Open Science Practices: Implementing a Reproducible Analysis Workflow
SoSe2024

Daniela Palleschi

2023-04-12

Table of contents

Welcome!

- the language of instruction is English
- you will need a laptop in class from May 6h
- and a recent installation of

    – R
    – RStudio
    – tinyTex

Course Moodle

Course name: Open Science Practices: Implementing a Reproducible Analysis Workflow Enrolment Key: r4repro

- course documents are available as HTML slides and PDF
- a course website will be up and running shortly

Learning objectives

- learn basic concepts of Open Science in the scope of language research

  - focus on reproducibile analysis

- create and maintain self-contained analyses projects
- develop good habits for project management
- learn how to transparently write up analyses for a paper or thesis
- learn about concepts like version control and containerization

What you will not learn in this book

- how to analyse data

- how to fit models

- how to plot data

- how to design an experiment

- we will not explicitly learn how to do these things, although we will be running simple code to achieve these goals

Course credits

- 4 LP

  - 1LP: participation and readings
  - 1LP: in-class exercises and preparation
  - 1LP: 2 in-class quizzes (0.5LP each)
  - 1LP: a pre-registration (1LP)

Session Info

Save your session info at the end of each document. Our results very often depend on the version of R/RStudio/a package we used. This is a great first step towards creating a reproducible workflow!

```
R version 4.3.0 (2023-04-21)
Platform: aarch64-apple-darwin20 (64-bit)
Running under: macOS Ventura 13.2.1

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRlapack.dylib;  

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

time zone: Europe/Berlin
tzcode source: internal

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base     

loaded via a namespace (and not attached):
 [1] compiler_4.3.0    fastmap_1.1.1     cli_3.6.2         tools_4.3.0
 [5] htmltools_0.5.7   rstudioapi_0.15.0 yaml_2.3.8        rmarkdown_2.25
 [9] knitr_1.45        jsonlite_1.8.8    xfun_0.42         digest_0.6.34
[13] rlang_1.1.3       evaluate_0.23
```

References