Reproducible Documents with `{rmarkdown}`

Some general tips and good practice

Instructor: Selina Baldauf

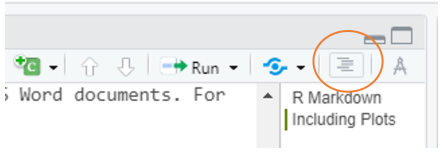
Freie Universität Berlin - Theoretical Ecology

2022-22-03 (updated: 2022-09-11)

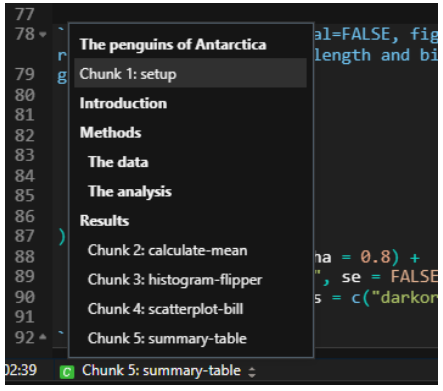
Tip 1: Keep your document clean

`.Rmd` documents can quickly become large and messy. To keep them clean, you can

- Use headers to mark sections in your document
- Navigate the file using the document outline



- Use names for your code chunks
- Navigate code chunks with the code and document outline (bottom left of script)



Tip 2: Source large data preparation scripts

- Related to Tip 1
- If it's not necessary for the document, do data preparation in a separate R Script
- Place that R Script in the project where the `.Rmd` is located
- Then source the script in a code chunk:

````
```{r prepare-data, warning=TRUE, message=TRUE}
source("path/to/script.R")
```
````

- This runs all the R code in `script.R` and loads the results into the `.Rmd` document

Tip 3: Split larger documents into multiple `.Rmd` files

- Related to Tip 1
- Write separate `.Rmd` files e.g. for Introduction, Methods and Results
- Have on main `.Rmd` file that
  - Combines the sections into one
  - Controls YAML options of the output
- You can load an `.Rmd` file into another one using the `child` chunk option

````
```{r load-child, child="path/to/child.Rmd"}
```
````

Tip 3: Split larger documents into multiple `.Rmd` files

- 3 separate files `Introduction.Rmd`, `Methods.Rmd`, `Results.Rmd`
- The separate files control everything that happens on the lower levels of the documents, e.g.

````
## First results

```{r result-plot, fig.width=3}
plot(1:10, 1:10)
```
````

- `Main.Rmd` (see right) controls
  - YAML options
  - Global setup options
  - Includes the sections via the `child` option

````
---
title: "My paper"
author: "Selina Baldauf"
output:
  pdf_document:
    toc: true
---

```{r global-setup, include = FALSE}
knitr::opts_chunk$set(echo = FALSE)
```

# Introduction

```{r intro, child="Introduction.Rmd"}
```

# Methods

```{r methods, child="Methods.Rmd"}
```

# Results

```{r results, child="Results.Rmd"}
```
````

Tip 4: Read through some online resources

- Read or scroll through some R Markdown books or tutorials to
    - See what is possible with R Markdown
    - Find thing that are relevant for your own documents
- I recommend to start with the two books:
    - R Markdown Cookbook
    - R Markdown - The Definitve Guide
- You can also find some resources on the workshop website

Tip 5: Use a project oriented workflow

- One directory with all files relevant for project
  - Scripts, data, plots, documents, ...
- An RStudio project is just a normal directory with an `*.Rproj` file
- Advantages of using RStudio projects
  - Easy to navigate in R Studio (`File` pane)
  - Easy to find and access scripts in RStudio
  - Project root is working directory
  - Open multiple projects simultaneously in separate RStudio instances

```
Project
|
|- data
|
|- doc
|    |
|    |- analysis.Rmd
|    |
|    |- publication.Rmd
|
|- analysis
|    |
|    |- clean_data.R
|    |
|    |- statistics.R
|
|- *.RProj
```

Example project structure

Tip 5: Use a project oriented workflow

Create a project from scratch:

1. `File -> New Project -> New Directory -> New Project`
2. Enter a directory name (this will be the name of your project)
3. Choose the Directory where the project should be initiated
4. `Create Project`