

Creating reports with Rmarkdown

Joselyn Chávez

06/21/2022

Applications of R Markdown



- Rendering scripts (knitr)
- Creating reports (pagedown-pdf)
- Websites (pagedown)
- Books (bookdown)
- Presentations (xaringan)
- CV (vitae, datadrivencv)
- Thesis (thesisdown, iheiddown)

R script vs R markdown

Script

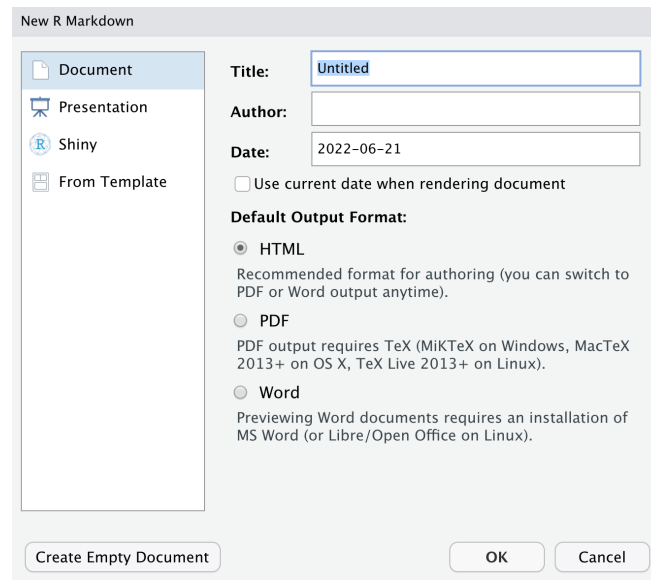
- Code is written directly on the script
- Comments are marked with '#'
- Output is printed on the console

R markdown

- Code is written inside R chunks
- Comments are written directly on the document
- Output on multiple formats

Starting a pipeline

- File > New file > R Markdown
- Assign a title
- Create empty document



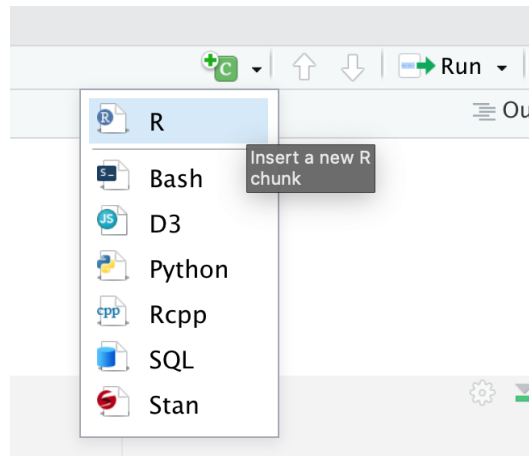
The screenshot shows the 'New R Markdown' dialog box. On the left, there is a sidebar with four options: 'Document' (selected), 'Presentation', 'Shiny', and 'From Template'. The main area contains the following fields and options:

- Title:** A text box containing 'Untitled'.
- Author:** An empty text box.
- Date:** A text box containing '2022-06-21'.
- ☐ Use current date when rendering document
- Default Output Format:**
 - ☒ **HTML**
Recommended format for authoring (you can switch to PDF or Word output anytime).
 - ☐ **PDF**
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).
 - ☐ **Word**
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

At the bottom, there are three buttons: 'Create Empty Document', 'OK', and 'Cancel'.

Starting a pipeline

- Open a new chunk



Starting a pipeline

- Start writing some code

```
library(tidyverse)
library(janitor)
sinai_covid <- read_csv("sinai_covid.csv")
sinai_covid <- clean_names(sinai_covid)
```

```
sinai_covid %>%
  select(age, ethnicity, facility) %>%
  head()

ggplot(sinai_covid, aes(x = age, color = facility)) +
  geom_line()
```

Chunk options

- echo
- eval
- warnings
- message
- out.width
- out.height
- fig.align
- fig.alt
- ...

Bullet lists

You can use bullets (+, -, *) or numbers (1, 1.1)

```
## Bullet lists
```

```
- First element  
+ Second element  
* Third element
```

```
1. A number  
2. Another number
```

```
1.1 A deep number  
1.2 A second number
```


Working with tables

Using the kableExtra package

```
library(kableExtra)
x <- sinai_covid %>%
  select(age, ethnicity, facility) %>%
  head(2)

kable(x, caption = "Head of sinai covid") %>%
  kable_styling()
```

Head of sinai covid

age	ethnicity	facility
63	NON-HISPANIC	THE MOUNT SINAI HOSPITAL
64	NON-HISPANIC	THE MOUNT SINAI HOSPITAL

More styling

```
kable(x, caption = "Head of sinai covid") %>%  
  kable_styling() %>%  
  kable_classic(full_width = F, html_font = "Cambria")
```

Head of sinai covid		
age	ethnicity	facility
63	NON-HISPANIC	THE MOUNT SINAI HOSPITAL
64	NON-HISPANIC	THE MOUNT SINAI HOSPITAL

Using the flextable package

```
library(flextable)
flextable(x) %>%
  autofit()
```

age	ethnicity	facility
63	NON-HISPANIC	THE MOUNT SINAI HOSPITAL
64	NON-HISPANIC	THE MOUNT SINAI HOSPITAL

The yaml section

- Add a title

```
---  
title: "My first pipeline"  
output:  
  html_document  
---
```

- Create a table of contents

```
---  
title: "My first pipeline"  
output:  
  html_document:  
    toc: true  
    toc_depth: 2  
---
```

- Float the table of contents

```
---  
title: "My first pipeline"  
output:  
  html_document:  
    toc: true  
    toc_float: true  
---
```

Animated plots with gganimate

```
library(gganimate)

influenza <- read.csv("influenza_vaccination.csv")
influenza <- clean_names(influenza)

influenza %>%
  group_by(mmwr_year, race_eth) %>%
  summarise(mean_vaccinated = mean(vaccinated)) %>%
  ggplot(aes(factor(mmwr_year), mean_vaccinated, color = race_eth))
  geom_point()
```

Adding time transition

```
influenza %>%  
  group_by(mmwr_year, race_eth) %>%  
  summarise(mean_vaccinated = mean(vaccinated)) %>%  
  ggplot(aes(race_eth, mean_vaccinated)) +  
  geom_col() +  
  theme(axis.text.x = element_text(angle = 90)) +  
  transition_states(mmwr_year)
```

Adding time as title

```
influenza %>%  
  group_by(mmwr_year, race_eth) %>%  
  summarise(mean_vaccinated = mean(vaccinated)) %>%  
  ggplot(aes(race_eth, mean_vaccinated)) +  
  geom_col() +  
  theme(axis.text.x = element_text(angle = 90)) +  
  labs(title = 'Year: {frame_time}') +  
  transition_time(mmwr_year)
```


Statistical tests

Nature of the data	Dependence of the data			
	Independent		Paired	
	2 data sets	≥3 data sets	2 data sets	≥3 data sets
Interval with assumptions	Unpaired t-test	Analysis of variance (ANOVA)	Paired t-test	Repeated measures analysis of variance (ANOVA)
Ordinal (or Interval without assumptions)	Mann-Whitney test	Kruskal-Wallis test	Wilcoxon test	Friedman test
Nominal	Chi-squared test (best referred as Fisher exact test)	Chi-squared test	McNemar test	Cochran Q test

Thanks!

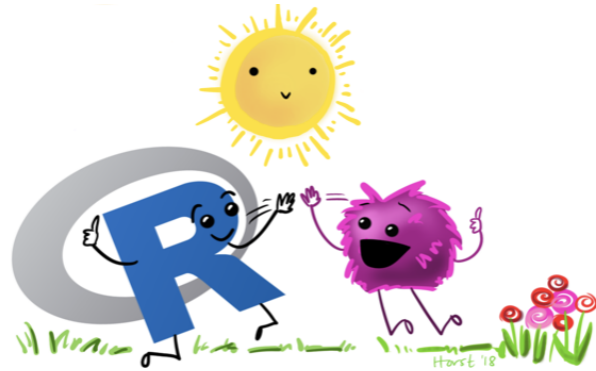


Illustration
by Allison
Horst