

Dynamic Documents 1

Intro to R Markdown

Daniel Anderson
Week 4, Class 1



Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

2 / 28

Agenda

- Questions
- R Markdown
- Lab

Learning objectives for today

- Understand how to render R Markdown documents & mix code with text
- Understand different chunk options
- Understand inline code evaluation

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

2 / 28

Agenda

- Questions
- R Markdown
- Lab

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

2 / 28

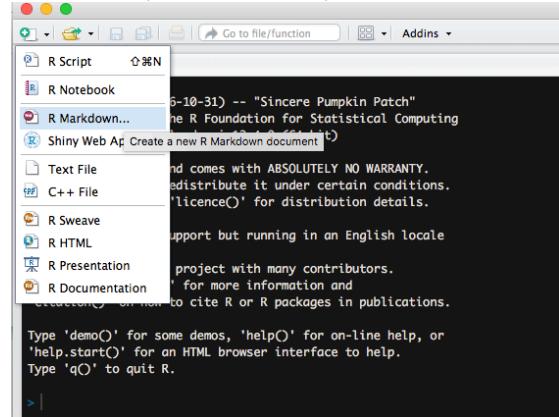
What questions do you have?

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

3 / 28

R Markdown

From within your R Studio Project:

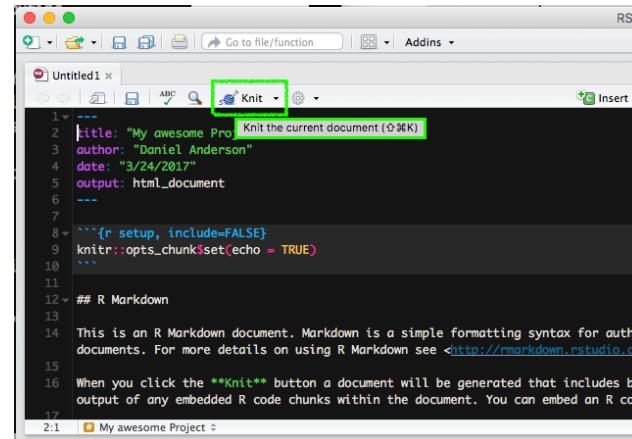


A screenshot of the R Studio interface showing the 'R Markdown' project structure. The left sidebar lists various document types: R Script, R Notebook, R Markdown..., Shiny Web App, Text File, C++ File, R Sweave, R HTML, R Presentation, and R Documentation. The 'R Markdown...' item is selected. The main pane displays the YAML front matter and the R code for the document.

```
---  
title: "Sincere Pumpkin Patch"  
author: "The R Foundation for Statistical Computing"  
date: "2010-10-31"  
output: html_document  
  
This document comes with ABSOLUTELY NO WARRANTY.  
You must redistribute it under certain conditions.  
See the 'licence()' for distribution details.  
  
Support for running in an English locale  
is provided by many contributors.  
For more information and  
instructions 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.
```

4 / 28

First thing: Render!



A screenshot of the R Studio interface showing the 'Untitled1.Rmd' file. The 'Knit' button in the toolbar is highlighted with a green box. The code in the file includes the YAML front matter and an R code chunk.

```
---  
title: "My awesome Project"  
author: "Daniel Anderson"  
date: "3/24/2017"  
output: html_document  
  
```{r setup, include=FALSE}  
knitr::opts_chunk$set(echo = TRUE)
```  
  
## R Markdown  
  
This is an R Markdown document. Markdown is a simple formatting syntax for authoring 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 output of any embedded R code chunks within the document. You can embed an R code chunk like this:  
  
2:1 My awesome Project
```

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

5 / 28

YAML Front Matter

```
---  
title: Example Markdown document  
author: Daniel Anderson  
date: "2015-09-17"  
---
```

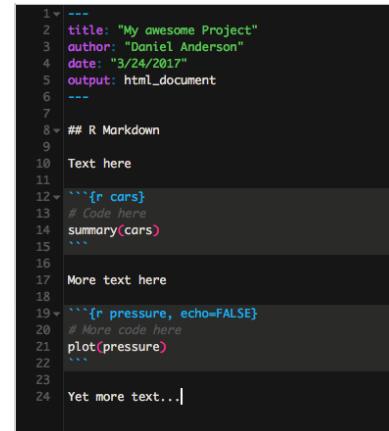
Example Markdown document

Daniel Anderson
2015-09-17

- Three dashes before and after the YAML fields
- Case sensitive
- Many other fields are possible.
 - For example, you may want to include an `output:` argument (`pdf_document`, `html_document`, `word_document`). Must be specified as it is rendered, if not supplied.

6 / 28

Code chunks versus text



A screenshot of the R Studio interface showing the 'Untitled1.Rmd' file. The code includes several code chunks and text blocks. The code chunks are indicated by triple backticks and labels like 'r cars', 'r pressure', etc.

```
---  
title: "My awesome Project"  
author: "Daniel Anderson"  
date: "3/24/2017"  
output: html_document  
  
## R Markdown  
  
Text here  
  
```{r cars}  
Code here
summary(cars)
```  
  
More text here  
  
```{r pressure, echo=FALSE}  
More code here
plot(pressure)
```  
  
Yet more text...|
```

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

7 / 28

Code chunks

Start a code chunk with ````{r}`, then produce some r code, then close the chunk with three additional back ticks `````.

```
```{r rCalc}
a <- 3
b <- 5

a + b * (exp(a)/b)
````
```

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

8 / 28

Headings and Lists

Not R-lists

```
# Level 1  
## Level 2  
### Level 3 (etc.)
```

- * Unordered list
 - inset
 - + inset more
 - etc.

1. Ordered list
 - a. blah blah
2. More stuff

Level 1

Level 2

Level 3 (etc.)

- Unordered list
 - Inset
 - inset more
 - etc.
 1. Ordered list
 - a. blah blah
 2. More stuff

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

9 / 28

Code chunks

Start a code chunk with ````{r}`, then produce some r code, then close the chunk with three additional back ticks `````.

```
```{r rCalc}
a <- 3
b <- 5

a + b * (exp(a)/b)
````
```

```
a <- 3  
b <- 5  
  
a + b * (exp(a)/b)  
  
## [1] 23.08554
```

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

8 / 28

Final Product

```
ExampleMarkdown.Rmd
```

```
1 <-- Example Markdown Rmd
2
3 title Example Markdown document
4 author Daniel Anderson
5 date "r Sys.Date()"
6 output html_document
7
8 ---
9
10 ```{r setup, include = FALSE}
11 library(knitr)
12
13 # Set global chunk options
14 opts_chunk$set(cache = TRUE, cache.comments = FALSE, autodrop = TRUE)
15
16 ## Determine caching dependencies automatically
17 dep_auto()
18
19 # Level 1
20
21 ## Level 2
22
23 ### Level 3 (etc.)
24
25
26 * Unordered list
27   - Inset
28     + Inset more
29   - etc.
30
31 1. Ordered list
32   a. blah blah
33   2. More stuff
34
35 ```{r ex_rCalc1}
36 a <- 1
37 b <- 5
38
39 a + b == (exp(a) * b)
40
41
42 This is an example of inline code, where I want to refer to the sum of 'a' and
43 'b', which is a + b.
```

file:///Users/Daniel/Dropbox/Teaching/Rcourse/Week1/Example/

| | | | | | | | | | | |
|----------|-------|-------|--------------|---------|----------|------|---------|------------|-------|------|
| inst Row | Apple | Gmail | MidMarketDev | UO Mail | Facebook | espn | Duckweb | Blackboard | Wells | Fair |
|----------|-------|-------|--------------|---------|----------|------|---------|------------|-------|------|

r - How to add subtext to an RMarkdown document... Example

Example Markdown document

Daniel Anderson
2015-09-17

Level 1

Level 2

Level 3 (etc.)

- Unordered list
 - item 1
 - item 2
 - item 3
- 1. Ordered list
 - a. first item
 - b. More stuff
- 2. Equations

$$a = b \rightarrow a = 5$$

$$a = b \rightarrow (\exp(a/b))$$

```
## [1] 23.49550
```

This is an example of inline code, where `w` is used to refer to the sum of `a` and `b`, which is 8.

You can show code without evaluating it, using `eval= FALSE`:

```
a = b = (exp(a/b))
```

Alternatively, you can evaluate the code without displaying it:

Relation between Miles Per Gallon and Horsepower

| Horsepower | MPG |
|------------|-----|
| 12 | 28 |
| 13 | 29 |
| 14 | 28 |
| 15 | 29 |
| 16 | 28 |
| 17 | 29 |
| 18 | 28 |
| 19 | 29 |
| 20 | 28 |
| 21 | 29 |
| 22 | 28 |
| 23 | 29 |
| 24 | 28 |
| 25 | 29 |
| 26 | 28 |
| 27 | 29 |
| 28 | 28 |
| 29 | 29 |
| 30 | 28 |
| 31 | 29 |
| 32 | 28 |
| 33 | 29 |
| 34 | 28 |
| 35 | 29 |

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

10 / 28

More advanced options

- Chunk options
- Setting global options
- Inline code evaluation

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

11 / 28

(and a few more)

| Options | Arguments | Default | Result |
|-------------------------|---------------------------|---------|--|
| 9 cache | logical, 0:3 | FALSE | Cache code chunks? |
| 10 cache.comments | logical | NULL | Cache invalidated by comment changes? |
| 11 dependson | char, num | NULL | Current chunk depend on prior cached chunks? |
| 12 autodep | logical | FALSE | Depends determined automatically? |
| 13 fig.height/fig.width | numeric | 7, 7 | Height and width of figure |
| 14 fig.show | asis, hold, animate, hide | asis | How the figure should be displayed |
| 15 interval | numeric | 1 | Animate speed |

For complete documentation, see <http://yihui.name/knitr/options/>

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

13 / 28

A few select chunk options

| Options | Arguments | Default | Result |
|---------|--------------------------|---------|--|
| eval | logical | TRUE | Evaluate the code? |
| echo | logical | TRUE | Show the code? |
| results | markup, asis, hold, hide | markup | Render the results |
| warning | logical | TRUE | Print warnings? |
| error | logical | TRUE | Preserve errors? (if FALSE, quit) |
| message | logical | TRUE | Print any messages? |
| include | logical | TRUE | Include any of the code or output or code? |
| tidy | logical | FALSE | Tidy code? (see formatR package) |

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

12 / 28

echo and eval

You can show code without evaluating it, using `eval = FALSE`.

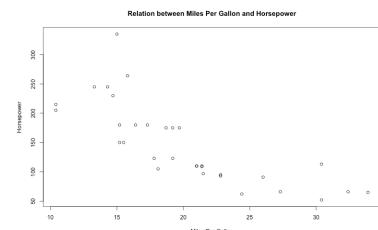
```
```{r ex_rCalc2, eval = FALSE}
a + b * (exp(a)/b)
```
```

```
a + b * (exp(a)/b)
```

Alternatively, you can evaluate the code without displaying it, using `echo = FALSE`.

```
```{r plotExample, echo = FALSE, fig.width = 6, fig.height = 3.8}
data(mtcars)
with(mtcars, plot(mpg, hp,
xlab = "Miles Per Gallon",
ylab = "Horsepower",
main = "Relation between Miles Per Gallon and Horsepower"))
```

```



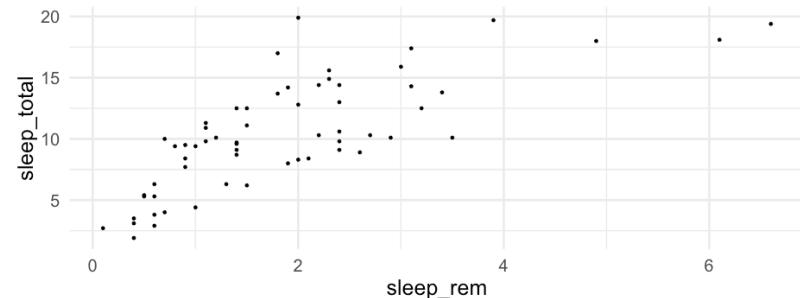
Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

14 / 28

warning

`Warning = FALSE`

```
ggplot(msleep, aes(sleep_rem, sleep_total)) +  
  geom_point()
```

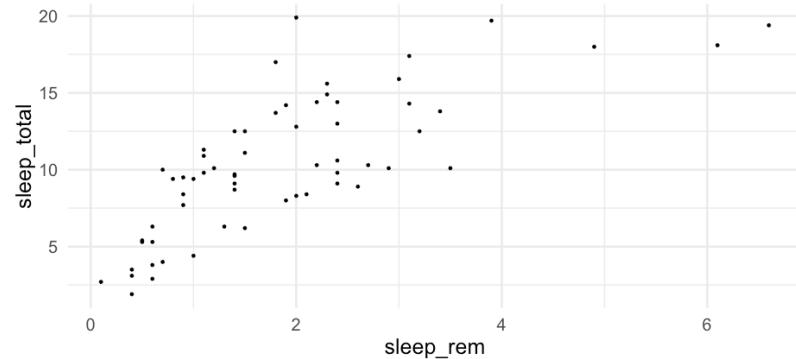


Warning is printed to the console when rendering.

`Warning = TRUE`

```
ggplot(msleep, aes(sleep_rem, sleep_total)) +  
  geom_point()
```

```
## Warning: Removed 22 rows containing missing values (geom_point).
```



Show errors

`error = TRUE`

```
ggplot(msleep, aes(sleep, sleep_total)) +  
  geom_point()
```

```
## Error: Aesthetics must be either length 1 or the same as the data (83): x
```

Show errors

`error = TRUE`

```
ggplot(msleep, aes(sleep, sleep_total)) +  
  geom_point()
```

```
## Error: Aesthetics must be either length 1 or the same as the data (83): x
```

If `error = FALSE`, the document won't render if it encounters an error.

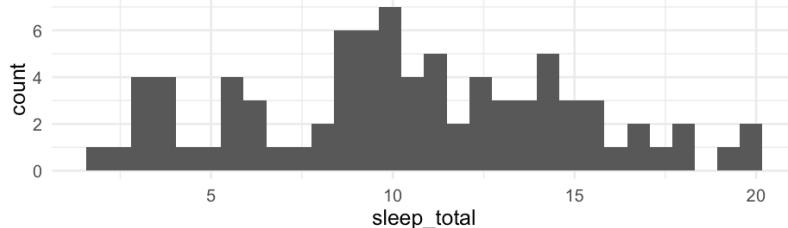
```
|.....| 67%  
|.....| 71%  
|.....| 76%  
|.....| 81%  
|.....| 86%  
label: showErrors (with options)  
List of 1  
$ error: logi FALSE  
Quitting from lines 300-303 (dynamicDocuments.Rmd)  
Error: Aesthetics must be either length 1 or the same as the data (83): x, y  
> |
```

Message

Some functions will return messages. You may want to suppress these.

`message = FALSE`

```
ggplot(msleep, aes(sleep_total)) +  
  geom_histogram()
```



Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

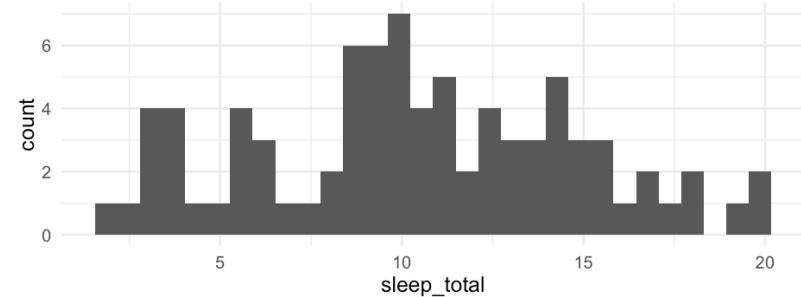
18 / 28

Message

`message = TRUE`

```
ggplot(msleep, aes(sleep_total)) +  
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

19 / 28

include

```
```{r setup, include = FALSE}  
library(knitr)

Set global chunk options
opts_chunk$set(cache = TRUE, cache.comments = FALSE, autodep = TRUE)

dep_auto()
```
```

The `include` argument is used to evaluate code that is not included in the document at all. For example, when setting up your global options.

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

20 / 28

Setting global options

Change the default behavior

```
opts_chunk$set(options)
```

For example, you can set `echo = FALSE` and `fig.width = 6.5` and `fig.height = 8` with the following code.

```
opts_chunk$set(echo = FALSE, fig.width = 6.5, fig.height = 8)
```

This is most useful when producing a report for somebody who doesn't use R and has no use or knowledge of the code.

You can always override the global options within a particular chunk, e.g.

```
```{r, chunkName, echo = TRUE}  

```
```

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

21 / 28

Other things to consider setting globally:

- `warnings = FALSE`
- `message = FALSE`
- `errors = TRUE`
- `echo = FALSE`
- Caching options (next slides)

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

22 / 28

Real example

```
```{r ell_grp_means}
ell_means <- with(d, tapply(reading, ell, mean, na.rm = TRUE))
```

Figure 2 represents a example of why the PP plot is such a powerful visualization of group differences. As previously stated, the *Monitor* group had the highest average achievement, scoring `r round(ell_means[2] - ell_means[3], 2)` points higher than *Non-ELL* students and `r round(ell_means[2] - ell_means[1], 2)` points higher than *Active* students, on average. However, the data visualization provides a more nuanced picture of the achievement differences. The *Monitor* curve is below the reference line on the lower end of the scale, but above the reference line at the upper end of the scale. In other words, students in the *Monitor* group only scored higher than students in the *Non-ELL* group at the bottom of the scale. At the upper end of the scale the effect was reversed, and students in the *Non-ELL* group had the higher achievement. In this case, the line crosses at essentially the 50th percentile of achievement for non-ELL students. Observing these achievement differences may help lead us to more refined research questions and research hypotheses. For example, a reasonable hypothesis stemming from Figure 2 may be that low-performing students are in need of additional attention, and that providing additional attention is beneficial to their academic achievement even if the intervention is not strictly focused on academics. Interestingly, although not reported here, when the equivalent plot is produced with the mathematics outcome, the overall group differences appear very similar at the bottom of the scale, but the groups are essentially indistinguishable above the 50th percentile.
```

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

24 / 28

Inline code

A single back tick followed by `r` produces inline code to be evaluated.

```
This is an example of inline code, where I want to refer to the sum of `a` and `b`, which is `r a + b`.
```

This is an example of inline code, where I want to refer to the sum of `a` and `b`, which is 8.

This is *extremely* useful in writing reports. Never have to update any numbers in text, regardless of changes to your models or data (if you are careful about it).

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

23 / 28

Want to customize further?

The YAML will control a lot of how a document looks. For example, if you wanted to render with a different syntax highlighter:

Standard Rmd

```
---
```

```
title: "Doc Title"
```

```
output: pdf_document
```

```
---
```

kate

```
---
```

```
title: "Doc Title"
```

```
output:
```

```
  pdf_document:
```

```
    highlight: kate
```

```
---
```

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

25 / 28

Slides available at: <http://www.datalorax.com/vita/ds/ds1-slides/w4p1/>

Last bit

Formats

- R Markdown will render to HTML and PDF really well
 - Word kinda-sorta supported
 - You can also create websites, slides (like these), books, etc.

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

26 / 28

Last bit

Formats

- R Markdown will render to HTML and PDF really well
 - Word kinda-sorta supported
 - You can also create websites, slides (like these), books, etc.
- Need a tex (pronounced tek) distribution
 - My recommendation for this class, and probs everything you'll ever need: {tinytex}

```
install.packages("tinytex")
tinytex::install_tinytex()
```

This is another amazing package by Yihui Xie. See more about it [here](#)

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

26 / 28

Last bit

Formats

- R Markdown will render to HTML and PDF really well
 - Word kinda-sorta supported
 - You can also create websites, slides (like these), books, etc.
- Need a tex (pronounced tek) distribution
 - My recommendation for this class, and probs everything you'll ever need: {tinytex}

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

26 / 28

Full distributions

- MacTeX: <http://www.tug.org/mactex/>
- MiKTeX: <https://miktex.org/download>
- TeXLive: <http://www.tug.org/texlive/>

Slides available at: <http://www.datolorax.com/vita/ds/ds1-slides/w4p1/>

27 / 28

Full distributions

- MacTex: <http://www.tug.org/mactex/>
- MikTex: <https://miktex.org/download>
- TexLive: <http://www.tug.org/texlive/>

Modify the YAML

Get the same document to render to different formats by modifying the YAML

```
output:  
  pdf_document:  
    highlight: kate  
  html_document:  
    highlight: tango  
  word_document: default
```

Lab