

Datenanalyse mit R

Praktische Übung – Jürgen Hermes – DH – SoSe 2022

Ziele dieses Kurses

- Erwerb der wichtigsten Grundlagen der Programmiersprache R
- Umgang mit der Entwicklungsumgebung RStudio
- Ermittlung interessanter Datenquellen zur Analyse
- Selbständige (Klein)Projektumsetzungen der Teilnehmer

Eingesetzte Mittel

- Vorstellung der wichtigsten Konzepte im Präsenzunterricht
- Verweis auf Online-Kurs beim HPI (zum Selbststudium)
- Bearbeitung kleinerer Hausaufgaben (Abgabe über ILIAS)
- Abgabe Projektumsetzung bis zum 31.07.2022

Die Programmiersprache R

- Freie, vollständige Programmiersprache, die meist für statistische Berechnungen und Visualisierungen eingesetzt wird.
- Publiziert 1993, aktuelle Version 4.1.3
- Benutzerfreundliche integrierte Entwicklungsumgebung RStudio
- Besonderheiten: Für statistische Berechnungen optimierte Datenstrukturen und Funktionen, besonderes Potential in der einfachen Grafikerzeugung.
- “To understand computations in R, two slogans are helpful: Everything that exists is an object. Everything that happens is a function call.” (John Chambers)
- Extrem große Nutzer:innen-Community, die viele vorgefertigte, nutzbare Pakete zur Verfügung stellt.

Hausaufgabe

- Installieren Sie sich R in einer Version ab 4.0 auf Ihrem Rechner
→ <https://cran.r-project.org/>
- Installieren Sie sich die Entwicklungsumgebung RStudio auf Ihrem Rechner
→ <https://www.rstudio.com/products/rstudio/download/>
- (optional) Treten Sie dem Online-Kurs “Programmieren mit R für Einsteiger” bei und bearbeiten Sie die Materialien von Woche 0 und Woche 1.
→ <https://open.hpi.de/courses/programmieren-r2022/>