gesis Leibniz Institute for the Social Sciences

Open Methodology in practice: Reproduzierbare Forschung mit R

Dr. Tobias Heycke

27.06.2019

Part 2: R Markdown

Schedule

- Basic rules of Markdown
- The YAML Header
- R Markdown
- Papaja
- Citation
- Tables

Introduction

Course material

All materials can be found in the folder: material

Markdown

- Simple Markup Language
- Written in text files and translated to final output format
- To make it easier, we will also use RStudio as IDE (which is not necessary)
- R Markdown = Markdown + R

Writing in R Markdown (.Rmd)

- Easier than LaTeX
- Reproducible and commented R Code/Manuscript
- Less flexible than LaTeX

Basic Rules of Markdown

Line break I

- One line per sentence (helps also with version control)
- Use empty line as line break

Line break II

```
This is an example sentence.
This is annother example sentence.

This is an example sentence.

This is annother example sentence.
```

This is an example sentence. This is another example sentence.

This is an example sentence.

This is another example sentence.

Header

- Use ‘#’ to write headers
- Add empty line above and below the header
- The more ‘#’ the lower the level of the header

Header II

```
# Level 1 Header

## Level 2 Header

### Level 3 Header
```

Level 1 Header

Level 2 Header

Level 3 Header

Bold and Italic I

To write text in bold or in italic, just add *'s around the word(s)

- one * = italic
- two ** = bold

Bold and Italic II

```
You can print one word **bold** and another *italic*.

Or *multiple words* can **be altered**.
```

You can print one word **bold** and another word *italic*.

Or *multiple words* can **be altered**.

Lists I

- Use a dash for lists (or numbers)
- Make sure there is an empty line above and below the list

Lists II

```
- This
- Is
- A
- List
```

- This
- Is
- A
- List

Lists III

```
1. This
2. Is
3. Another
4. List
```

1. This
2. Is
3. Another
4. List

Overview rules (in RStudio)

For more see: `Help > Cheatsheets > R Markdown Reference Guide`

Task I

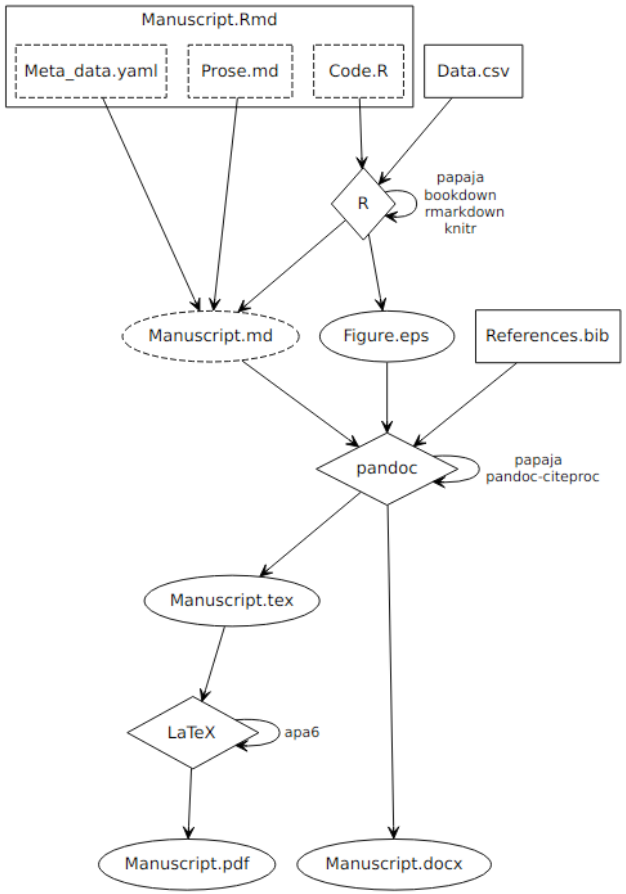
Go to hackmd.io > Click on ‘Use for free’ > Write Markdown!

R Markdown

What is possible?

- Scientific papers (which we will do later)
- Slides
- HTML sites
- CVs
- …

From Rmd to PDF



YAML header

Info YAML

- YAML = “YAML Ain’t Markup Language”
- Human readable configuration information
- At the top of each markdown document
- starts and ends with: ---

Simple YAML

```
---
title: "My first Markdown"
author: "My name"
output: pdf_document
---
```

Extended YAML

```
---
title: "GESIS Panel Wave Report"
output:
  pdf_document:
    keep_tex: yes
    includes:
      in_header: style/my_style.sty
  bibliography: literature.bib
fontsize: 11pt
params:
  wave: fe
  wrauthor: David Bretschi, Tobias Heycke
  date: September 2018
---
```

R Markdown

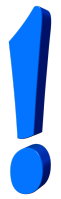
Your first Markdown (with the help of RStudio)

- R Studio is making it easier to knit a final document
- use UTF-8 encoding
- Let us take a look at a Rmd template by R Studio

Your first Markdown: general information

- A new environment is created when R markdown is knitted
- The working directory is always set to the folder of the .Rmd file
- You can knit Markdown documents by clicking on the knit button or pressing `Ctrl + Shift + k`.

Task II



Open the file: task02.Rmd and write some content and knit the file (and save the Rmd file)!

R Markdown

R and Markdown

- Change from prose to R code (and back)
- Easy to read and spot where results come from
- Many useful functions

Code chunks

R code chunks

- Code chunks include code (mostly R but python is possible)
- Code is executed when knitting a R Markdown file
- In R Studio you can use the shortcut `Ctrl + alt + i` to inset a code block
- No underscore in chunk name
- No name twice

````
```{r name, options=...}

```
````

Chunk options

- Useful options that will be used by `knitr`
- See handout for all options
- Check this for more details: https://yihui.name/knitr/options/

Useful chunk options

- eval = FALSE
- echo = FALSE (see next slide)
- warnings = FALSE
- cache = TRUE
- dependson = “chunkname”
- child = “rmdfilename.Rmd”

Global chunk options

Adjust all chunk options in a Rmd file with the following command:

```
knitr::opts_chunk$set(echo = TRUE)
```

In line R code

- Code can be run in the text
- Makes most sense to ‘print’ R output in the Text
- Result = Dynamic document
- In the text just write:

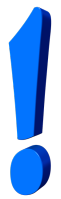
```
`r 2 + 2`
```

Task III

Open the file: task03.Rmd

- add a chunk calculating the sd the speed of cars (and save it to a new object)
- Print the results in a sentence below the chunk
- Knit the PDF
- Not happy with result? Round it!

Task IIIb



Add the task02.Rmd as a child to task03.Rmd and knit!

Try out cache and dependson!


LATEX

LaTeX I

You can add LaTeX code to your text.

- Use math mode for formulas etc.

```
You can simply add a formular: $\frac{SD}{\sqrt{N}}$
```

Result:

You can simply add a formula: $\frac{SD}{\sqrt{N}}$

LaTeX II

Things I often use:

```
\newpage
\includegraphics
\begin{center}
\end{center}
```

If you want to use a LaTeX package, load it in the YAML header (see later)

Math mode and R code

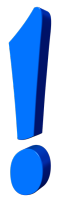
You can use R within math mode

```
sd.1 <- 15
n.1 <- 5
```

```
$\frac{`r sd.1`}{\sqrt{`r n.1`}}=
  `r round(sd.1/(sqrt(n.1)), 2)`$
```

$\frac{15}{\sqrt{5}} = 6.71$

Task IV



Add some LaTeX to your Rmd file!

For help with formulas, check out:
codecogs.com/latex/eqneditor.php and mathpix.com/

LaTeX in YAML

```
header-includes:
    - \usepackage{tikz}
    - \usepackage{graphicx}
    - \usepackage{titlesec}
    - \usepackage{xcolor}
```

papaja

papaja

- Template + filters + functions
- Helps writing APA6 articles
- Still in developement
- More information: https://github.com/crsh/papaja

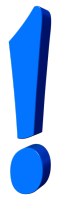
Installing papaja

In the R console run the following:

- install.packages("devtools")
- devtools::install_github("crsh/papaja", ref = "devel")

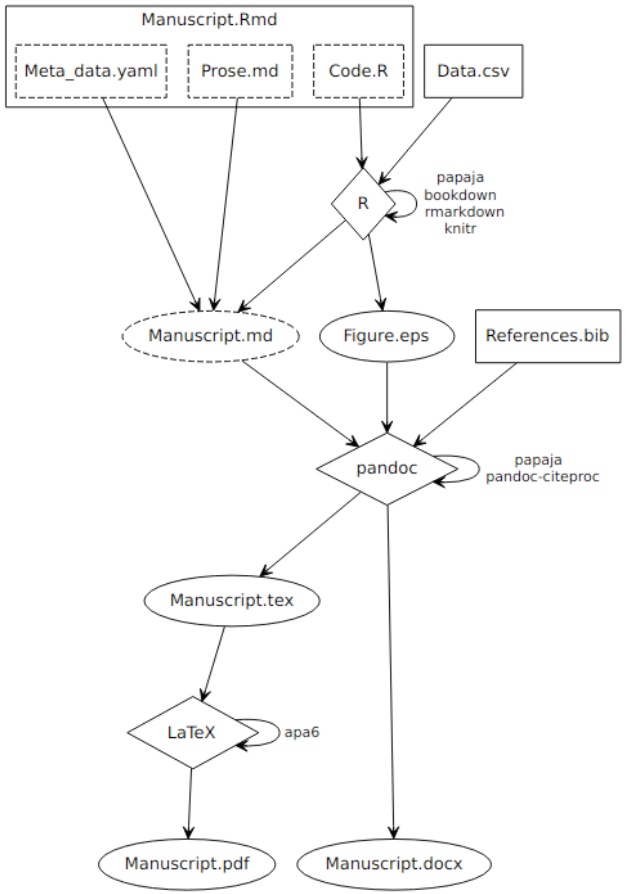
File > New File > R Markdown > From Template > APA article

Task V



- Set up a folder structure for one of your projects!
- Add some (not all) of your R code into a papaja document!
- Adjust the code, to standardize everything
- Knit the document!

From Rmd to PDF



Sources and citing

What you need to cite

- A biblatex file of your references (e.g., from Zotero)
- Add .bib file name to YAML

  `bibliography: nameoffile.bib`
- One way: Look up citation key and insert to text with `@` in front

Citr

Another way is using citr

- Save bib file and add to YAML
- Install citr (RStudio plug-in)
- Set citr shortcut
- More information: github.com/crsh/citr

(Even better: citr + better BibTeX; see link above)

Task VI

Add sources from the markdown_course.bib file to your text in the Rmd file!

(If you have your own bib files, you can use these)

Task VIb

Change the citation style to a style of your choosing!

Tables & Figures

Figures

- apa_barplot & apa_beeplot
- Very handy if you have a factorial design (compare to `afex`)
- In papaja, figures can be in text or below (see YAML)

Figure reference and caption

````
```{r some-fig, fig.cap = "Some text here. \\label{fig:figure1}"}
papaja::apa_barplot(
    data = npk
    , id = "block"
    , dv = "yield"
    , factors = c("N")
)
```

As can be seen in Figure \@ref(fig:figure1).
````

Tables: general information

- Tables can be produced with R code in R chunks (e.g., with kable or apa_table)
- Tables can be written in Markdown directly

apa_table

- Important: add results='asis' to chunk options

References to tables

````
```{r some-table, results='asis'}
my_table <- t(apply(cars, 2, function(x) # Create data
  round(c(Mean = mean(x), SD = sd(x), Min = min(x), Max = max(x)), 2)
))

apa_table(
  my_table
  , align = c("l", rep("r", 3))
  , caption = "\\label{tab:table1}A summary table of the cars dataset."
  , escape = FALSE
)
```

As can be seen in Table \@ref(tab:table1).
````

Task VII

Add a table and a figure to your Markdown document!

Don’t forget to refer to it in the text!

More papaja / markdown

Comments

You can add comments that will not be displayed in final PDF (or html)

```
<!-- this is a comment -->
```

Also useful for collaborating!

R Studio features

- Show document outline
- Shortcut to Chunks

Read data from git

From: https://gist.githubusercontent.com/crsh/bd4d1f62d300462ea0c0f44b9ad38616/raw/edd9c74e24b68f42433c2526cafc888509b1b8bc/batch_download_github.R

- Benefit: Folders not needed, sharing RMD alone can be enough
- Downside: Max 1000 files can be read from github (with this method)
- Downside: Might take a while

papaja to MS Word

You can change `output: papaja::apa6_pdf` to `output: papaja::apa6_word` to create MS word files from Rmd

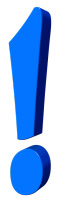
but:

- LaTeX does not work (e.g., math mode or `\@ref`)
- Tables will look different
- Generally not as pretty

For the future: Redoc

- Convert Rmd to Word and back
- Idea: also convert track changes and comments back
- More information: https://github.com/noamross/redoc
- Still in developement

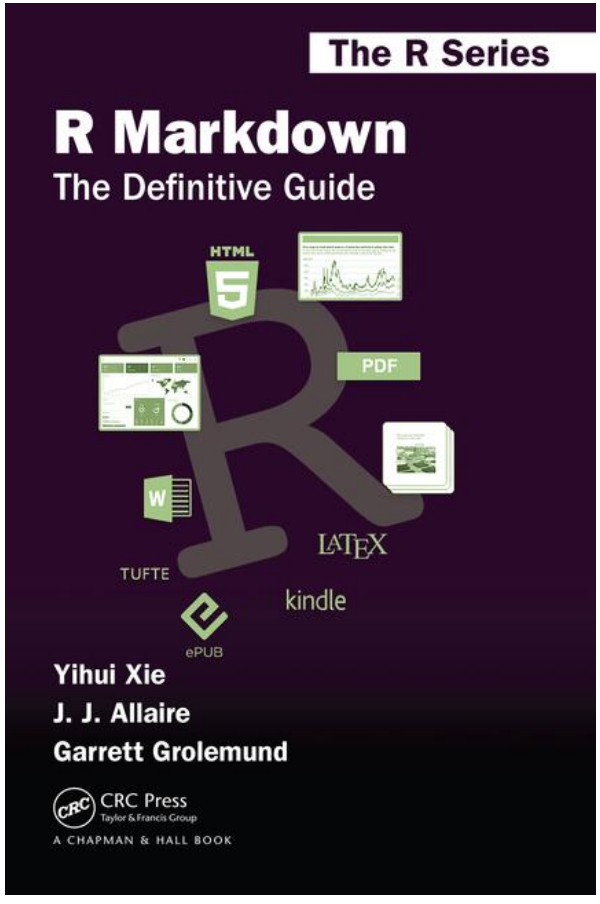
Task



Depending on time: Work on your Rmd file

Feel free to work on your Rmd at home tonight (not neccesary)

R Markdown Guide



https://bookdown.org/yihui/rmarkdown/

End of day 1