Essential Research Toolkit for the Humanities

Week 7: Creating reports with `RMarkdown` and `knitr`

Anna Pryslopska

May 23, 2022

Psycholinguistics and Cognitive Modeling Lab

Questions?

New environment, who dis?

Include the code to create the data frames you are using and the session info.

`adjectives_summary` ??

`adjectives_clean` ??

`mean_value` ??

`adjectives_clean1` ??

`adj_final` ??

`library(??)`

Create **three** different plot types

You have until the end of today to add the missing plots.

Write scripts and assign values

Do not copy the output of the console.

```
> ggplot()+
+ geom_bar()+
+ scale_fill_manual()+
+ labs()
```

Nohing happens without assignment

```
read.csv("adjectives.csv")
```

open the file and chill

Read errors

If you try code and get an error → something went wrong → you didn't finish → more info please.

```
stat_count() can only have an x or y aesthetic.
```

```
Error: unexpected symbol in "ggplot(adjectives,
aes(x=Value)) + geom_bar(fill="indianred1",
color="#FF9999") + labs(title = adjectives
summarise"
```

You have until the end of today to fix these errors.

`esquisse` bug on Windows 10: `esquisse::esquisser(iris, viewer = "browser")`

Don’t leave your arguments hanging

```
adjectives_clean <- adjectives %>%
  na.omit() %>%
  filter(ADVERB!="123" & ADVERB!="dghdhffhg" & age>=17 &
         LIST %in% 1:6 & Value %in% 1:7) %>%
  select(Value, ADJECTIVE, ADVERB) %>%
  group_by(ADVERB) %>%
  summarise(mean = mean(Value),
         count = n())
arrange(ADVERB)
```

→ missing pipe

→ R gets angry, what is ADVERB?

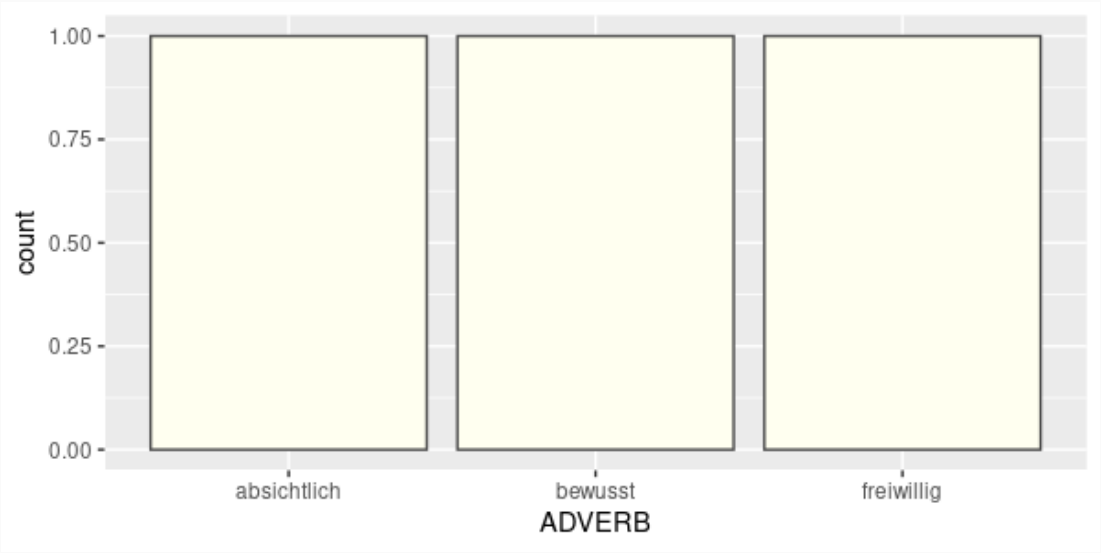
```
ggplot() +
geom_line()
```

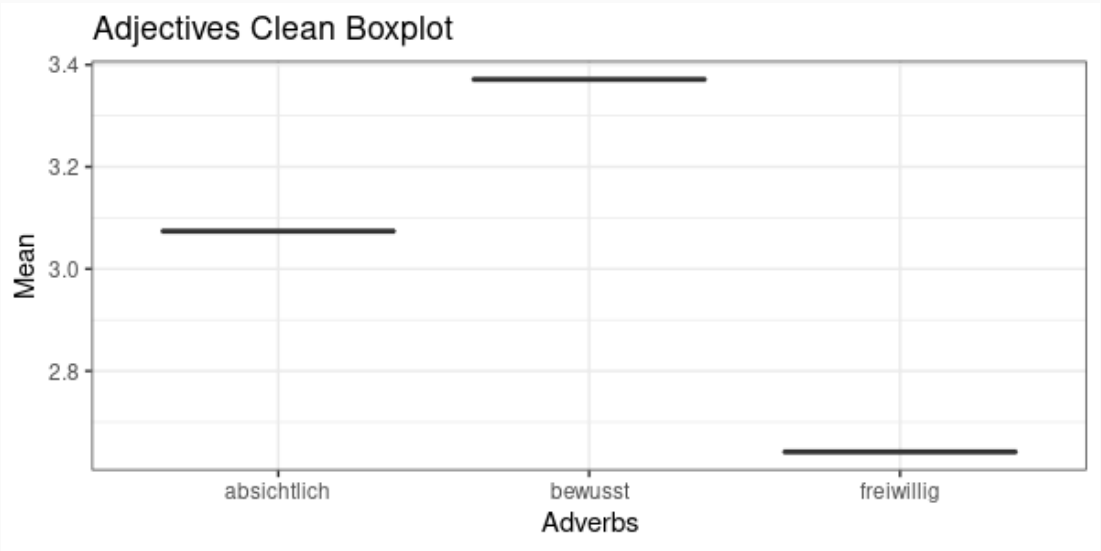
☺ 👍

```
ggplot()
+ geom_line()
```

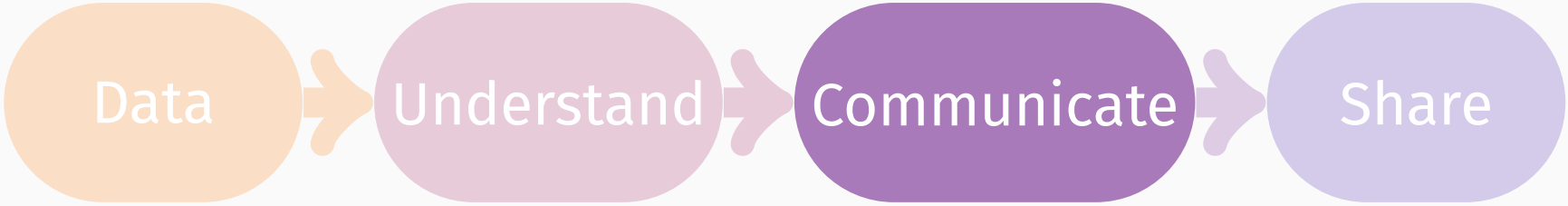
☹ 👎

The right tool for the job





Communicating



import, clean, transform, export, visualize, model

document, create clean and beautiful reports

connect, collaborate, backup

Table of contents

1. Documentation

2. Reporting with `RMarkdown` and `knitr`

3. Homework assignment

Documentation

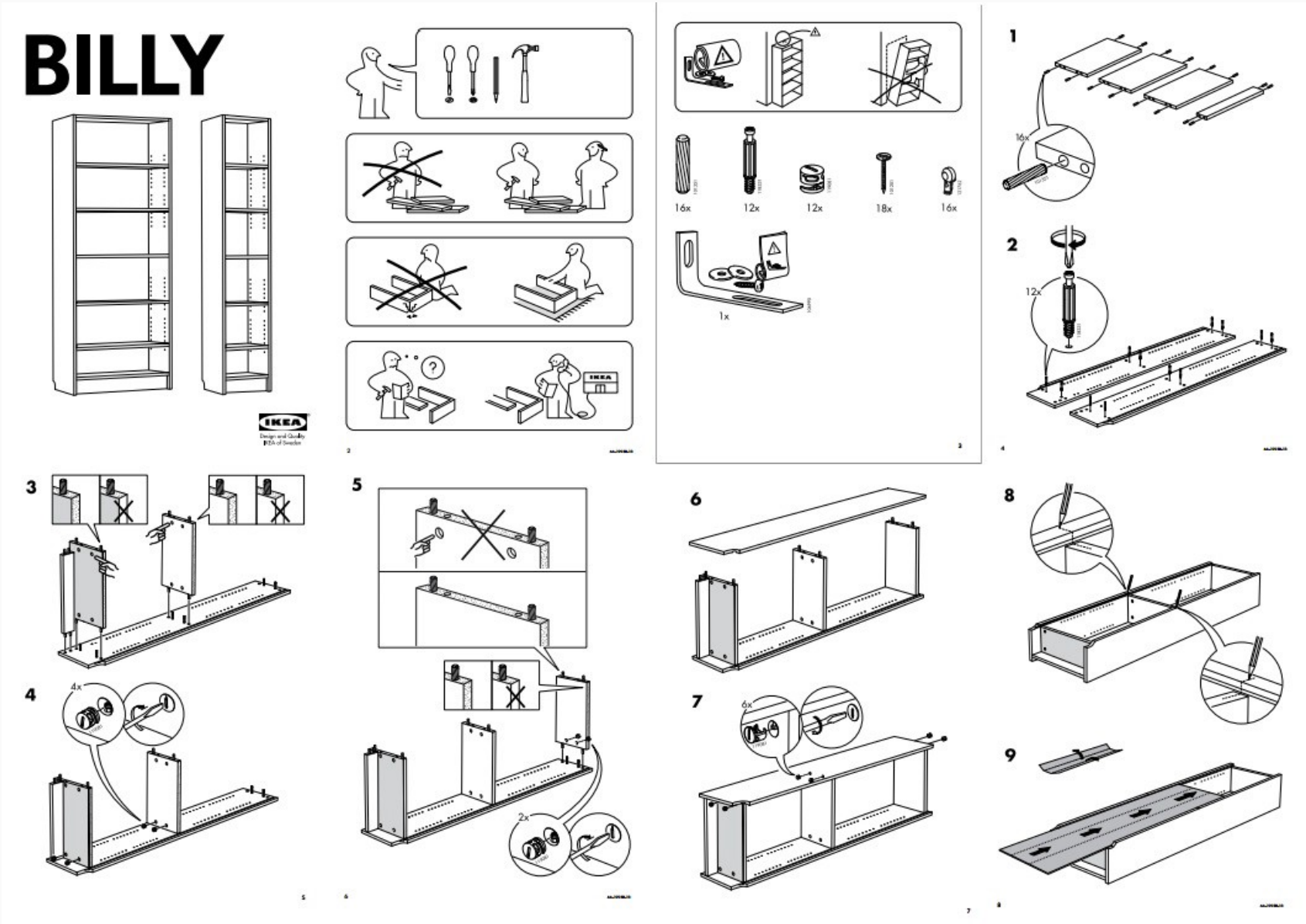
Keeping records

Why bother?

Someone WILL revisit your work → you or someone else

Up to date	keep track of all aspects
Reconsider	you think about what you're doing
Improve	quality (data, code, analysis, etc.)
Share	communication and collaboration
Remember	decisions, procedures, options, details, etc.
Reproducibility and transparency	easy access to information
Troubleshooting	what isn't doing its job

How to document

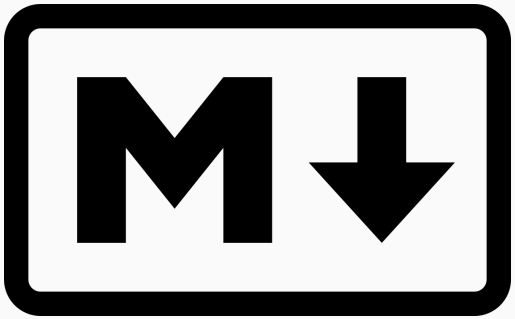


Make your documentation

Understandable	to your audience (at least yourself)
Short	tl;dr
Descriptive	what is **x** and what does it do
Focused	get to the point
Complete	what data, packages, etc.
Fresh	memory decays fast
Timestamped	when did this happen?

Reporting with RMarkdown and knitr

Markdown



Lightweight markup language.

“write using an easy-to-read and easy-to-write plain text format”

Key goal: readability (compare e.g. RTF or HTML)

Widely used (Wikipedia, blogs, forums, instant messaging, GitHub, readme files, software documentation etc.)

Pandoc

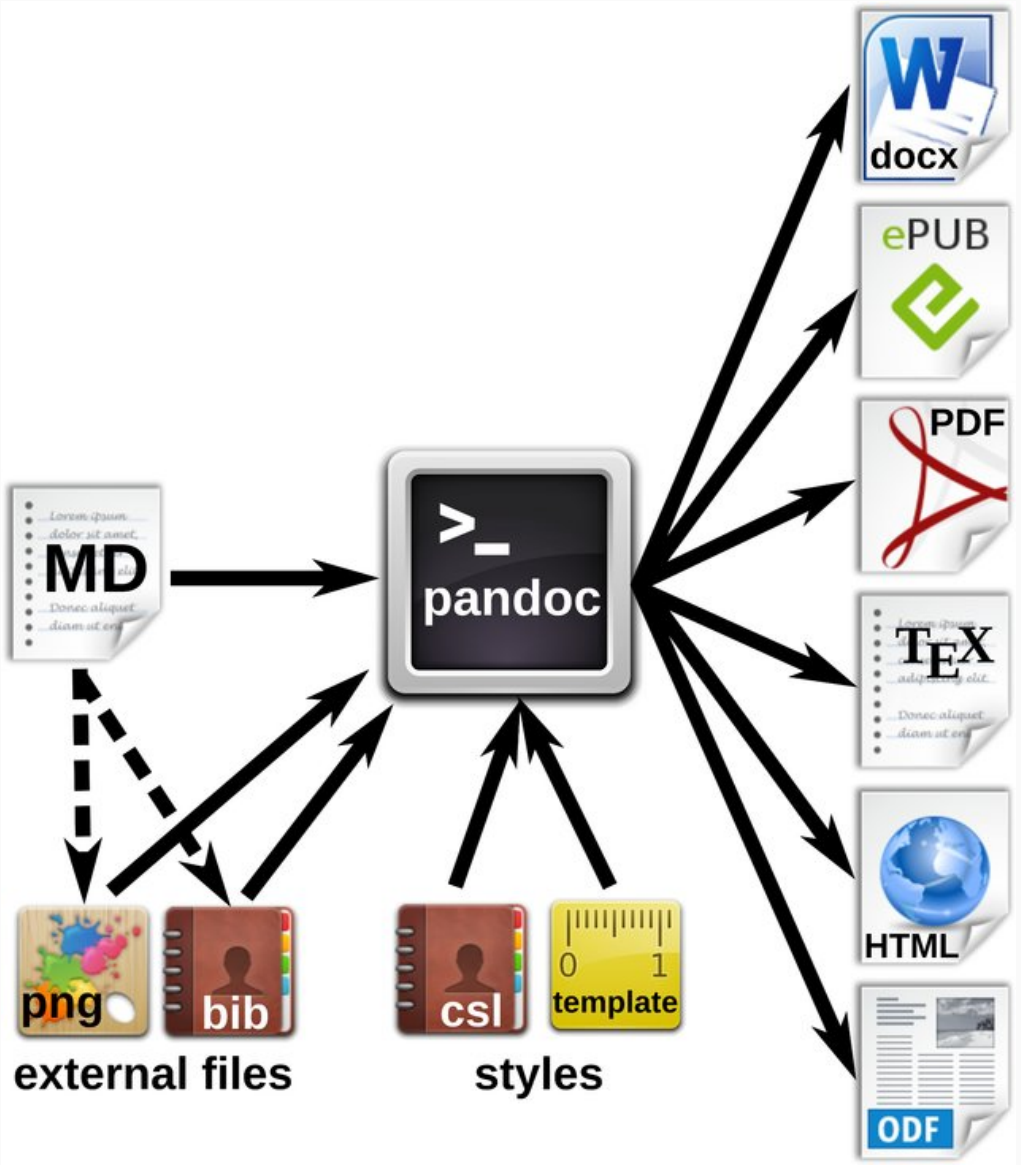
Enhanced Markdown

Free and open source document/markup format converter.

Converts a document from one format to another (`https://pandoc.org`).

Basic functionality (looks are less important).

Supports bibliographies.



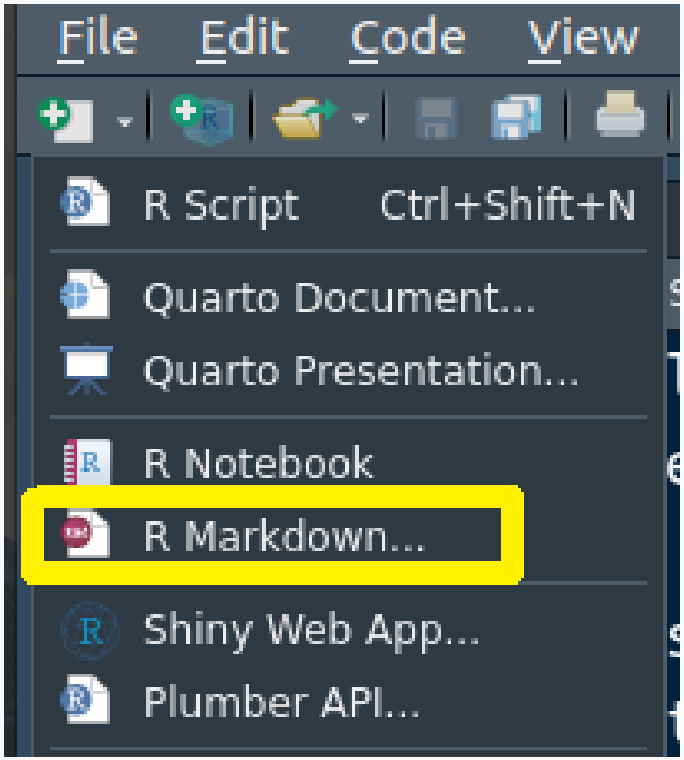
Krewinkel & Winkler (2017) `10.7717/peerj-cs.112`

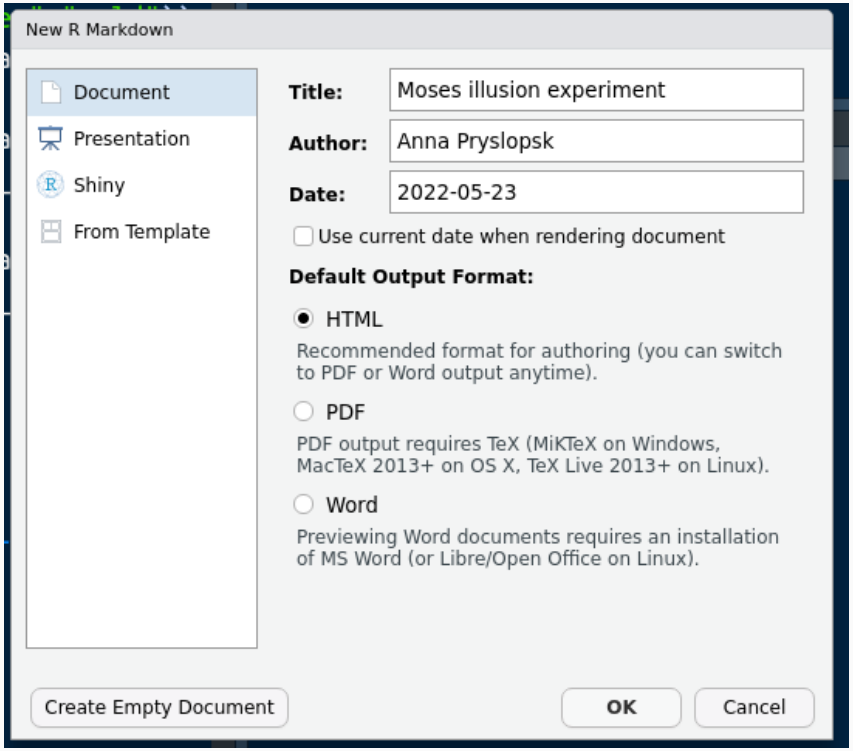
Report with `RMarkdown` and `knitr`

`RMarkdown` for syntax, `knitr` for creating reports.

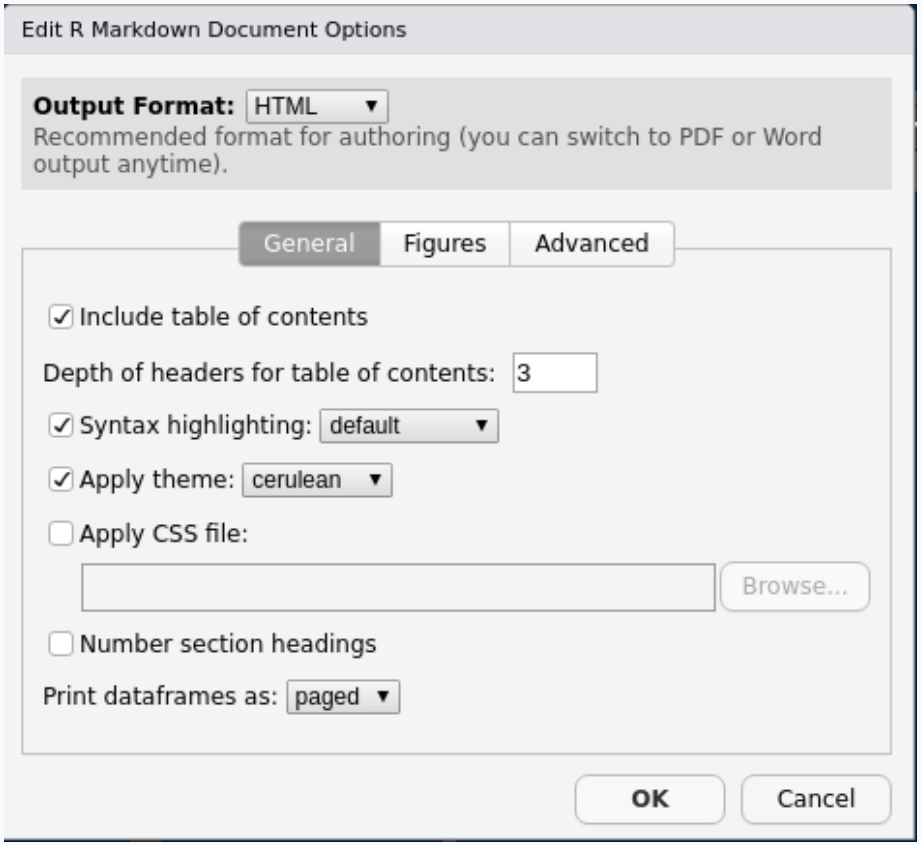
Why analyze data, then report if you can analyze data + report!

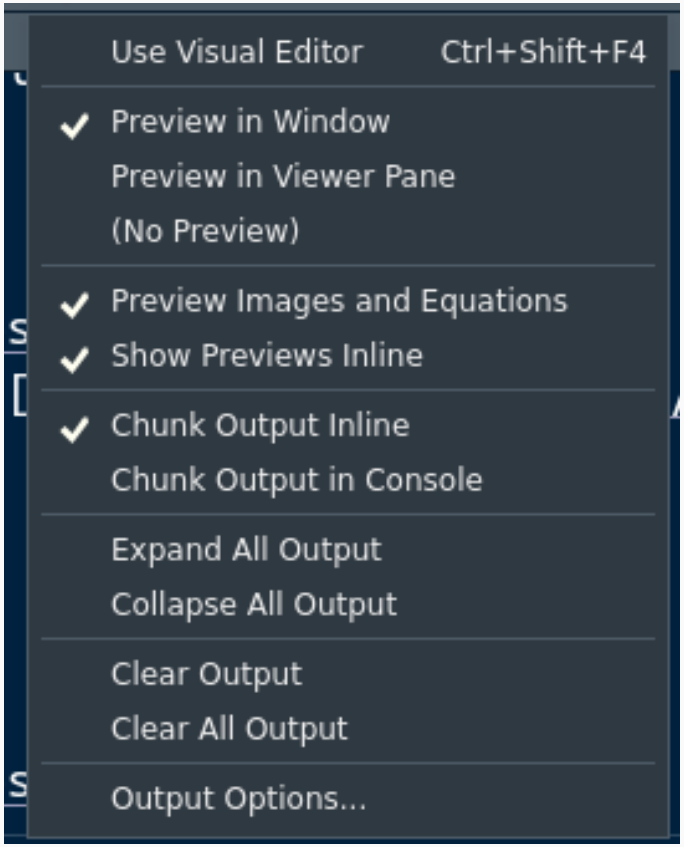
→ harder, better, faster, stronger





Customize


Edit R Markdown Document Options
Output Format: HTML
Recommended format for authoring (you can switch to PDF or Word output anytime).
General
Figures
Advanced
Include table of contents
Depth of headers for table of contents: 3
Syntax highlighting: default
Apply theme: cerulean
Apply CSS file:
Browse...
Number section headings
Print dataframes as: paged
OK
Cancel


Use Visual Editor Ctrl+Shift+F4
Preview in Window
Preview in Viewer Pane
(No Preview)
Preview Images and Equations
Show Previews Inline
Chunk Output Inline
Chunk Output in Console
Expand All Output
Collapse All Output
Clear Output
Clear All Output
Output Options...

RMarkdown shortcuts

Use shortcuts (e.g. select + * = *italics*)

Have a reference guide open

Try out the Visual Editor

- similar to commercial text editors
- see what you get but may get messy

Questions?

Homework assignment

Homework assignment due May 30

Submit as RMarkdown and HTML files.

- Report on the Moses illusion experiment.
- Include R code chunks and plain text.
- Make at least one table and create one plot of the data.

Use scientific article structure (be very brief, this isn't a term paper)

1. **Background** (what is the Moses illusion, include sources)
2. **Methods**
    - **Design** (of experiment = 1 factor within-design, i.e. Moses illusion vs. normal sentence) **and materials** (what different groups were in the experiment, example sentences → ILIAS)
    - **Participants** (who took part and how many participants)
    - **Procedure** (what did you have to do)
    - **Predictions** (what was the expected outcome for each condition/group in experiment)
3. **Results and discussion** (main findings and conclusion)