

Turning your PhD into a package

Ella Kaye

Warwick YRM, 21st Nov 2017

Why?

- Keep all your data/code/analysis together and easily accessible
- Reproducible
- Easy to share (e.g. with supervisors)
- It's easier than you may think!

What I'm going to talk about

- How to set up a (PhD) **package**
- Basic package structure
- devtools and usethis
- bookdown for authoring vignettes/reports
- What I've done for my PhD

What I'm not going to talk about*

*but which you should definitely consider as part of a package workflow

- git/GitHub
- packrat

Setting up a package (demo)

- Use the latest version of RStudio (1.1.338 or higher)
- File > New Project > New Directory > R Package

devtools and usethis

- devtools is a set of package development tools primarily written by Hadley Wickham to make it easy for Hadley Wickham to write packages
- usethis automates many common package and analysis setup tasks (splits off from devtools)
- library(devtools) (first)
- library(usethis) (second)

Data (demo)

- `usethis::use_data_raw`
- `usethis::use_data`
- `devtools::load_all()` (Shift + Cmd + L)

Document your data (demo)

- data.R (in R folder)
- Can now **use markdown**. Add the following to the DESCRIPTION:
Roxygen: list(markdown = TRUE)
- devtools::document()

In data-raw folder

- Data creations scripts
- Raw data (though these don't have to be in the package)

In R folder

- Data import functions (for use in data creation scripts)
- `data_documentation.R` for documenting your data
- Analysis functions
 - Be sure to document these too!

Markdown

- **Markdown:**
 - sections, lists, links, images
- **Pandoc:**
 - adds footnotes, **citations**, tables
 - raw *L^AT_EX* (including the ability to use *L^AT_EX* packages)
 - makes it possible to generate HTML, LaTeX/pdf, Word and slides
- **R Markdown:**
 - Uses Pandoc markdown
 - adds code chunks

bookdown

A package by Yihui Xie that **extends** R Markdown with features that actually make it feasible to write a statistics PhD.

- Number and reference equations
- Theorems and proofs
- Special headers
 - (PART)
 - (APPENDIX)
- Text references

(show bookdown_examples.Rmd)

bookdown

A package by Yihui Xie that **extends** R Markdown with features that actually make it feasible to write a statistics PhD.

- Number and reference equations
- Theorems and proofs
- Special headers
 - (PART)
 - (APPENDIX)
- Text references

(show bookdown_examples.Rmd)

bookdown can be used to bring a thesis together

- Typically, each chapter is a separate .Rmd file
- Switch between HTML and pdf

bookdown

A package by Yihui Xie that **extends** R Markdown with features that actually make it feasible to write a statistics PhD.

- Number and reference equations
- Theorems and proofs
- Special headers
 - (PART)
 - (APPENDIX)
- Text references

(show bookdown_examples.Rmd)

bookdown can be used to bring a thesis together

- Typically, each chapter is a separate .Rmd file
- Switch between HTML and pdf

Goodbye *L^AT_EX*!

Analysis as vignettes

Keeps everything together.

- `usethis::use_vignette()`
- modify the `.Rmd` file
- `devtools::build_vignettes()`
- Build and reload
- `browseVignettes("package_name")`

Analysis as vignettes

Keeps everything together.

- `usethis::use_vignette()`
- modify the `.Rmd` file
- `devtools::build_vignettes()`
- Build and reload
- `browseVignettes("package_name")`

Downsides

- Knitting only gets you a temp html file
- Building vignettes can be time-consuming
- Have to build all the vignettes together

Analysis as separate files

- Make a 'reports' folder (can use Terminal in RStudio)
- `usethis::use_build_ignore("reports")`
- create .Rmd files and knit as usual

A look at my Admissions package

Main links

R packages: An (online) book on R packages by Hadley Wickham

R packages cheatsheet: Also available within RStudio from Help > Cheatsheets

bookdown: A bookdown generated book about bookdown by Yihui Xie

Happy git with R: Superb bookdown book on git and GitHub for the useR.