Final Project

Navigate to 06_final project

Open variation.png

Open hirschey.Rmd

Knit the document, and review the output

Final Project Template

Navigate to 06_final project

Open 06_final_template.Rmd

Final Project Assignment

3 Documents

1. last_name.Rmd with full code

2. Knitted report

3. image.png

Final Presentation

2-3 minute oral presentation:

- Review the data
- Output file
- Salient observations (conclusion)
- Strategy to test the hypothesis

Final Presentation Feedback

You will receive feedback on your oral presentation based on:

- Content: Comments on the data and insight that were presented
- Visuals: Any ways the visuals could have been more clear?
- Mechanics of presentation: Comments on the student presentation
- Overall

Final Grades

As described in the syllabus,

“A final grade is an important metric summarizing your fluency and aptitude with the concepts presented in class. In this course, I will ask each student to be the judge of these criteria and submit their proposed grade and a justification for it after the conclusion of the class. The following metrics should be considered when proposing a grade:

40% Code, R markdown file, and summary interpretation of findings

30% In-class presentation of above findings

20% Help vignette

5% Attendance

5% Participation

While the instructors will ultimately determine the final grade for each student, thoughtful reflection and justification for a letter grade provided by each student will be heavily weighted. Self-critique is an important skill for scientists, and this class presents and opportunity for it."

Concluding Thoughts

Data science enables **data-driven** information gathering and hypothesis generation

➡Scientific Research

➡Reviews

Data science enables the ability to ask new types of questions

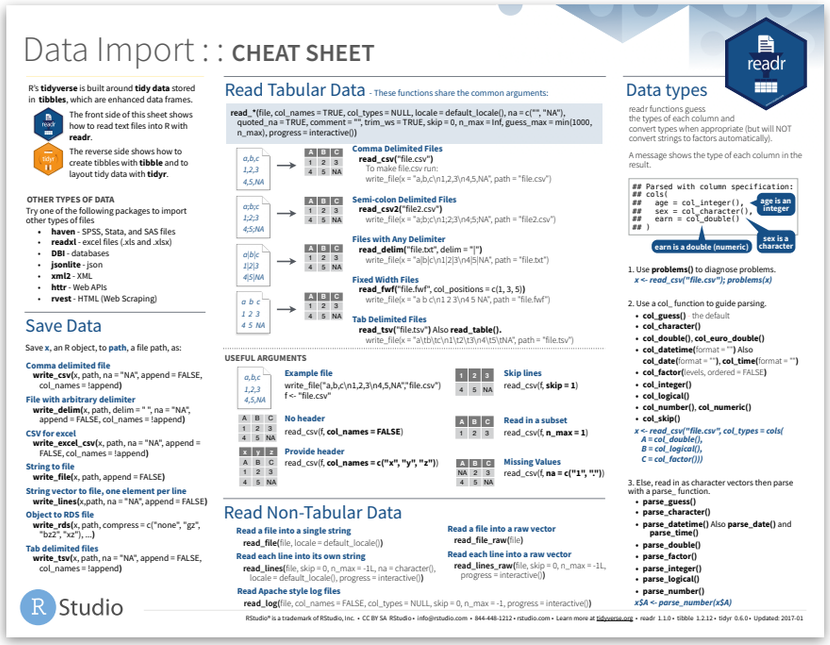
Process centric, not necessarily gene centric

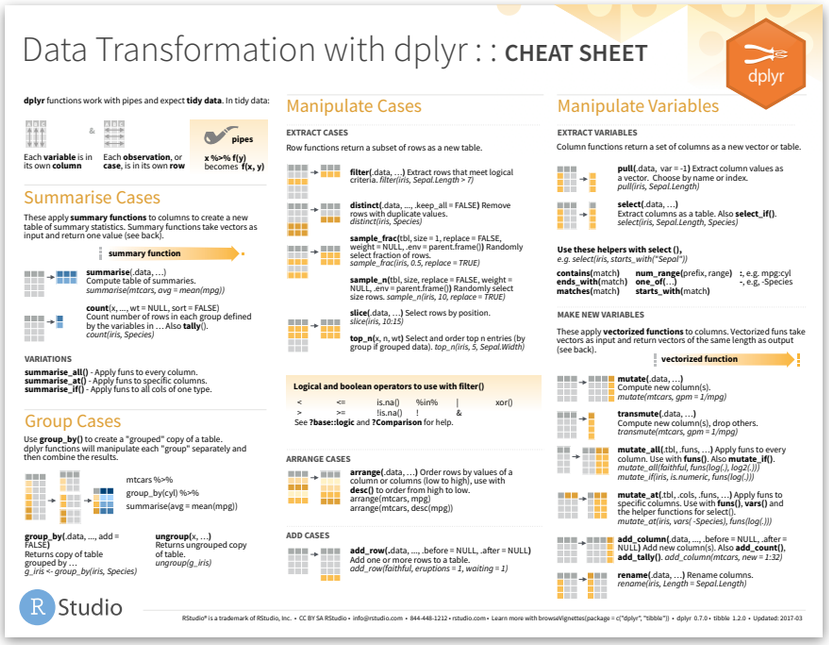
Making things computable makes them actionable at zero marginal cost.

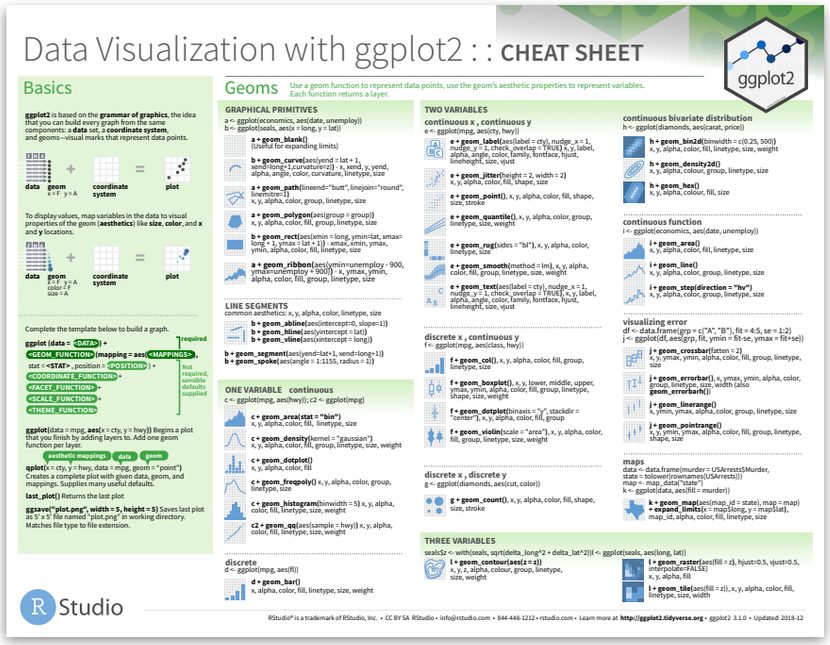
Workflows save time, achieve reproducibility

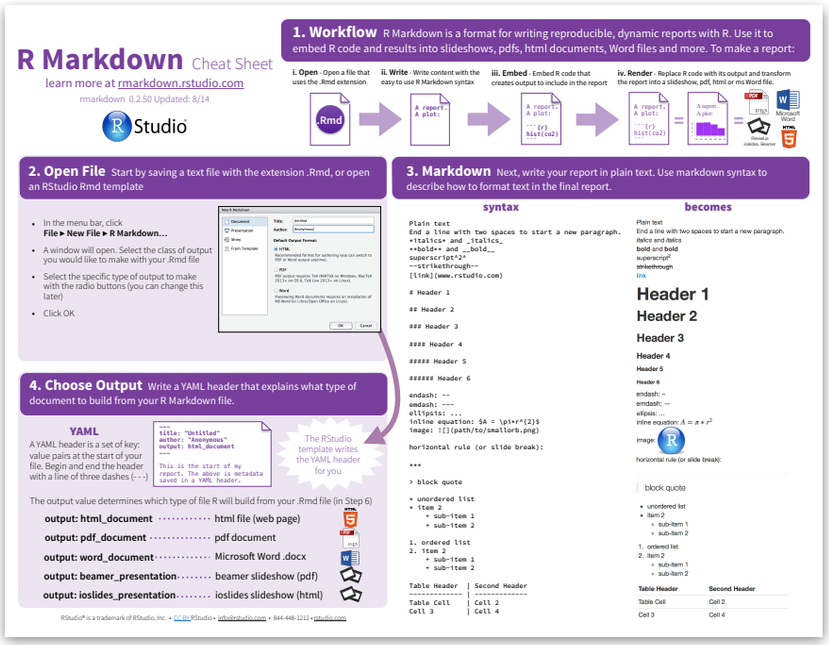
Resources: Cheat Sheets

Cheat sheets make it easy to learn about and quickly refer to function in some of the common packages.

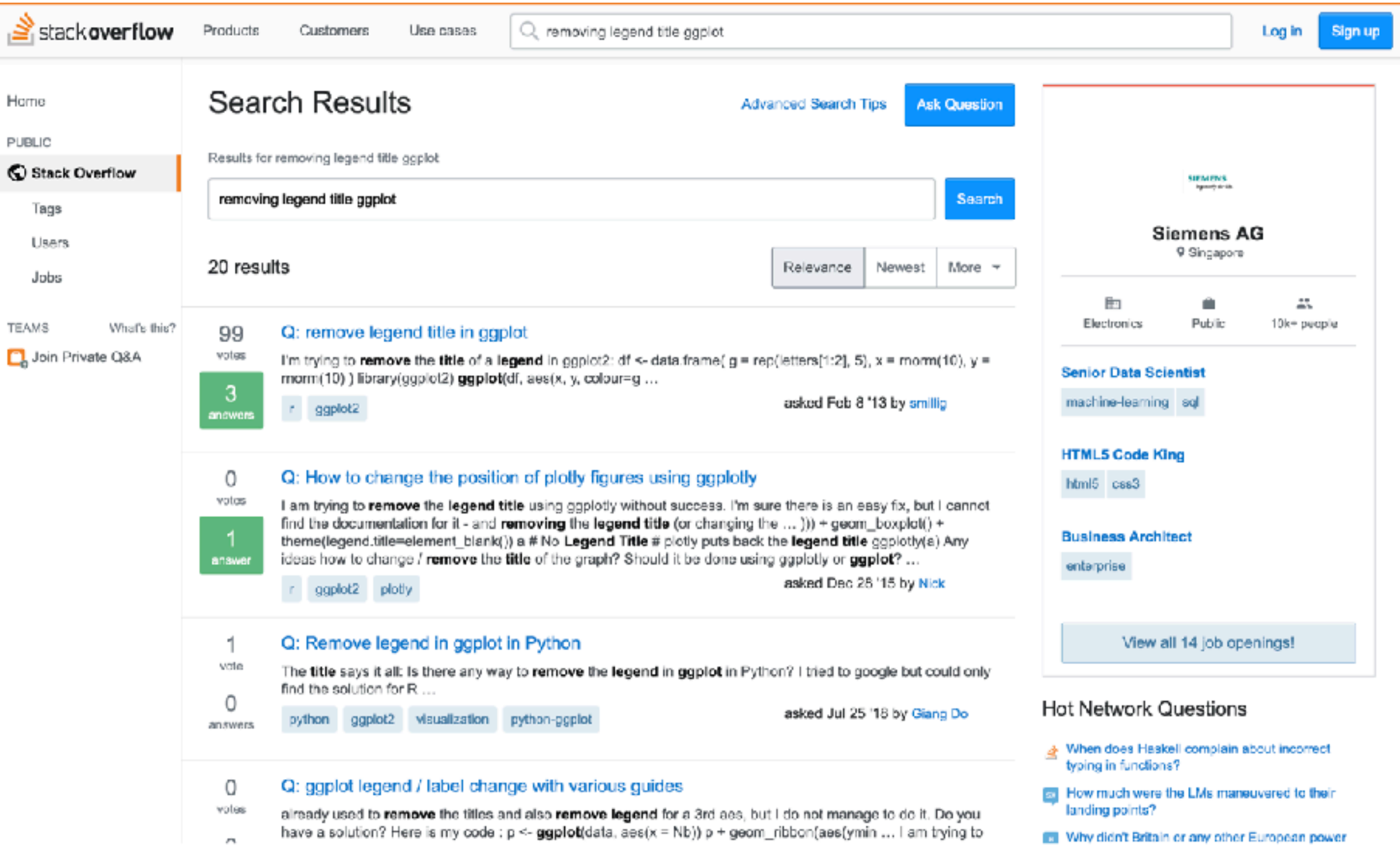








Resources: Stackoverflow



https://r4ds.had.co.nz/index.html

Resources

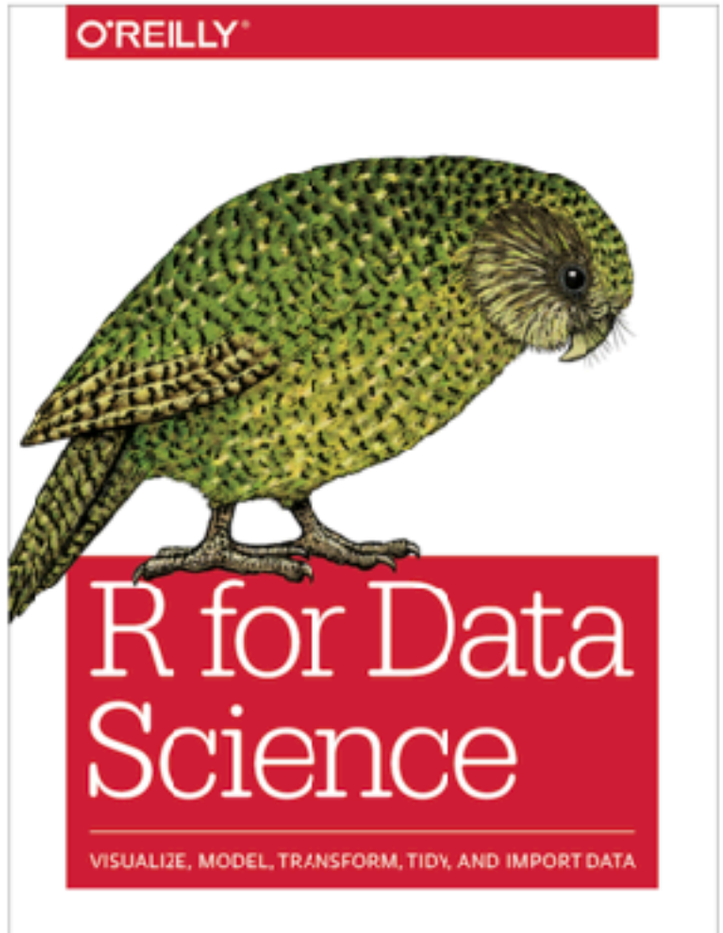
R for Data Science

Garrett Grolemund

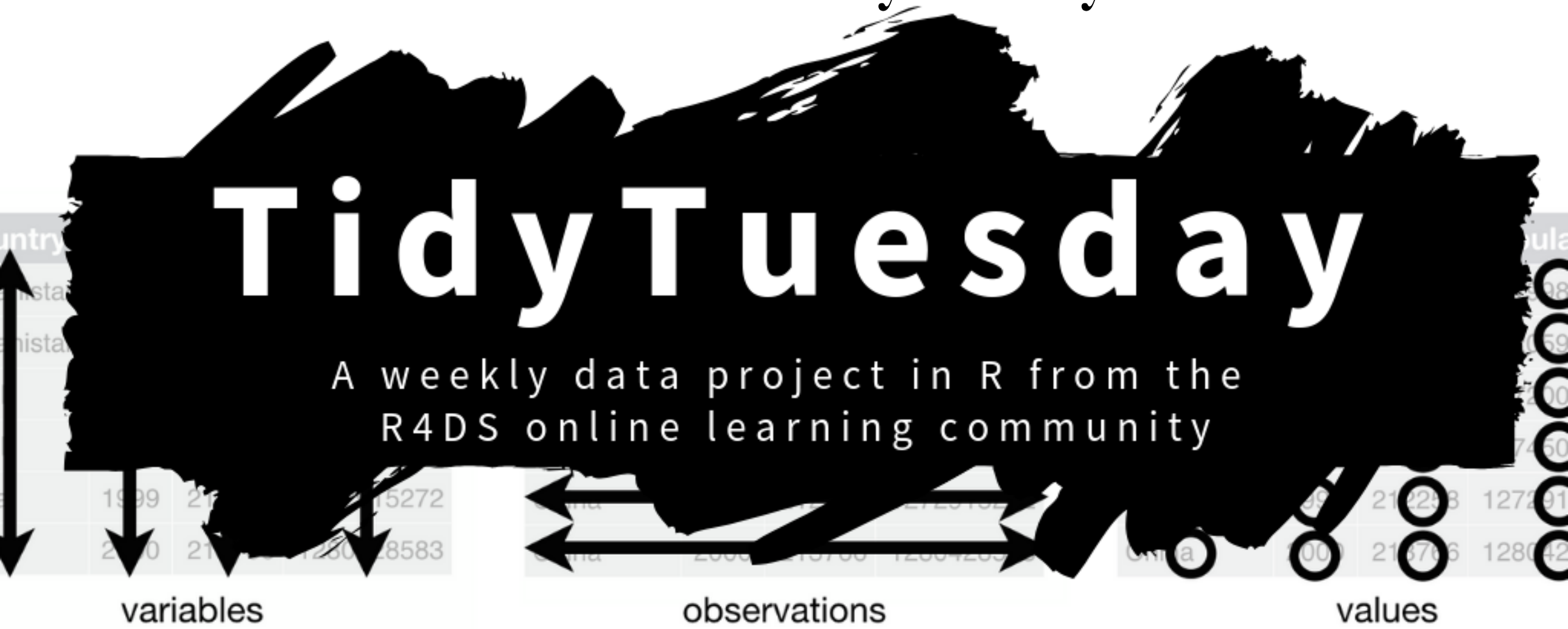
Hadley Wickham

Welcome

This is the website for **"R for Data Science"**. This book will teach you how to do data science with R: You'll learn how to get your data into R, get it into the most useful structure, transform it, visualise it and model it. In this book, you will find a practicum of skills for data science. Just as a chemist learns how to clean test tubes and stock a lab, you'll learn how to clean data and draw plots —and many other things besides. These are the skills that allow data science to happen, and here you will find the best practices for doing each of these things with R. You'll learn how to use the grammar of graphics, literate programming, and reproducible research to save time. You'll also learn how to manage cognitive resources to facilitate discoveries when wrangling, visualising, and exploring data.



Resources: #tidytuesday



rfordatascience / tidytuesday

Unwatch 258 | Unstar 960 | Fork 457

Code | Issues 48 | Pull requests 0 | Projects 0 | Wiki | Security | Insights

Official repo for the #tidytuesday project

380 commits | 2 branches | 11 releases | 8 contributors | MIT

Branch: master | New pull request | Create new file | Upload files | Find File | Clone or download

Acknowledgements

Teaching Assistants

Allie Mills, Ph.D.

Akshay Bareja, D.Phil.

Inspiration, ideas, packages, code

#R4DS (Garrett Grolemund and Hadley Wickham)

Mine Çetinkaya-Rundel (datasciencebox.org)

Chester Ismay and Albert Y. Kim (Modern Dive)

Garrett Grolemund (Remastering the Tidyverse)

Tidyverse devs and community

Rstudio

@matthewhirschey

www.hirscheylab.org

github.com/hirscheylab