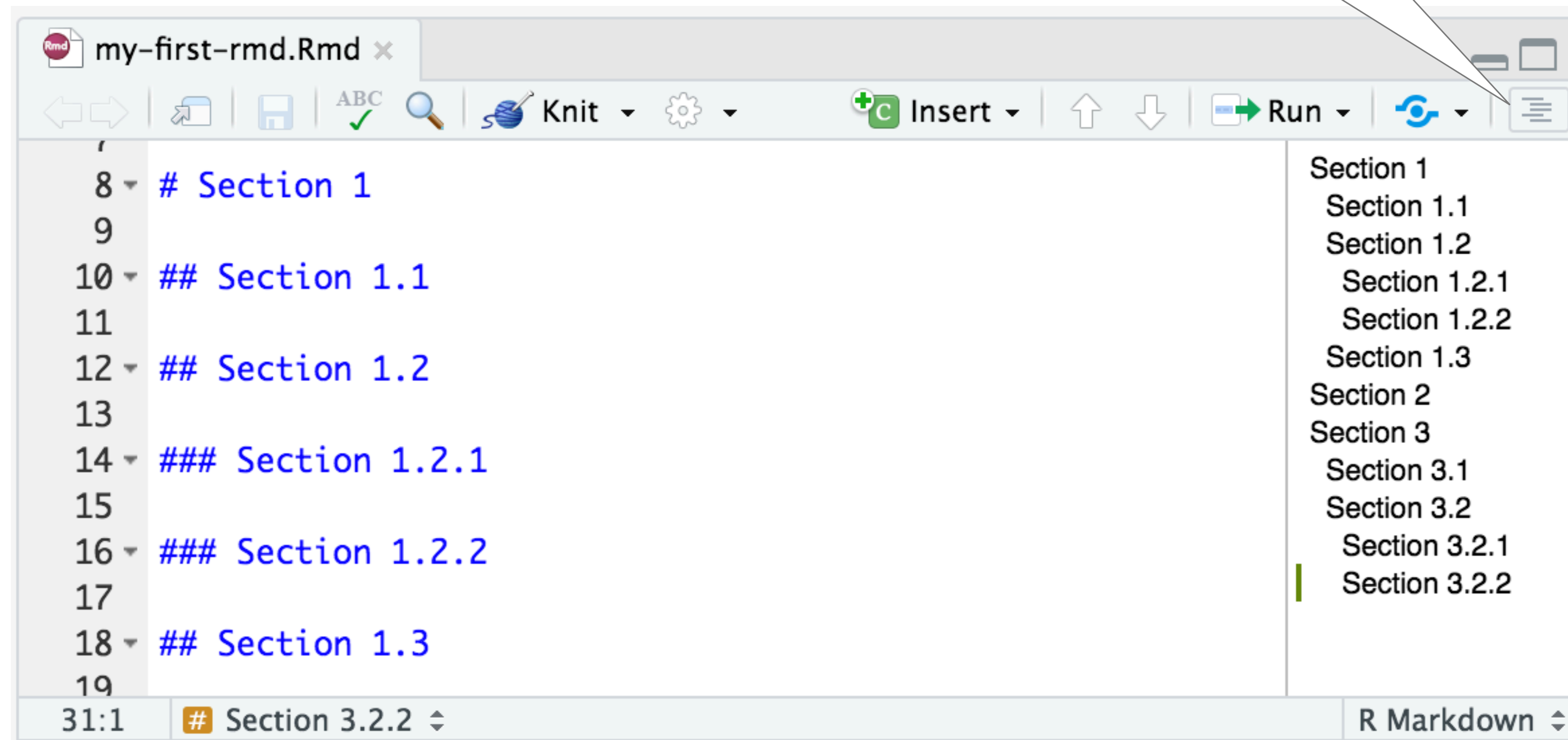


Text, hyperlinks, headers

- ▶ Text can be plain text or decorated as `*italic*` or `**bold**`
- ▶ Links can be plain http to address or add a link to a phrase:
 - ▶ `http://example.com`
 - ▶ `[linked phrase](http://example.com)`
- ▶ Headers use `#`s
 - `# Header 1`
 - `## Header 2`
 - `### Header 3`



Show / hide
document
outline



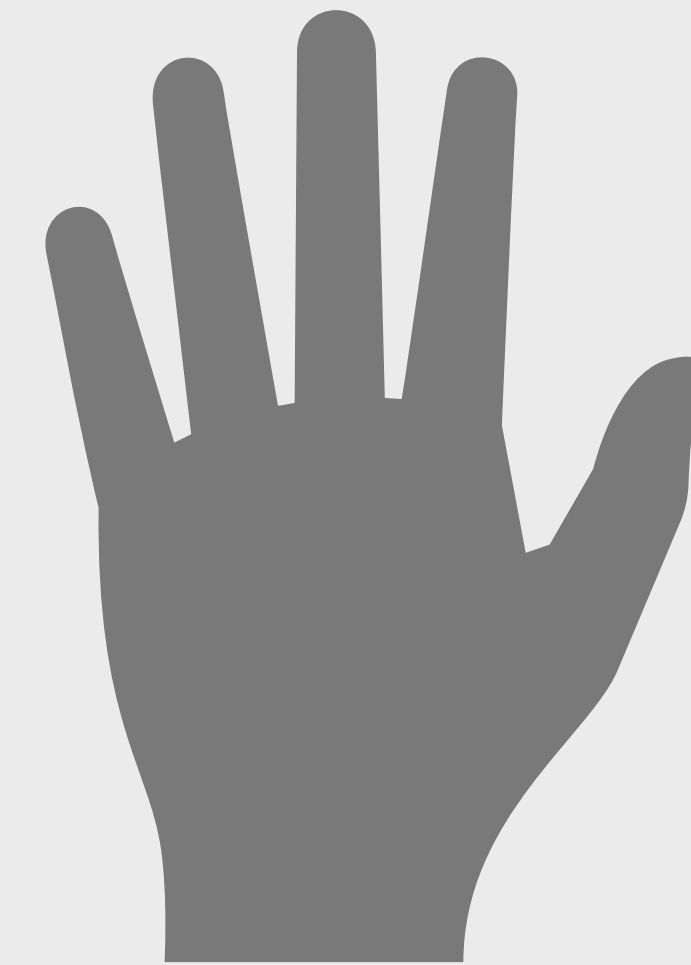
The screenshot shows the RStudio interface with a document titled "my-first-rmd.Rmd". The document content is as follows:

```
8 # Section 1
9
10 ## Section 1.1
11
12 ## Section 1.2
13
14 ### Section 1.2.1
15
16 ### Section 1.2.2
17
18 ## Section 1.3
19
```

The document outline on the right side of the editor shows the following structure:

- Section 1
 - Section 1.1
 - Section 1.2
 - Section 1.2.1
 - Section 1.2.2
 - Section 1.3
- Section 2
- Section 3
 - Section 3.1
 - Section 3.2
 - Section 3.2.1
 - Section 3.2.2

The status bar at the bottom indicates the current position is 31:1 and the document is in R Markdown mode.



TIP

Lists

► Lists can be ordered

1. Item 1
2. Item 2
 - + Item 2a
 - + Item 2b

► And lists can be unordered

- * Item 1
- * Item 2
 - + Item 2a
 - + Item 2b

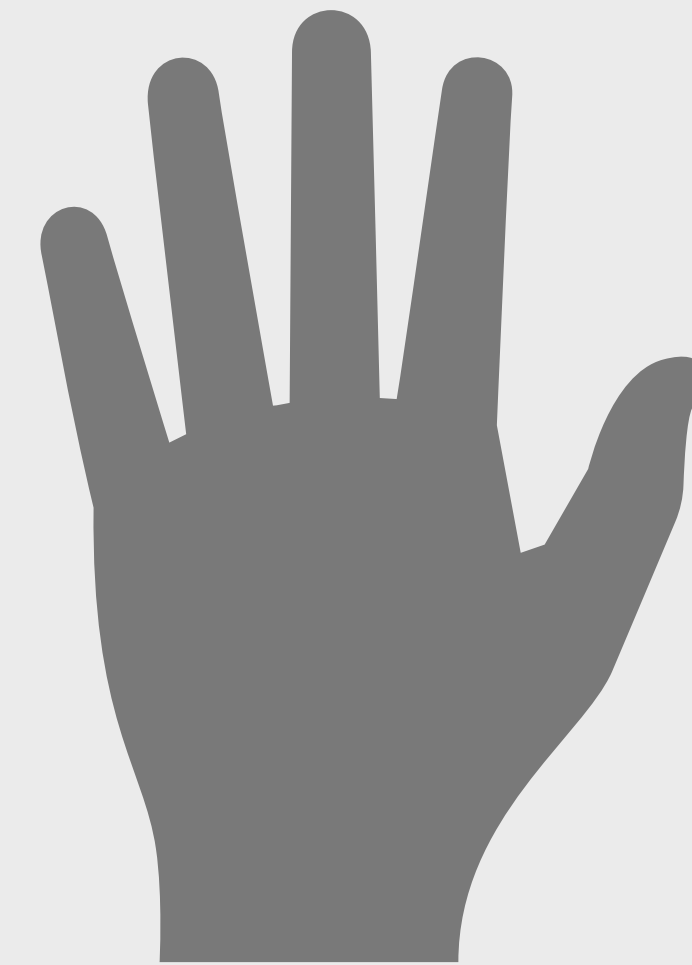
Math text

- ▶ If you already know some LaTeX, you're good to go
- ▶ Equations can be inline: `$equation$`
- ▶ And equations can be centered in a new line: `$$ equation $$`

Images

- ▶ Including an image is very similar to hyperlinking
- ▶ Images can be on the web:
`![alt text](http://example.com/logo.png)`
- ▶ Or they can be local:
`![alt text](figures/img.png)`

- To improve the accessibility of your document, always add `alt` text to you images.



TIP

Tables

- ▶ Tables are always a bit of a pain...
- ▶ You can use markdown syntax to generate them

| First Header | Second Header |
|--------------|---------------|
| Content Cell | Content Cell |
| Content Cell | Content Cell |

- ▶ Or you can format the table your R code spits out (more on this later)

Bibliographies and citations

- ▶ Include a linked phrase with an in the document body:

`[linked phrase][id]`

- ▶ And include this at the bottom of the document:

`[id]: http://example.com/ "Title"`

- ▶ You can also do styling of references, or read references in from a `.bib` file

Markdown Quick Reference

Help -> Markdown Quick Reference

Markdown Quick Reference

R Markdown is an easy-to-write plain text format for creating dynamic documents and reports.
See [Using R Markdown](#) to learn more.

Emphasis

```
*italic*    **bold**  
_italic_    __bold__
```

Headers

```
# Header 1  
## Header 2  
### Header 3
```

Lists

Unordered List

```
* Item 1  
* Item 2  
  + Item 2a  
  + Item 2b
```



- ▶ Go back to `my_first_rmd.rmd`
- ▶ Create a table with 5 columns and 6 rows, where the first two columns are left aligned, next is center aligned, and the last two are right aligned.
- ▶ See https://www.tablesgenerator.com/markdown_tables for markdown help.



EXERCISE

5_m 00_s

Embedding R Code



Code chunks

- ▶ Insert new chunks with
 - ▶ the Add Chunk button in the editor toolbar
 - ▶ typing the chunk delimiters ``{r}`` and ```
 - ▶ the keyboard shortcut **Ctrl + Alt + I** (OS X: **Cmd + Option + I**)
- ▶ When you render your .Rmd file, R Markdown will run each code chunk and embed the results beneath the code chunk in your final report.

02a-analysis.Rmd x

burritos x

← →

📄

📁

ABC ✓

🔍

🌐 Knit

⚙️

+

Insert

↑

↓

🏠 Run

🔄

☰

7

8

9

10

11

12

13

Load packages

```
{r load-packages}
library(tidyverse)

```

⚙️

▾

▶

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

The data

data comes from <https://www.kaggle.com/srcole/burritos-in-san-diego>,

```
{r load-data}
burritos <- read_csv("../data/burritos_01022018.csv")

```

⚙️

▾

▶

Mexican cuisine is often the best food option in southern California. And the burrito is the hallmark of this food: tasty, cheap, and filling. Appropriately, an effort was made to collect burritos across the county and analyze them. At this time, the data set contains information on the following variables:

```
r `r nrow(burritos)` burritos fromd `r burritos %>%
)` restaurants.
```

ons of the San Diego burrito. * Volume * Tortilla
at quality * Non-meat filling quality * Meat-to-filling

02A - Analysis

Load packages

Chunk 1: load-packages

The data

Chunk 2: load-data

31:1

The data

R Markdown

Code chunk

Engine

Chunk label

Navigation



02a-analysis.Rmd x burritos x

← → ↶ ↷

ABC 🔍

Knit ⚙

Insert ⌵

⬆ ⬇

Run ⌵

🔄 ⌵

☰

7

8 ## Load packages

9

10 ```{r load-packages}

11 library(tidyverse)

12 ```

13

14 ## The data

15

16 The data come from [Kaggle.com](https://www.kaggle.com/srcole/burritos-in-san-diego),

17

18 ```{r load-data}

19 burritos <- read_csv("../data/burritos_01022018.csv")

20 ```

21

22 Mexican cuisine is often the best food option in southern California. And the

23 burrito is the hallmark of delicious taco shop food: tasty, cheap, and filling.

24 Appropriately, an effort was launched to critique burritos across the county

25 and to the lay burrito consumer. At this time, the data set

26 contains the following variables: `r nrow(burritos)` burritos fromd `r burritos %>%`

27

28 The data

29

31:1

☰

⚙

⬇

▶

Run all chunks up to this point

Run this chunk

Options

02A - Analysis

Load packages

Chunk 1: load-packages

The data

Chunk 2: load-data

The data ⌵

R Markdown ⌵

The logo for R Markdown, featuring a red hexagon with a white quill pen icon and the word "rmarkdown" in white lowercase letters.

Chunk options

- ▶ `include = FALSE` prevents code and results from appearing in the finished file. R Markdown still runs the code in the chunk, and the results can be used by other chunks.
- ▶ `echo = FALSE` prevents code, but not the results from appearing in the finished file. This is a useful way to embed figures.
- ▶ `message = FALSE` prevents messages that are generated by code from appearing in the finished file.
- ▶ `warning = FALSE` prevents warnings that are generated by code from appearing in the finished.

- ▶ In RStudio Cloud, make a copy of the project called “San Diego Burritos”
- ▶ Open the R Markdown document called `sd-burritos.Rmd`
- ▶ Set each of the chunk options and knit your document.



EXERCISE

More chunk options

```
40 ▾ ## Exploratory data analysis
41
42 ▾ ```{r fig.height = 4, fig.width = 8, fig.cap = "Five most common reviewers"}
43 burritos %>%
44   mutate(Reviewer = fct_lump(Reviewer, n = 5)) %>%
45   ggplot(mapping = aes(x = Reviewer)) +
46     geom_bar()
47   ```
48
49 ▾ ```{r fig.height = 4, fig.width = 8, fig.cap = "Relationship between Google and Yelp
    review scores"}
50 ggplot(data = burritos, mapping = aes(x = Cost, y = overall)) +
51   geom_point()
52   ```
53
54
```

52:4 # Exploratory data analysis ↕

R Markdown ↕



Complete set of chunk options

<https://www.rstudio.com/resources/cheatsheets/>



R Markdown Reference Guide

Learn more about R Markdown at rmarkdown.rstudio.com

Learn more about Interactive Docs at shiny.rstudio.com/articles

Contents:

1. **Markdown Syntax**
2. Knitr chunk options
3. Pandoc options

Syntax

Plain text

End a line with two spaces
to start a new paragraph.

italics and *_italics_*

****bold**** and **__bold__**

superscript^{^2^}

~~~~strikethrough~~~~

[link] ([www.rstudio.com](https://www.rstudio.com))

### Becomes

Plain text

End a line with two spaces to start a new paragraph.

*italics* and *italics*

**bold** and **bold**

superscript<sup>2</sup>

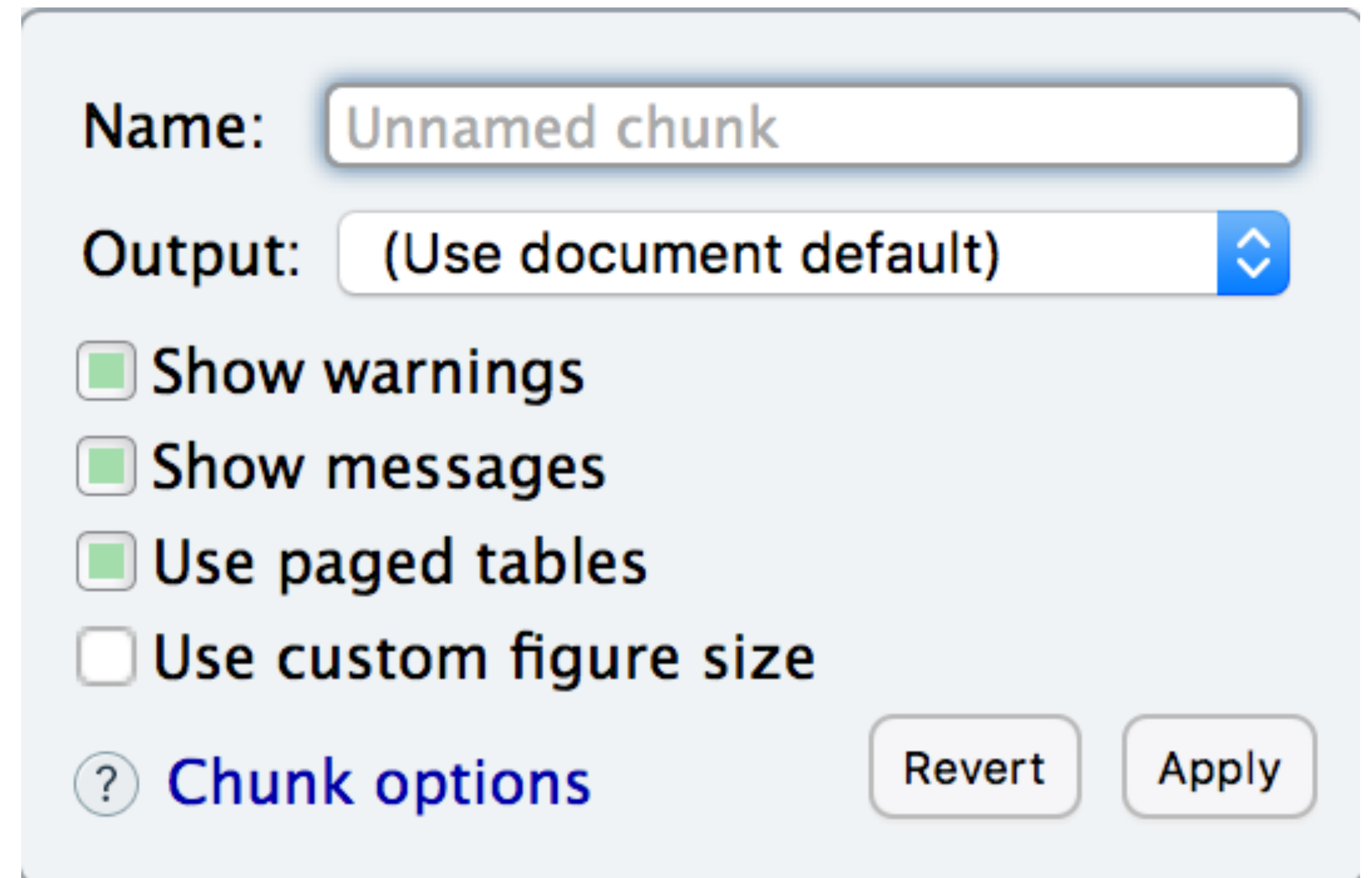
~~strikethrough~~

[link](https://www.rstudio.com)



# Setting chunk options via GUI

Some of the chunk options can be set via a handy GUI that you can access by clicking on the gear icon on a given chunk.



The screenshot shows a 'Chunk options' dialog box with a light blue background. At the top, there is a 'Name:' label followed by a text input field containing 'Unnamed chunk'. Below this is an 'Output:' label followed by a dropdown menu showing '(Use document default)' with a blue arrow icon on the right. Underneath the dropdown are four checkboxes, each followed by a label: 'Show warnings' (checked), 'Show messages' (checked), 'Use paged tables' (checked), and 'Use custom figure size' (unchecked). At the bottom left is a link with a question mark icon and the text 'Chunk options'. At the bottom right are two buttons: 'Revert' and 'Apply'.

Name:

Output:

☒ Show warnings

☒ Show messages

☒ Use paged tables

☐ Use custom figure size

? [Chunk options](#)

# Global options

- ▶ To set global options that apply to every chunk in your file, call `knitr::opts_chunk$set` in a code chunk.
- ▶ Knitr will treat each option that you pass to `knitr::opts_chunk$set` as a global default that can be overwritten in individual chunk headers.

Hide all code chunks

```
7  
8 ```{r}  
9 knitr::opts_chunk$set(echo=FALSE)  
10 ```  
11
```

- ▶ Add a new code chunk and set some options for that particular chunk. You could do plotting or summary or just basic calculation in the chunk.
- ▶ Remove the figure height and width options from individual chunks and set them as global options.



# EXERCISE

5<sub>m</sub> 00<sub>s</sub>

# Caching

- ▶ If document rendering becomes time consuming due to long computations you can use caching to improve performance
- ▶ If `cache = TRUE` is set:
  - ▶ Cached chunks are skipped, but objects created in these chunks are (lazy-) loaded from previously saved databases (.rdb and .rdx) files
  - ▶ These files are saved when a chunk is evaluated for the first time, or when cached files are not found
  - ▶ Results of the code will still be included in the output even when cache is used, because knitr also caches the printed output of a code chunk as a character string





# Inline code

- Code results can be inserted directly into the text of a .Rmd file by enclosing the code with ``r``

```
26 Mexican cuisine is often the best food option is southern California. And the
27 burrito is the hallmark of delicious taco shop food: tasty, cheap, and filling.
28 Appropriately, an effort was launched to critique burritos across the county
29 and make this data open to the lay burrito consumer. At this time, the data set
30 contains ratings from over `r nrow(burritos)` burritos fromd `r burritos %>%
count(Location) %>% nrow()` restaurants.
```

Mexican cuisine is often the best food option is southern California. And the burrito is the hallmark of delicious taco shop food: tasty, cheap, and filling. Appropriately, an effort was launched to critique burritos across the county and make this data open to the lay burrito consumer. At this time, the data set contains ratings from over 385 burritos fromd 102 restaurants.

# Inline code styling

- ▶ R Markdown will always
  - ▶ display the results of inline code, but not the code
  - ▶ apply relevant text formatting to the results
- ▶ As a result, inline output is indistinguishable from the surrounding text
- ▶ Inline expressions do not take knitr options

- ▶ Add an inline R chunk to your document. This should be something that results in a numerical value or a character string.
- ▶ What would a figure defined in an inline chunk look like?



## EXERCISE



# Other languages



# Other languages

- ▶ knitr can execute code in many languages besides R. Some of the available language engines include:
  - ▶ Python
  - ▶ SQL
  - ▶ Bash
  - ▶ Rcpp
  - ▶ Stan
  - ▶ JavaScript
  - ▶ CSS

```
1 ---
2 title: "Simple Language Demos"
3 output: html_document
4 ---
5
6 You can write code in languages other than R with R Markdown, e.g.
7
8 ## Bash
9
10 ```{bash}
11 ls *.Rmd
12 ```
13
14 ## Python
15
16 ```{python}
17 x = 'hello, python world!'
18 print(x.split(' '))
19 ```
20
21
```

Engine

# Simple Language Demos

You can write code in languages other than R with R Markdown, e.g.

## Bash

```
ls *.Rmd
```

```
## 1-example.Rmd
## 2-chunks.Rmd
## 3-inline.Rmd
## 4-languages.Rmd
```

## Python

```
x = 'hello, python world!'
print(x.split(' '))
```

```
## ['hello,', 'python', 'world!']
```

# Output formats



# Documents

- ▶ Most commonly used output is `html_document` — HTML document w/ Bootstrap CSS

- ▶ Other document options are as follows:

```
1 ---  
2 title: "02A - Analysis"  
3 author: "Mine Cetinkaya-Rundel"  
4 date: "1/23/2018"  
5 output: html_document  
6 ---
```

- ▶ `html_notebook` - Interactive R Notebooks
- ▶ `pdf_document` - PDF document (via LaTeX template)
- ▶ `word_document` - Microsoft Word document (docx)
- ▶ `odt_document` - OpenDocument Text document
- ▶ `rtf_document` - Rich Text Format document
- ▶ `md_document` - Markdown document (various flavors)



- Convert `html_document` output to `html_notebook`. What changed?



## EXERCISE

# Presentations

- ▶ `ioslides_presentation` - HTML presentation with ioslides
- ▶ `revealjs::revealjs_presentation` - HTML presentation with reveal.js
- ▶ `slidy_presentation` - HTML presentation with W3C Slidy
- ▶ `beamer_presentation` - PDF presentation with LaTeX Beamer
- ▶ `xaringan::moon_reader` - remark.js slides

# Other

- ▶ `flexdashboard::flex_dashboard` - Interactive dashboards
- ▶ `tufte::tufte_html` - HTML handouts in the style of Edward Tufte
- ▶ `html_vignette` - R package vignette (HTML)
- ▶ `github_document` - GitHub Flavored Markdown document





# Output options



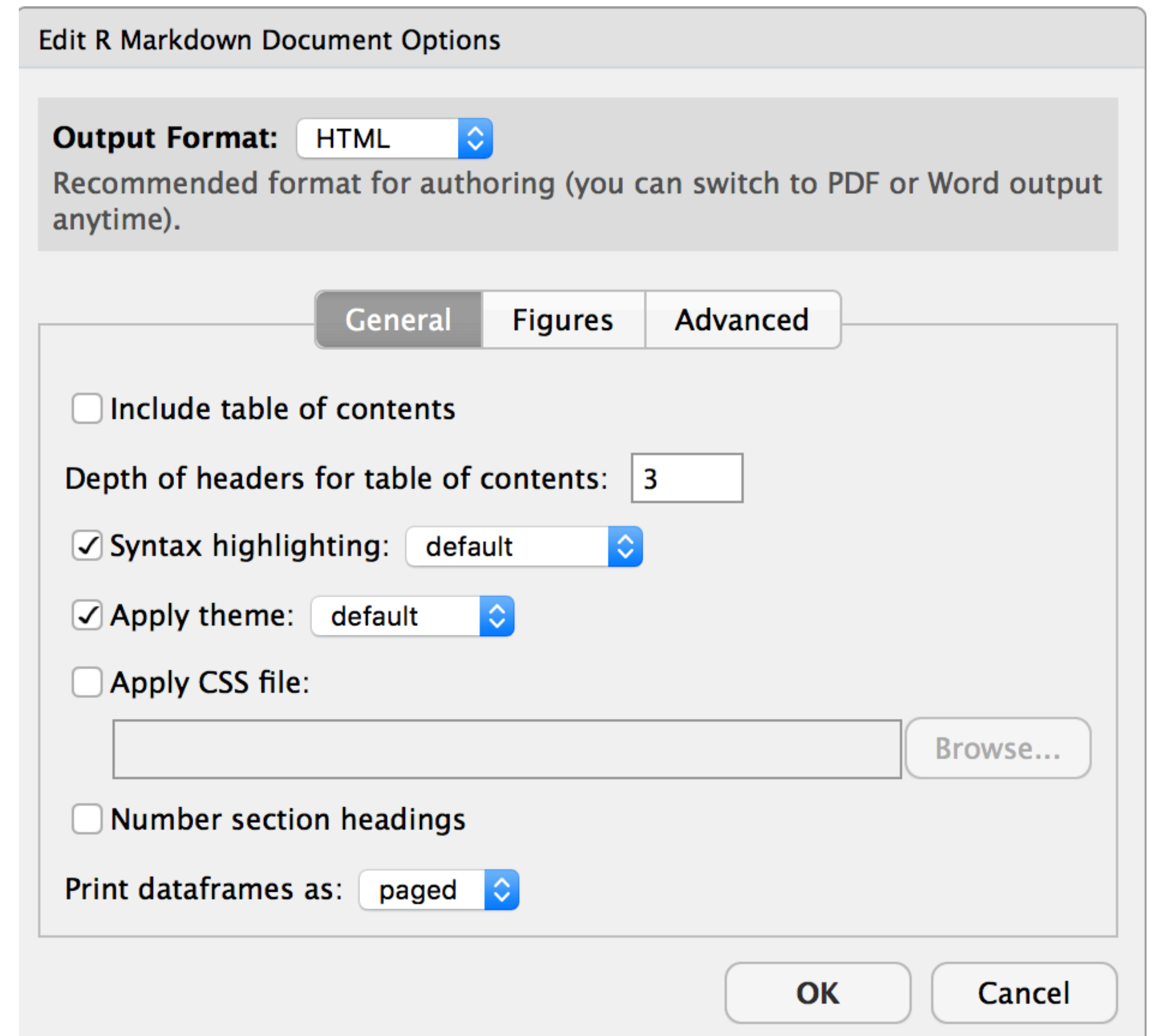
# Output options

- ▶ Output options are defined in the YAML
- ▶ They are similar to setting global knitr options
- ▶ To learn which arguments a format takes, read the format's help page in R, e.g. `?html_document`

```
html_document(toc = FALSE, toc_depth = 3, toc_float = FALSE,  
  number_sections = FALSE, section_divs = TRUE, fig_width = 7,  
  fig_height = 5, fig_retina = 2, fig_caption = TRUE, dev = "png",  
  df_print = "default", code_folding = c("none", "show", "hide"),  
  code_download = FALSE, smart = TRUE, self_contained = TRUE,  
  theme = "default", highlight = "default", mathjax = "default",  
  template = "default", extra_dependencies = NULL, css = NULL,  
  includes = NULL, keep_md = FALSE, lib_dir = NULL,  
  md_extensions = NULL, pandoc_args = NULL, ...)
```


# Setting output options via GUI

Some of the output options for some of the output types can be set via a handy GUI that you can access by clicking on the gear icon on the toolbar.




The screenshot shows the 'Edit R Markdown Document Options' dialog box with the 'General' tab selected. The 'Output Format' is set to 'HTML'. Below this, there is a note: 'Recommended format for authoring (you can switch to PDF or Word output anytime)'. The 'General' tab contains several options: 'Include table of contents' (unchecked), 'Depth of headers for table of contents' (set to 3), 'Syntax highlighting' (checked, set to 'default'), 'Apply theme' (checked, set to 'default'), 'Apply CSS file' (unchecked, with a text field and a 'Browse...' button), 'Number section headings' (unchecked), and 'Print dataframes as' (set to 'paged'). At the bottom right are 'OK' and 'Cancel' buttons.


**Edit R Markdown Document Options**

**Output Format:** HTML   
Recommended format for authoring (you can switch to PDF or Word output anytime).

**General** **Figures** **Advanced**


☐ Include table of contents  
Depth of headers for table of contents: 3

☒ Syntax highlighting: default 

☒ Apply theme: default 

☐ Apply CSS file:  
 **Browse...**

☐ Number section headings

Print dataframes as: paged 

**OK** **Cancel**

- ▶ Add a floating table of contents to the html document output.
- ▶ Also set all figures to be 4 x 7.

