dropR: Analyze and Visualize Dropout

Annika Tave Overlander, Matthias Bannert & Ulf-Dietrich Reips

2024-07-09

dropR

What is dropout (DO)?

- Psychological online research is great! 🧠 (Reips 2000)

  - We can reach many people → *high statistical power*

  - We can ask questions in several formats, e.g. 5-point rating scales or visual analog scales (similar to sliders)

  - There is less social pressure, so participants may as well *be honest*

  - There is less social pressure, so participants may as well *leave*
    ☹

What to do about dropout?

There are several ways to deal with dropout (Reips 2002a, 2002b):

- Avoid/ Control: give incentives to stay (e.g. study credits for students or $$$)

- Use as dependent variable for analysis!

  - Do participants drop out more in one experimental condition vs. others?

  - Can give information on motivation of participants for different versions of the same survey (Kaufmann and Reips 2024)

Why is dropout analysis important?

- **In-lab** there is little dropout due to social pressure
  → Valuable information might be lost

- **Online** we can use dropout to improve survey methods

  - A condition shows higher rates of dropout? Maybe it is too difficult or boring

  - Females show higher rates of dropout? Maybe our survey is too specifically male

  - Older participants drop out more? Maybe text is too small or people don't have necessary knowledge

Introducing dropR

Dropout analysis should be a routine part of analyzing online data and reported as such.

- dropR is a free R package and Shiny App to facilitate dropout analysis for everyone in the scientific community

  - Statistically solid analyses *for everyone*

  - Clean and simple reporting conducted & downloaded from the App *for everyone*

  - More advanced/ customized outputs from the R package *for more advanced R users*
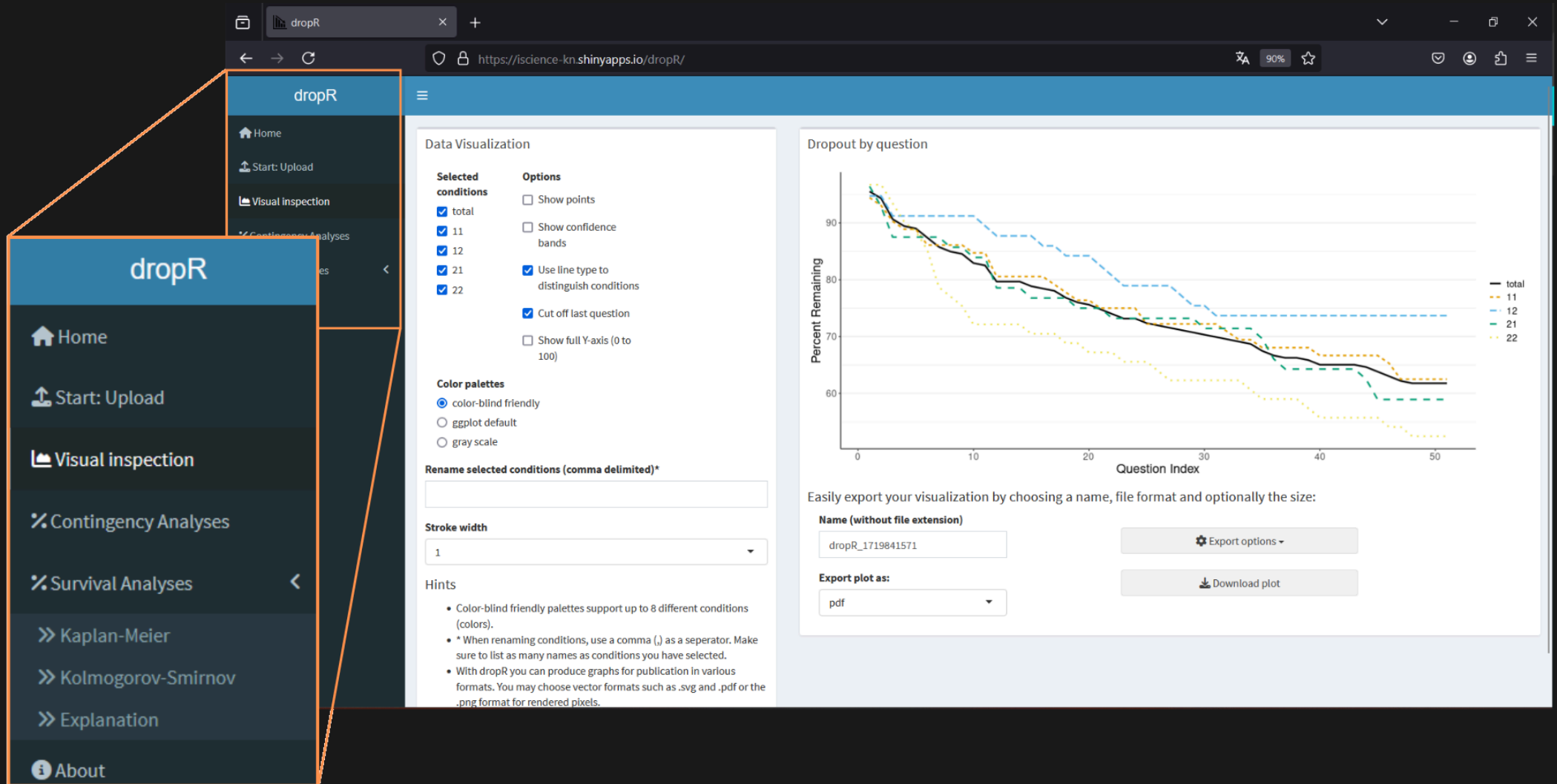
dropR - the Shiny App

dropr.eu

(https://iscience-kn.shinyapps.io/dropR/)

- Can also be started locally in R directly with
  `dropR::start_app()`

- Leads users through analysis process with tabs

  - Start: Upload

  - Visual Inspection

  - Contingency Analyses

  - Survival Analyses

Shiny App

dropR: Visual Inspection Tab

dropR - the R package

- Install the official CRAN package with `install.packages("dropR")`

- You can always install the developer version from GitHub using `devtools::install_github("iscience-kn/dropR")`

- All features from the app can be used as individual functions from the package

dropR

dropR Workflow

There are 2-3 preparatory steps:

1. Load the `dropR` package using `library(dropR)`

2. Add a dropout index to your data using the function `add_dropout_idx()` and specify your research questions

3. Compute the dropout statistics with `compute_stats()` either only for the overall data or, in most cases, by experimental condition

Load dropR → Add Dropout Index → Compute Dropout Stats → Further Analyses & Data Viz

dropR

dropR Workflow

Your initial data should look something like the `dropRdemo` demo dataset:

ID	condition	item_1	...	item_51	item_52
7a9f33	11	1	...	1	1
e11f94	22	1	...	NA	NA
e72a50	22	1	...	1	1

- Wide, tidy data

- You can use `tidyverse` workflow as a nice way to conduct a whole analysis + visualization in basically one step

dropR Workflow Example
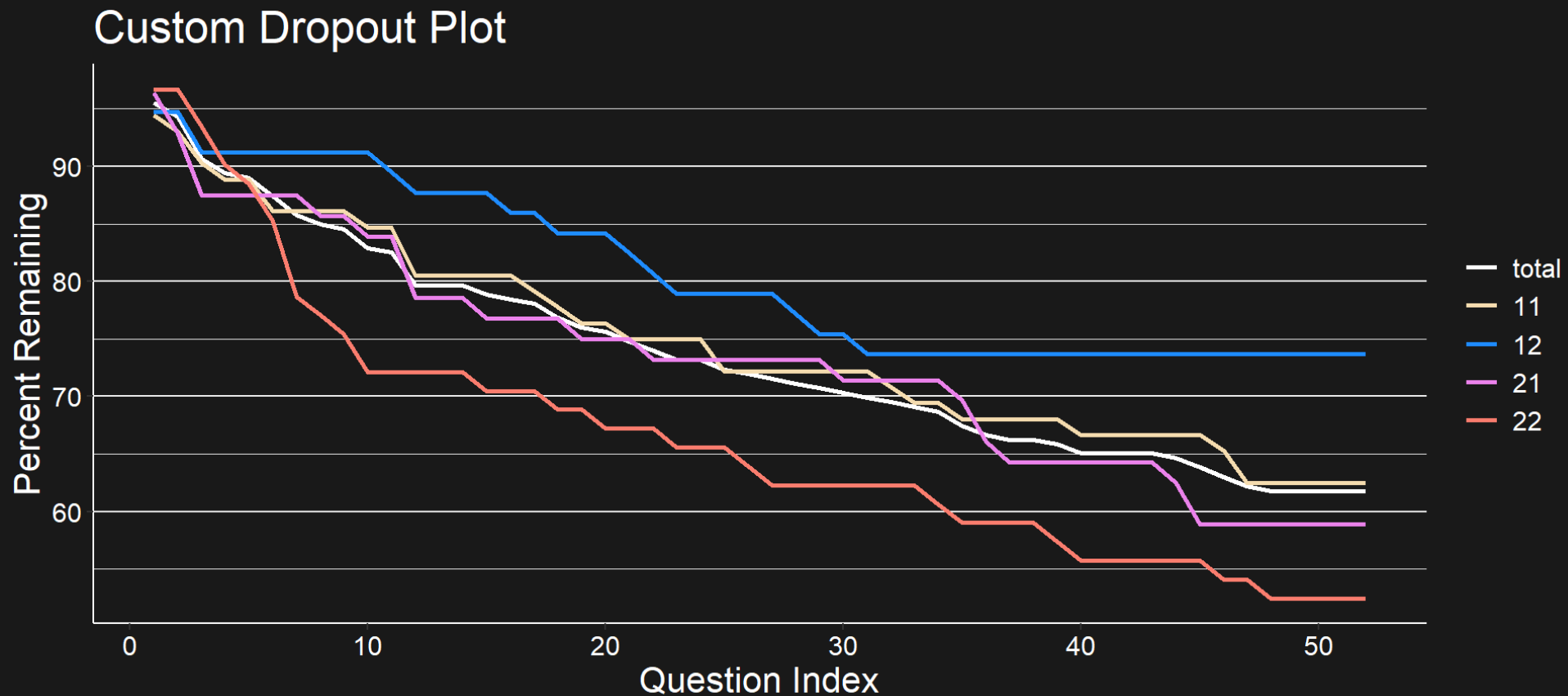
```
1  do_stats <- dropRdemo %>%
2    add_dropout_idx(q_pos = 3:54) %>%
3    compute_stats(by_cond = "experimental_condition", no_of_vars = 52
4  plot_do_curve(do_stats, full_scale = FALSE)
```

dropR Workflow Example
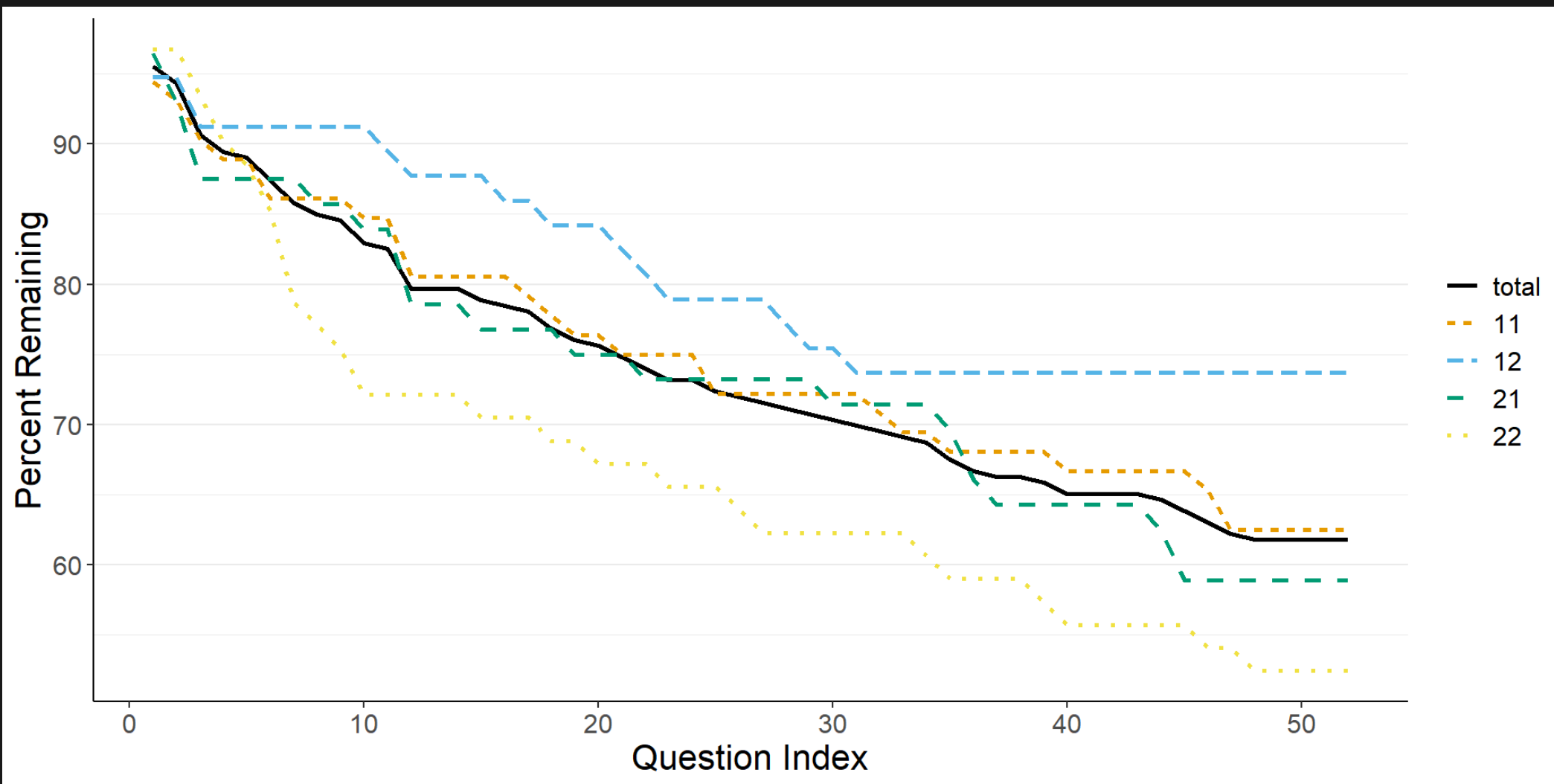
With a bit of customization…

```
do_stats  %>%
    plot_do_curve(full_scale = F, linetypes = F) +
    labs(title="Custom Dropout Plot") +
    scale_color_manual(values = c("white", "wheat", "dodgerblue", "violet
```


Custom Dropout Plot

dropR

Advanced Analyses

Easy to use in the workflow:

- Kaplan-Meier Estimation for survival differences: `do_kpm()` & `plot_do_kpm()`

- Chi^2 analysis identifies difference between conditions at a specific question: `do_chisq()`

- Kolmogorov-Smirnov test of the most different conditions at a specific question: `do_ks()` & `plot_do_ks()`

Most Different Conditions
Kolmogorov-Smirnov Test & Visualization
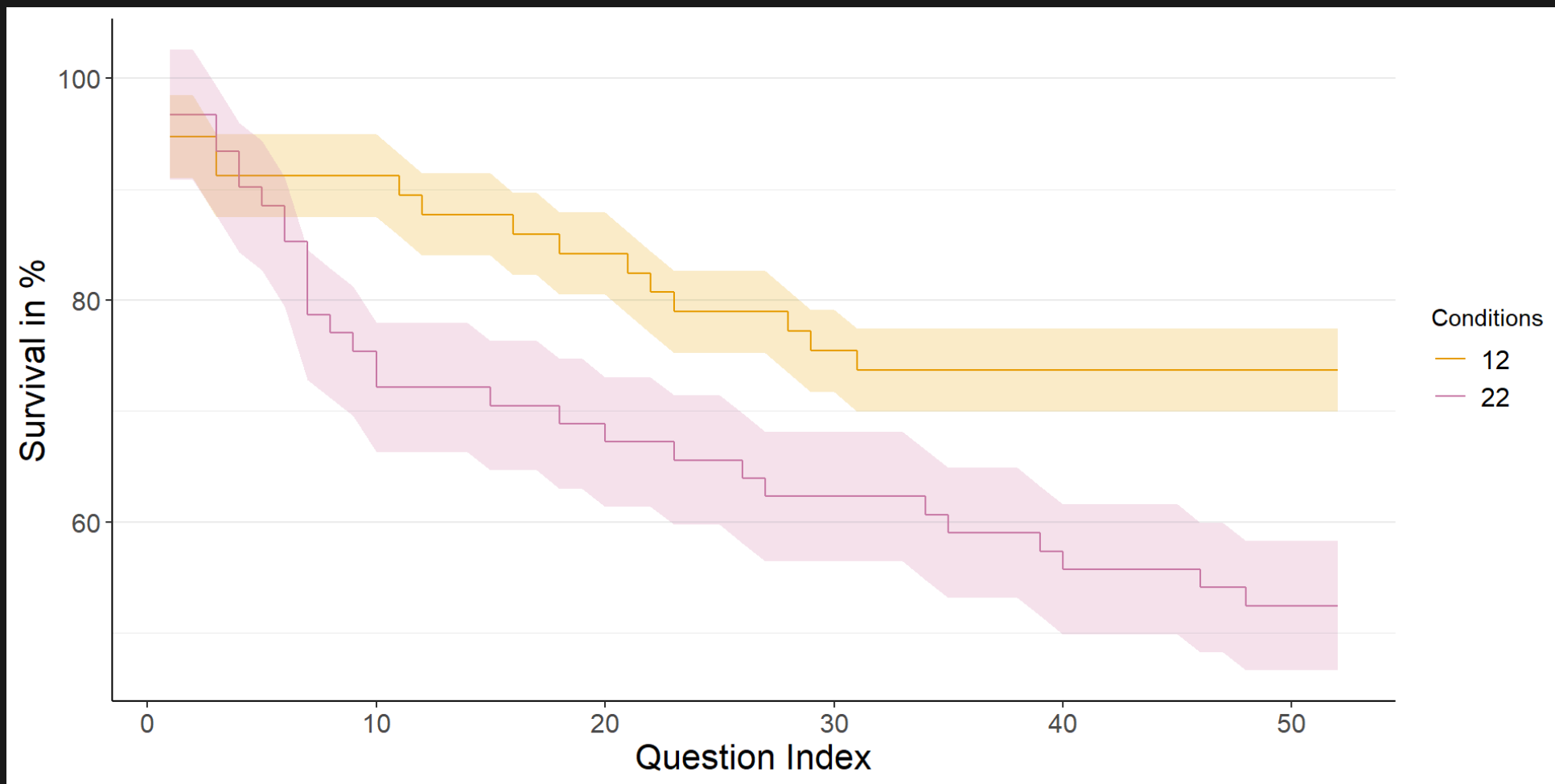
dropR

So, why dropR?

- dropout is a fairly well-known phenomenon in data collection (Reips 2009)

- dropout analysis, however, is not so common

- We are currently aware of one other standardized tool, the R package dropout

  - Also helps with basic analysis

  - *Lacks plotting tools and advanced analyses*

Conclusion

- Dropout can provide a lot of information that other measures cannot

  - dropout analysis should be a routine part of analyzing online data

- `dropR` facilitates such analyses for everyone

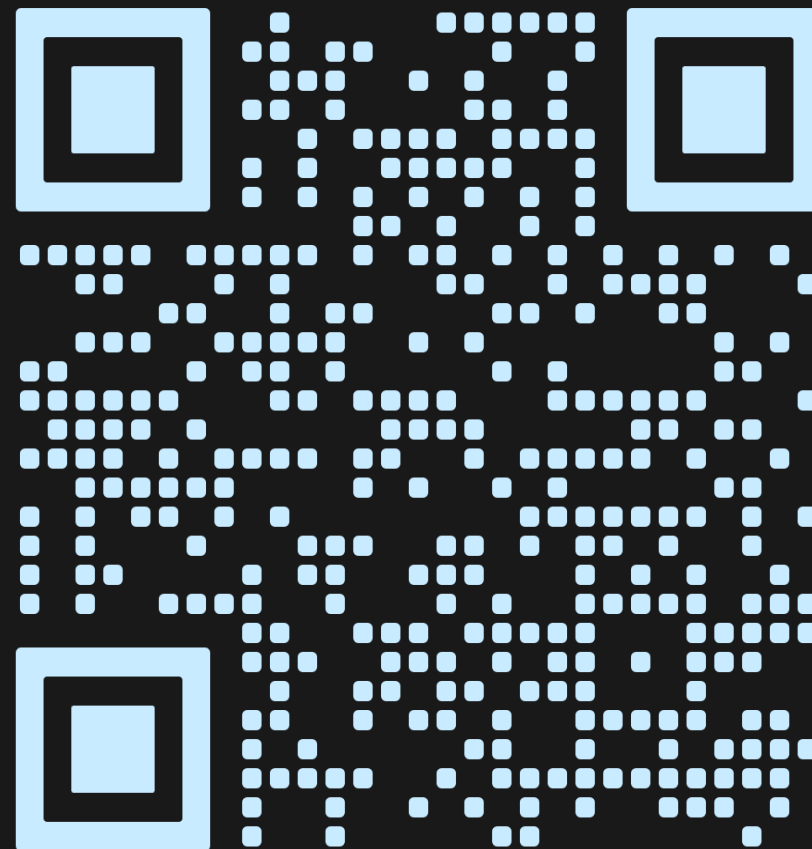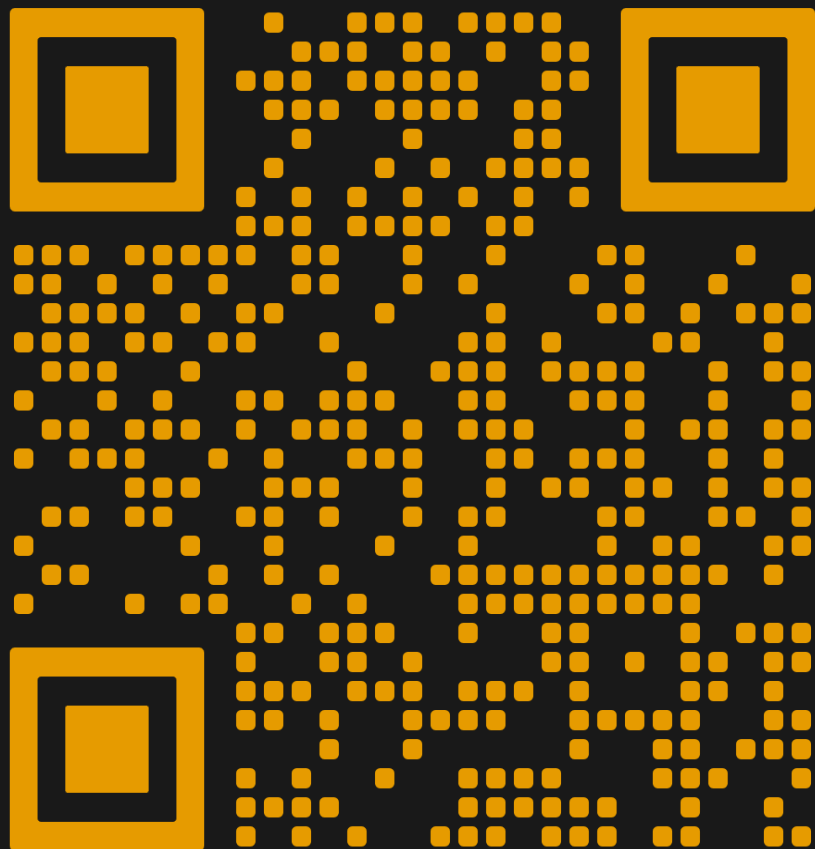 - Duality of App and Package offers balance between accessibility and advanced options

Outlook

- Make code used in shiny app available to copy (similar to reading in data with `readr`)

- Potentially migrate from shinydashboard to bslib

- Hopefully many valuable ideas and contributions from the community!

Sources

Kaufmann, Esther, and Ulf-Dietrich Reips. 2024. "Meta-Analysis in a Digitalized World: A Step-by-Step Primer." *Behavior Research Methods*, April. https://doi.org/10.3758/s13428-024-02374-8.

Reips, Ulf-Dietrich. 2000. "The Web Experiment Method: Advantages, Disadvantages, and Solutions." In *Psychological Experiments on the Internet*, edited by Michael H. Birnbaum, 89–117. San Diego: Academic Press. https://doi.org/10.1016/B978-012099980-4/50005-8.

———. 2002a. "Internet-Based Psychological Experimenting - Five Dos and Five Don'ts." *Social Science Computer Review* 20 (September): 241–49.

———. 2002b. "Standards for Internet-Based Experimenting." *Experimental Psychology (Formerly Zeitschrift für Experimentelle Psychologie)* 49 (4): 243–56. https://doi.org/10.1027//1618-3169.49.4.243.

———. 2009. "Internet Experiments: Methods, Guidelines, Metadata." In *Human Vision and Electronic Imaging XIV*, 7240:53–61. SPIE. https://doi.org/10.1117/12.823416.

Find us after the talk!

dropR on GitHub:
https://github.com/iscience-kn/dropR

iscience website: https://iscience.uni-konstanz.de/

dropR