Institut zur Qualitätsentwicklung im Bildungswesen
Forschungsdatenzentrum
WISSENSCHAFTLICHE EINRICHTUNG DER LÄNDER AN DER HUMBOLDT-UNIVERSITÄT ZU BERLIN E.V.

Workshop: Reproducible Data Analysis Workflow with R

Dozierender: Aaron Peikert, M.Sc. Psych. (Max-Planck-Institut für Bildungsforschung, Berlin)

Termin: Teil 1: Mittwoch, 27.09.2023, 9:00 – 13:00 Uhr

Teil 2: Donnerstag, 28.09.2023, 9:00 – 13:00 Uhr

Abstract

Computational reproducibility represents a key challenge for making one's own work accessible and easily understandable for others—but achieving this goal is not easy. Often, researchers must invest considerable efforts to find out what the end goal of such a reproducible workflow can look like and how they can acquire the skills to get there. This workshop gives an overview of how to create reproducible data analyses in R (and beyond). Specifically, this workshop will showcase one way to automate the entire process from raw data to publishable manuscripts in a pre-set online environment. This automation combines dynamic document generation (via R Markdown), version control (via Git), workflow orchestration (via Make) and software management (via Docker). All these tools and, therefore, the entire automatic reproduction of results are available on any machine with Docker installed. The resulting workflow is hence ensured to reproduce over time, locally and in the cloud.

Inhalte

Continues delivery of research products using automated reproduction via R(Markdown)/Git(Hub)/Make/Docker.

Voraussetzungen

Stable internet connection, GitHub Account, OpenAI Account

Literatur.

Peikert, A., & Brandmaier, A. M. (2021). A Reproducible Data Analysis Workflow. *Quantitative and Computational Methods in Behavioral Sciences*, *1*, e3763. https://doi.org/10.5964/qcmb.3763

Peikert, A., van Lissa, C. J., & Brandmaier, A. M. (2021). Reproducible Research in R: A Tutorial on How to Do the Same Thing More Than Once. *Psych*, *3*(4), 836–867. https://doi.org/10.3390/psych3040053

Van Lissa, C. J., Brandmaier, A. M., Brinkman, L., Lamprecht, A.-L., Peikert, A., Struiksma, M. E., & Vreede, B. M. I. (2021). WORCS: A workflow for open reproducible code in science. *Data Science*, *(in press)*, 1–21. https://doi.org/10.3233/DS-210031

Software

We will use a Cloud Environment (Posit Cloud + GitHub Actions) but later you are welcome to use a local installation of R/RStudio/Git/Make/Docker if you want.