R for the Excel User

Julie Lowndes & Allison Horst

2019-09-26

Contents

1 Prerequisites **7**

2 Overview **11**

3 `readxl` **13**
- 3.1 Summary (a few sentences) . . . 13
- 3.2 Objectives (more detailed, bulletpoints?) . . . 13
- 3.3 Resources . . . 14
- 3.4 Lessons teaching for each objective..... (objectives, examples) . . 14
- 3.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) . . . 14
- 3.6 Interludes (deep thoughts/openscapes) . . . 14
- 3.7 Our Turn Your Turn 1 . . . 14
- 3.8 Our Turn Your Turn 2 . . . 14
- 3.9 Efficiency Tips . . . 14

4 RMarkdown **15**
- 4.1 Summary (a few sentences) . . . 15
- 4.2 Objectives (more detailed, bulletpoints?) . . . 15
- 4.3 Resources . . . 15
- 4.4 Lessons teaching for each objective..... (objectives, examples) . . 15
- 4.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) . . . 15
- 4.6 Interludes (deep thoughts/openscapes) . . . 15
- 4.7 Our Turn Your Turn 1 . . . 15
- 4.8 Our Turn Your Turn 2 . . . 15
- 4.9 Efficiency Tips . . . 15

5 Dplyr and Pivot Tables **17**
- 5.1 Summary (a few sentences) . . . 17
- 5.2 Objectives (more detailed, bulletpoints?) . . . 17
- 5.3 Resources . . . 17
- 5.4 Lessons teaching for each objective..... (objectives, examples) . . 17

5.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) 17
5.6 Interludes (deep thoughts/openscapes) 17
5.7 Our Turn Your Turn 1 17
5.8 Our Turn Your Turn 2 17
5.9 Efficiency Tips 17

6 Dplyr and vlookups **19**
6.1 Summary (a few sentences) 19
6.2 Objectives (more detailed, bulletpoints?) 19
6.3 Resources 19
6.4 Lessons teaching for each objective..... (objectives, examples) . . 19
6.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) 19
6.6 Interludes (deep thoughts/openscapes) 19
6.7 Our Turn Your Turn 1 19
6.8 Our Turn Your Turn 2 19
6.9 Efficiency Tips 19

7 Tidying **21**
7.1 Better practices [needs a better name] 21
7.2 Summary (a few sentences) 21
7.3 Objectives (more detailed, bulletpoints?) 21
7.4 Resources 22
7.5 Lessons teaching for each objective..... (objectives, examples) . . 22
7.6 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) 23
7.7 Interludes (deep thoughts/openscapes) 23
7.8 Our Turn Your Turn 2 23
7.9 Efficiency Tips 23

8 Formatting and Sharing **25**
8.1 Summary (a few sentences) 25
8.2 Objectives (more detailed, bulletpoints?) 25
8.3 Resources 25
8.4 Lessons teaching for each objective..... (objectives, examples) . . 25
8.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) 25
8.6 Interludes (deep thoughts/openscapes) 25
8.7 Our Turn Your Turn 1 25
8.8 Our Turn Your Turn 2 25
8.9 Efficiency Tips 25

9 Synthesis **27**
9.1 Summary (a few sentences) 27
9.2 Objectives (more detailed, bulletpoints?) 27

9.3 Resources 27
9.4 Lessons teaching for each objective..... (objectives, examples) . . 27
9.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections) 27
9.6 Interludes (deep thoughts/openscapes) 27
9.7 Our Turn Your Turn 1 27
9.8 Our Turn Your Turn 2 27
9.9 Efficiency Tips 27

Chapter 1

Prerequisites

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports, e.g., a math equation $a^2 + b^2 = c^2$.

The **bookdown** package can be installed from CRAN or Github:

```
install.packages("bookdown")
# or the development version
# devtools::install_github("rstudio/bookdown")
```

Remember each Rmd file contains one and only one chapter, and a chapter is defined by the first-level heading `#`.

To compile this example to PDF, you need XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): https://yihui.name/tinytex/.

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter 3.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 1.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 1.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

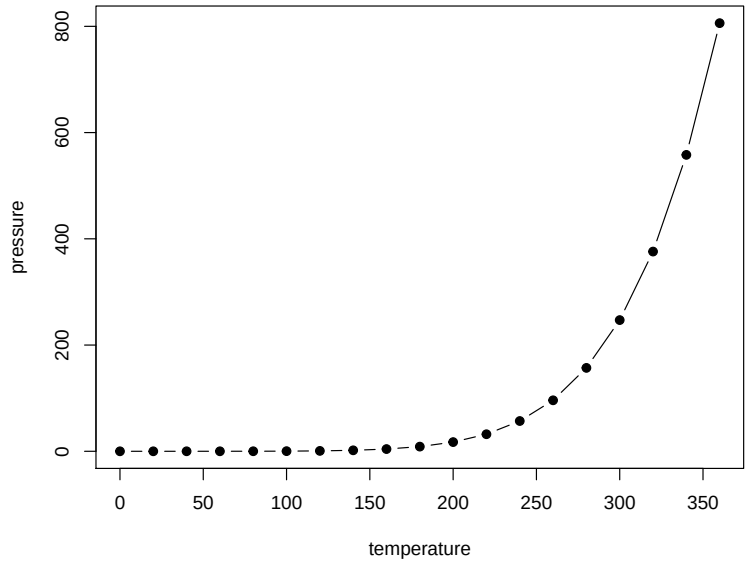

Figure 1.1: Here is a nice figure!

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2019) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

Table 1.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 2

Overview

References Brainstorm

- Broman & Woo 2017: https://www.tandfonline.com/doi/full/10.1080/00031305.2017.1375989

Chapter 3

`readxl`

Note for Allison: the `{#readxl}` above lets us reference it from another chapter…not sure how much we'll use it but it's possible (and doesn't hurt to have). Cool, thanks Julie!

3.1 Summary (a few sentences)

The **readxl** package makes it easy to import tabular data from Excel spreadsheets (.xls or .xlsx files) and includes several options for cleaning data during import. **readxl** has no external dependencies and functions on any operating system, making it an OS- and user-friendly package that simplifies getting your data from Excel into R.

3.2 Objectives (more detailed, bulletpoints?)

1. Use readxl::read_excel() to read in an Excel worksheet
2. Use readxl::read_excel() to read in parts of a worksheet (by cell range)
3. Specify column names when importing Excel data
4. Replace a specific string/value with 'NA'
5. Skip n rows when importing an Excel worksheet
6. Read in multiple Excel worksheets (note to Julie: I actually think this is fine to include…purrr::map() can be explained in a really simple way)
7. Something about readxl workflows (e.g. %>% write_csv() to keep a copy of the raw data in a project)??

3.3 Resources

- https://readxl.tidyverse.org/

3.4 Lessons teaching for each objective..... (objectives, examples)

3.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

3.6 Interludes (deep thoughts/openscapes)

3.7 Our Turn Your Turn 1

3.8 Our Turn Your Turn 2

3.9 Efficiency Tips

Chapter 4

RMarkdown

4.1 Summary (a few sentences)

4.2 Objectives (more detailed, bulletpoints?)

4.3 Resources

4.4 Lessons teaching for each objective..... (objectives, examples)

4.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

4.6 Interludes (deep thoughts/openscapes)

4.7 Our Turn Your Turn 1

4.8 Our Turn Your Turn 2

4.9 Efficiency Tips

Chapter 5

Dplyr and Pivot Tables

5.1 Summary (a few sentences)

5.2 Objectives (more detailed, bulletpoints?)

5.3 Resources

5.4 Lessons teaching for each objective..... (objectives, examples)

5.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

5.6 Interludes (deep thoughts/openscapes)

5.7 Our Turn Your Turn 1

5.8 Our Turn Your Turn 2

5.9 Efficiency Tips

Chapter 6

Dplyr and vlookups

6.1 Summary (a few sentences)

6.2 Objectives (more detailed, bulletpoints?)

6.3 Resources

6.4 Lessons teaching for each objective..... (objectives, examples)

6.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

6.6 Interludes (deep thoughts/openscapes)

6.7 Our Turn Your Turn 1

6.8 Our Turn Your Turn 2

6.9 Efficiency Tips

Chapter 7

Tidying

7.1 Better practices [needs a better name]

How to be a nimble useR Modern useRs are nimble internet useRs something clever about cleaning I am the worst at naming things

7.2 Summary (a few sentences)

R ecosystem evolves and improves due to contributed work by the community, and this is a good thing. Being a nimble useR means being able to navigate/keep tabs on this ecosystem and find what you need. It also means working reproducibly, so you can re-run and update things more easily. Here we will teach you how to expect things and help yourself. Pay attention to urls.

7.3 Objectives (more detailed, bulletpoints?)

- expect there is a better way, how and where to look (20 mins)
  - CRAN
  - Twitter #rstats
  - rOpenSci
  - RStudio
  - Example: how to Google.
- hands-on with janitor (30+ mins)
  - discovery and quality assurance
  - installing from GitHub
  - big payoff for little effort

- hands-on with another excel-useful example: skimr?

- reproducibility (20 mins)
  - it's important, scripted

7.4 Resources

- Wilson et al. 2014 "Good enough practices"

7.5 Lessons teaching for each objective..... (objectives, examples)

7.5.1 Expect there's a better way chat

- give time for them to google?

7.5.2 Janitor

janitor & other things that will make your life easier with limited effort Janitor: up till now the column names have been fine. Until now.

7.5.2.1 Our turn your turn

Walk through and example and leave our code up, and have you do it but clean another dataset. Work with a neighbor.

7.5.3 Example: How to Google

Pay attention to URLs, build github/rmarkdown savviness (ex: raw.githubusercontent.com)

- I read this blog: https://blog.revolutionanalytics.com/2018/08/how-to-use-r-with-excel.html
- I've never heard of click on `openxlsx`, what is it
- Takes me here https://www.rdocumentation.org/packages/openxlsx/versions/4.1.0.1, but I want more info. How recently was it worked on? Does it interface with tidyverse? Click on "news"
- Takes me here. https://raw.githubusercontent.com/awalker89/openxlsx/master/NEWS . Not useful. But from this URL, - I see the username so I can edit this url to be https://github.com/awalker89/openxlsx/

- 1st thing: most recent commit was a year ago. Can poke around more, are there issues open, are they taken care of? Etc. I will probably not pursue using this right now. But good to have learned about it.

7.6 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

7.7 Interludes (deep thoughts/openscapes)

7.8 Our Turn Your Turn 2

7.9 Efficiency Tips

- browser efficiency tips
  - Rmd/github anchors for urls
  - press command to open a new tab

Reproducibility is important (this might be new to some people) Example: run everything start to finish and then closing it all and trying to do again In excel Vs R If your computer shuts off are you nervous to close it? Recreate it "What they didn't forget to teach you about R" WTDF. uncool

Chapter 8

Formatting and Sharing

8.1 Summary (a few sentences)

8.2 Objectives (more detailed, bulletpoints?)

8.3 Resources

8.4 Lessons teaching for each objective..... (objectives, examples)

8.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

8.6 Interludes (deep thoughts/openscapes)

8.7 Our Turn Your Turn 1

8.8 Our Turn Your Turn 2

8.9 Efficiency Tips

Chapter 9

Synthesis

9.1 Summary (a few sentences)

9.2 Objectives (more detailed, bulletpoints?)

9.3 Resources

9.4 Lessons teaching for each objective..... (objectives, examples)

9.5 Fun facts (quirky things) - making a note of these wherever possible for interest (little "Did you know?" sections)

9.6 Interludes (deep thoughts/openscapes)

9.7 Our Turn Your Turn 1

9.8 Our Turn Your Turn 2

9.9 Efficiency Tips

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2019). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.11.