PUBLICATION-READY DOCUMENTS USING R

Materials can be found at

https://github.com/kristenbhunter/presentations/tree/master/2025/ORNL

Kristen Hunter

April 2, 2025

ABOUT ME

- Lecturer in Statistics and Data Science
- University of New South Wales, Sydney, Australia
- Enthusiastic R user
- Research interests: experimental design, causal inference, environmental science & policy
- Always happy to connect with people!

WHY USE R MARKDOWN?

- reproducibility
- easy integration of R code
- easy to update plots and tables
- version control & collaboration through github

RTICLES PACKAGE

- R markdown templates for a variety of journals and publishers
- What if your intended journal isn't listed?
  - open source: add the template yourself!
  - LaTeX fragment

SETTING UP YOUR DOCUMENT

- useful defaults
- can be overwritten for a particular chunk

```
knitr::opts_chunk$set(
  cache = TRUE,
  warning = FALSE,
  message = FALSE,
  fig.height = 4,
  fig.width = 4,
  fig.align = "center"
)
```

SETTING UP YOUR DOCUMENT

- set a seed
- set `kable` options
  - `options(knitr.kable.max_rows = 30)`
  - `options(knitr.kable.NA = '')`
- set a ggplot theme

GOOD PRACTICE: CHECK YOURSELF

- Periodically clear cache
- Turn on all warnings and errors
- Make sure it still compiles!
- Consider changing the seed

RESTRAIN YOURSELF

You can in theory write a combination of:

- markdown
- LaTeX
- CSS
- HTML
- R

Example: tables

- R
  - `kable(data)`
- HTML (and CSS)
  - `column 1 | column 2`
- LaTeX table
  - `\begin{table}`

RESTRAIN YOURSELF

My personal recommendations

- kable for tables
- Default to LaTeX for everything else
- Avoid HTML and CSS for PDF documents
- Most customizable and powerful

Easy LaTeX features

- Referencing other document sections dynamically (e.g. In Section 2, we discuss…)
- Easy to change citation format
- Easy to change document-level formatting

GOOD CODING PRACTICE

One output per chunk

- Each figure, table, or output should have its own chunk

Name your chunks

- Helps with debugging and identifying slow points
- Saves out figures with meaningful names

DEMOS

- ORCID
- Citations using natbib
- Beautiful tables
- Referencing tables
- Referencing figures
- Generating tables in a loop
- Stargazer

DEBUGGING: STEP ONE

- Add `keep_tex: true` to YAML header
- Use a LaTeX distribution to debug the .tex file directly to give you the line number

```
output:
  pdf_document:
    keep_tex: true
```

- Never hurts to delete all the generated files (.tex, .aux, .log, .etc)

DEBUGGING: WHEN ALL ELSE FAILS

Two strategies:

- Top-down: Start with all code chunks. Remove one chunk at a time until it knits successfully.
- Bottom-up: Start with no chunks. Add one chunk at a time until it knits successfully.

EXTRA NOTES

- Advanced: You can re-use bits of R markdown across multiple files.
  - See this blog post for more info.
- R markdown and Quarto also produce word documents
- You can write journal articles in quarto, but infrastructure is less developed

REFERENCES

Materials

- https://github.com/kristenbhunter/presentations/tree/master/2025/ORNL

Original blog post

- CARES Blog

REFERENCES

Useful websites

- rticles package
- gallery of rticles templates
- chunk options
- gallery of ggplot themes
- quarto for scientists
- reproducible publishing with quarto
- quarto journal templates