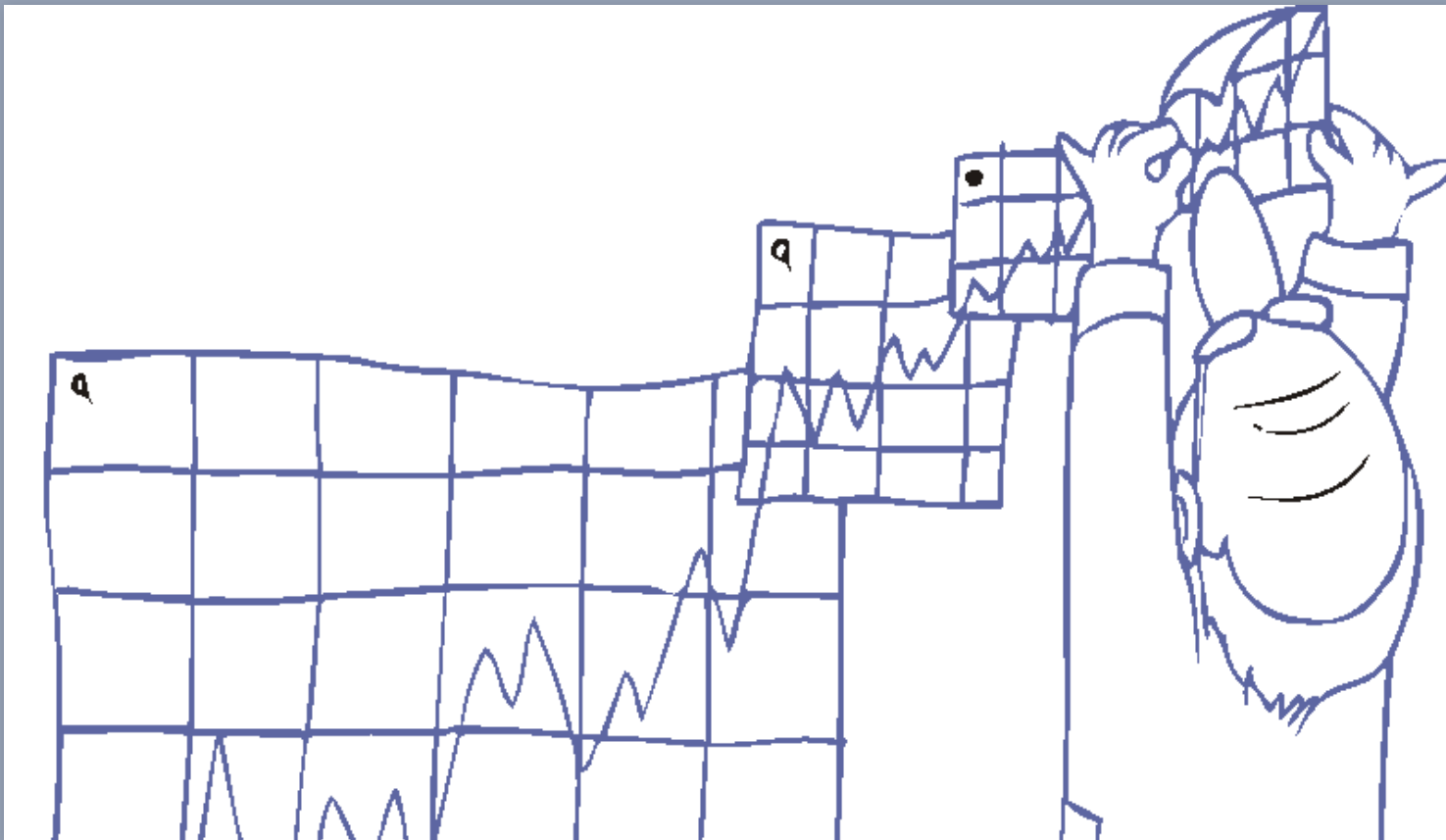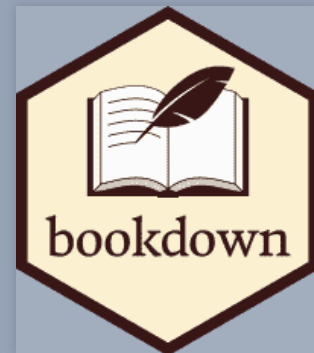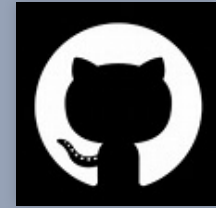# R from Start to Finish: A Case Study

Overview

- GitHub & R project

- feather: "a fast, lightweight, and easy-to-use binary file format for storing data frames"

  short-term as part of analysis[1]

- flextable
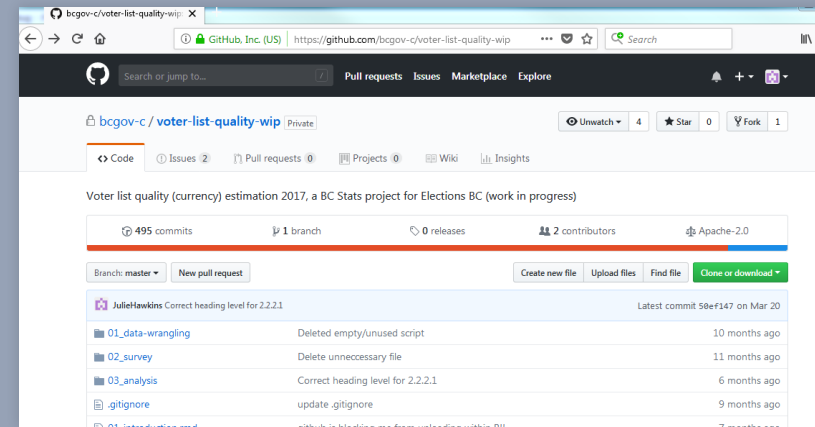
- bookdown @ https://bookdown.org/

1: https://blog.cloudera.com/blog/2016/03/feather-a-fast-on-disk-format-for-data-frames-for-r-and-python-powered-by-apache-arrow/

GitHub & R project

- connected GitHub repo to an R project to keep files together and organized

- used .gitignore file to ensure that confidential data files did not push to GitHub

- provided version control and issue notification for project with more than one contributor

Feather

- project required working with massive datasets (@1.1 GB text file)

  - needed to sample from N = 1,168,572; reading in 450 MB text file with read_csv took @9 minutes

  - post-survey, read in 551 MB worked on csv file (took @12 minutes)

  - read in 740 MB text file with read_csv (took @15 minutes)

  - merge of these two files with some additional information took HOURS to write_csv

- `library(feather)`

Flextable

@ https://cran.r-project.org/web/packages/flextable/vignettes/overview.html
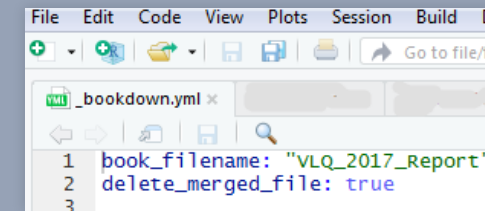
Bookdown

@ https://bookdown.org/yihui/bookdown/

- install the development versions of `bookdown` from GitHub: `devtools::install_github("rstudio/bookdown")`

- `library(bookdown)` built on R Markdown and knitr

- `bookdown` (and `knitr`) package author: Yihui Xie, a software engineer at RStudio

- project required Word document output

  - `bookdown` meant to output formats such as HTML, PDF, and e-books
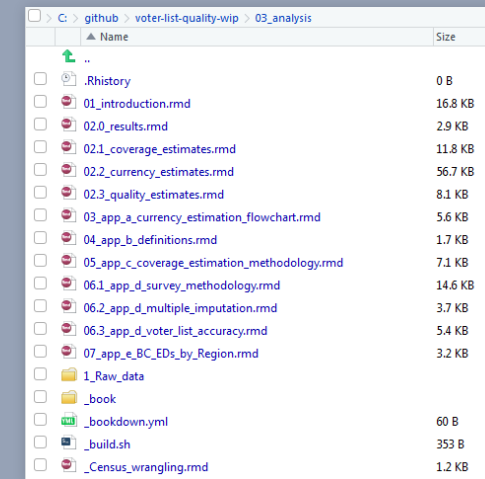
Bookdown Learnings

-

- site: "bookdown::bookdown_site" is REQUIRED to knit all rmds together

- you must have either an `index.rmd` OR `_bookdown.yml` file for bookdown to work

- `bookdown` will include all .rmd files that are not named starting with an underscore

- number the rmd files for `bookdown` to knit them in order

- each .rmd file that you want in the book must have one and only one chapter

Conclusion

Thanks!

Any questions?

Julie.Hawkins@gov.bc.ca