

Dynamic documents in R

reproducible research with R Markdown

2021-10-12

Rmarkdown

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



Artwork by @allison_horst

R Markdown



Authoring framework: code and text in same document

Reproducible: re-run your analysis

Flexible: Output to different formats easily



knitting

Your turn 1

Create a new R Markdown file. Go to File > New File > R Markdown. Press OK. Save the file and press the "Knit" button above.

```
~/Documents/rmarkdown - gh-pages - RStudio
>Addins
```

1-example.Rmd x

1 ---
2 title: "Viridis Demo"
3 output: html_document
4 ---
5
6 ```{r include = FALSE}
7 library(viridis)
8 ...
9
10 The code below demonstrates two color palettes in the
[viridis](<https://github.com/sjmgarnier/viridis>) package. Each
plot displays a contour map of the Maunga Whau volcano in
Auckland, New Zealand.
11
12 ## Viridis colors
13
14 ```{r}
15 image(volcano, col = viridis(200))
16 ...
17
18 ## Magma colors
19
20 ```{r}
21 image(volcano, col = viridis(200, option = "A"))
22 ...
23

Environment History Build Git

Files Plots Packages Help Viewer

YAML Metadata } Plain text } Code chunk

1:1 Viridis Demo ▾ R Markdown ▾

Console

R Markdown



Prose

Code

Metadata

R Markdown

Prose = Markdown

Code

Metadata



Visual R Markdown

Filtering joins match observations in the same way as [mutating joins](#), but affect the observations, not the variables¹. There are two types:

<code>semi_join(x, y)</code>	$x \times y$	Keeps all observations in x that have a match in y
<code>anti_join(x, y)</code>	$x > y$	Drops all observations in x that have a match in y

Graphically, a semi-join looks like this:

```
{r, echo = FALSE, out.width = NULL}
knitr::include_graphics("diagrams/join-semi.png")
```

key	val_x
1	x1
2	x2

Only the existence of a match is important; it doesn't matter which observation is matched. This means that filtering joins never duplicate rows like mutating joins do:

Basic Markdown Syntax

italic **bold**

italic __bold__

Basic Markdown Syntax

```
# Header 1
```

```
## Header 2
```

```
### Header 3
```

Basic Markdown Syntax

`http://example.com`

`[linked phrase](http://example.com)`

**Learn more about Markdown Syntax
with the ten-twenty minute tutorial on
markdown at
[https://commonmark.org/help/tutorial.](https://commonmark.org/help/tutorial)**

Your turn 2 (open exercises.Rmd)

Read this short introduction to Visual R Markdown:

<https://rstudio.github.io/visual-markdown-editing/#/intro?id=getting-started>

Use Visual R Markdown to stylize the text from the Gapminder website below. Experiment with bolding, italicizing, making lists, etc.

R Markdown

Prose

Code = R code chunks

Metadata



Code chunks

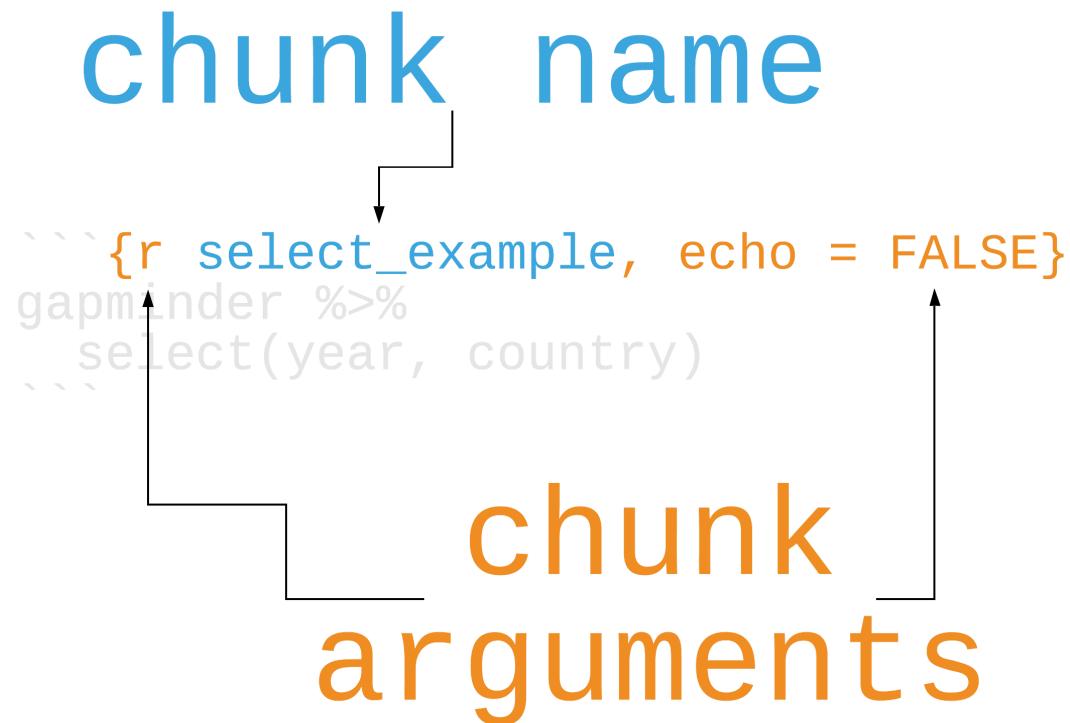
```
```{r select_example, echo = FALSE}
gapminder %>%
 select(year, country)
```
```

Code chunks

fences (3
backticks)

```
r select_example, echo = FALSE}  
gapminder %>%  
  select(year, country)
```

Code chunks



Chunk options

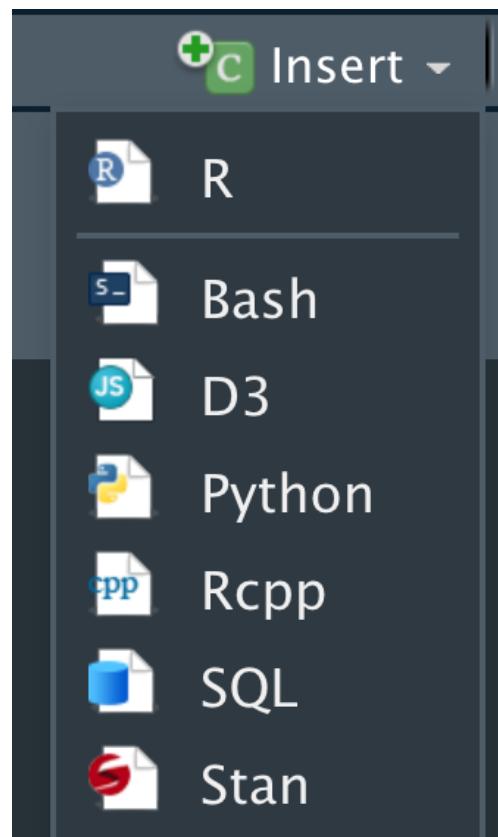
| Option | Effect |
|-----------------------------------|--|
| <code>include = FALSE</code> | run the code but don't print it or results |
| <code>eval = FALSE</code> | don't evaluate the code |
| <code>echo = FALSE</code> | run the code and output but don't print code |
| <code>message = FALSE</code> | don't print messages (e.g. from a function) |
| <code>warning = FALSE</code> | don't print warnings |
| <code>fig.cap = "Figure 1"</code> | caption output plot with "Figure 1" |

See the [knitr web page](#)

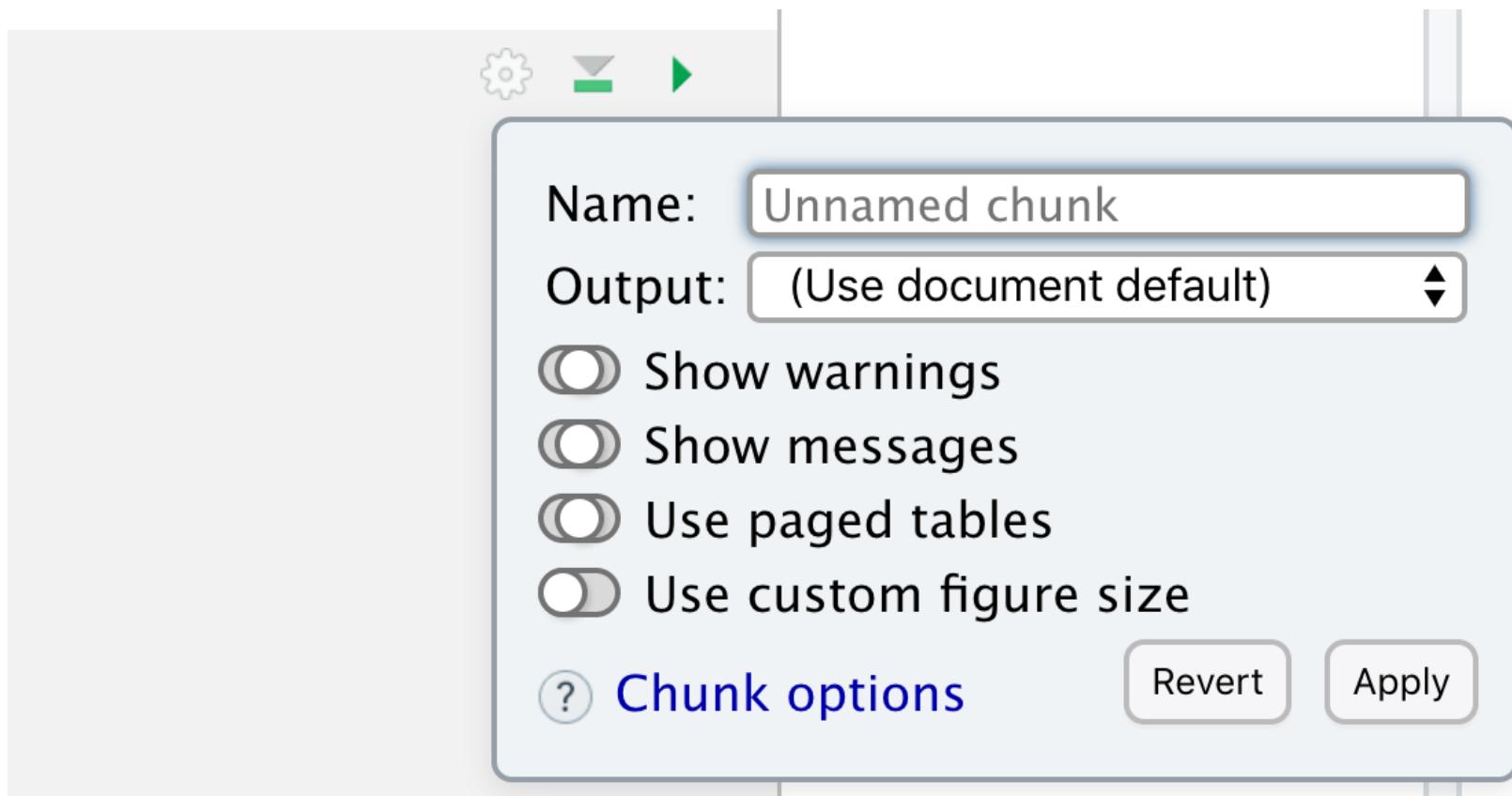
Engines

52! Including **Python, Julia, C++, SQL, SAS, and Stata**

Insert code chunks with cmd/ctrl + alt/option + I



Edit code chunk options



Your turn 3

Create a code chunk. You can type it in manually, use the keyboard short-cut (Cmd/Ctrl + Option/Alt + I), or use the "Insert" button above. Put the following code in it:

```
gapminder %>%
  slice(1:5) %>%
  knitr::kable()
```

Knit the document

Your turn 4

Add echo = FALSE to the code chunk above and re-knit

Remove echo = FALSE from the code chunk and move it to knitr::opts_chunk\$set() in the setup code chunk. Re-knit. What's different about this?

Make sure to remove knitr::opts_chunk\$set(echo = FALSE)

Inline Code

Lore*m ipsum dolor sit
amet, consetetur
sadipscing
`r max(gapminder\$year)`
elitr, sed diam nonumy
eirmod tempor invidunt*

Inline Code

The diagram illustrates the use of backticks in R. A large orange box labeled "backticks" contains the text "anet, consectetur + r sadipscing". Below this, a blue box labeled "any R code" contains the R command `r max(gapminder\$year)`. Two arrows point from the "backticks" box to the "any R code" box: one arrow points from the "anet, consectetur" part to the first character of the command, and another arrow points from the "sadipscing" part to the closing brace of the command.

```
r max(gapminder$year)
```

Your turn 5

Remove eval = FALSE so that R Markdown evaluates the code.

Use summarize() and n_distinct() to get the the number of unique years in gapminder and save the results as n_years.

Use inline code to describe the data set in the text below the code chunk and re-knit.

R Markdown

Prose

Code

Metadata = YAML



YAML Metadata

```
---
author: Malcolm Barrett
title: Quarterly Report
output:
  html_document: default
  pdf_document:
    toc: true
---
```

title: "Annual report"

author: Malcolm Barrett

date: "`r Sys.Date()`"

output:

pdf_document:

toc: true

```
title: "Annual report"
author: Malcolm Barrett
date: "r Sys.Date()"

output:
  pdf_document:
    toc: true
```

```
title: "Annual report"
```

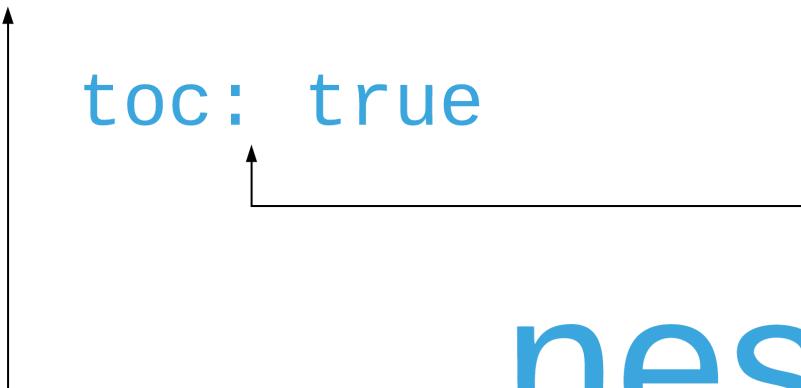
```
author: Malcolm Barrett
```

```
date: "``r Sys.Date()``"
```

```
output: ← top level
```

```
pdf_document:
```

```
  toc: true
```



```
nested
```

```
title: "Annual report"  
author: Malcolm Barron  
date: "r Sys.Date()"  
output:  
pdf_document:  
  toc: true
```

The diagram illustrates the flow of parameters from the top section to the bottom section. A bracket on the left groups 'output', 'pdf_document:', and 'toc: true'. An arrow points from this bracket down to the 'pdf_document:' line. Another arrow points from 'pdf_document:' up to the word 'arguments'.

output function

output arguments

```
title: "Annual report"  
author: Malcolm Barrett  
date: "`r Sys.Date()`"  
output:  
  pdf_document:
```

```
    toc: true
```

pdf_document(toc = TRUE)

Output formats

| Function | Outputs |
|--|---------------------|
| <code>html_document()</code> | HTML |
| <code>pdf_document()</code> | PDF |
| <code>word_document()</code> | Word .docx |
| <code>odt_document()</code> | .odt |
| <code>rtf_document()</code> | .rtf |
| <code>md_document()</code> | Markdown |
| <code>slidy_presentation()</code> | Slidy Slides (HTML) |
| <code>beamer_presentation()</code> | Beamer Slides (PDF) |
| <code>ioslides_presentation()</code> | ioslides (HTML) |
| <code>powerpoint_presentation()</code> | Powerpoint Slides |

Your turn 6

Set figure chunk options to the code chunk below, such as dpi, fig.width, and fig.height. Run `knitr::opts_chunk$get()` in the console to see the defaults.

Add your name to the YAML header using author: Your Name.

Change the YAML header above from output: html_document to output: distill::distill_article.

Set distill::distill_article to use the toc and code_folding options and re-knit

`ymlthis`

check out the `ymlthis` package for tools
and documentation for working with
`YAML`

<https://r-lib.github.io/ymlthis/>

Parameters

```
---  
params:  
  param1: x  
  param2: y  
  data: df  
---
```

```
params$param1  
params$param2  
params$data
```

Your turn 7

Change the params option in the YAML header to use a different continent. Re-knit

```
gapminder %>%
  filter(continent == params$continent) %>%
  ggplot(aes(x = year, y = lifeExp, group = country, color = country
  geom_line(lwd = 1, show.legend = FALSE) +
  scale_color_manual(values = country_colors) +
  theme_minimal(14) +
  theme(strip.text = element_text(size = rel(1.1))) +
  ggtitle(paste("Continent:", params$continent))
```

Bibliographies and citations

Bibliographies and citations

Bibliography files: .bib, Zotero, others

Bibliographies and citations

Bibliography files: .bib, Zotero, others

Citation styles: .csl

Bibliographies and citations

Bibliography files: .bib, Zotero, others

Citation styles: .csl

[@citation-label]

Or just use Visual R Markdown's citation wizard!

Including bibliography files in YAML

```
---
```

```
bibliography: file.bib
```

```
csl: file.csl
```

```
--
```

Visual R Markdown can also manage this for you.

Your turn 8

Cite the Causal Inference book in text below. Using the citation wizard, find the right citation under My sources > Bibliography.

Add the American Journal of Epidemiology CSL to the YAML using csl: aje.csl

Re-knit

Make cool stuff in R Markdown!

`bookdown`

`blogdown`

`these slides!`

Resources

R Markdown: A comprehensive but friendly introduction to R Markdown and friends. Free online.

R for Data Science: A comprehensive but friendly introduction to the tidyverse. Free online.

R Markdown for Scientists: R Markdown for Scientists workshop material.