

# Die R-Lösung oder:

wie man mit R umfassend statistische Datenanalysen betreibt

Isabelle Borucki, Yannick Diehl, Daniel Moosdorf, Stine Ziegler  
Ort: Philipps-Universität Marburg, Creative Space Bahnhofstr. 7,  
Club Behring (Raum +1/0020)

**Reading Week: 20.11. - 24.11.2023**

## Überblick

Soziale Interaktion findet zunehmend in digitalen Räumen statt. Folglich verändern sich nicht nur die sozialwissenschaftlichen Forschungsgegenstände, sondern auch die zur Verfügung stehenden Daten. Das eröffnet neue Möglichkeiten für spannende Analysen, stellt Forschende aber auch vor besondere Herausforderungen bei der Auswahl, Sammlung, Aufbereitung und Analyse digitaler Datenquellen.

Rein auf statistische Analysen fokussierte Software wie SPSS oder Stata geraten hier schnell an ihre Grenzen, weshalb in den Datenwissenschaften und Computational Social Sciences längst mit anderen Programmen gearbeitet wird. Die kostspielige Lizenzierung kommerzieller Software stellt eine weitere Hürde dar. Die Open-Source-Programmiersprache R und die integrierten Entwicklungsumgebung RStudio/Posit decken die Funktionalitäten klassischer Statistik-Software vollumfänglich ab und können durch ihre Flexibilität und Dynamik immer wieder an die aktuell zu lösenden Aufgaben angepasst werden. Zusammen mit einer sehr aktiven und hilfsbereiten Community sind in R gute Grundlagen für autodidaktisches Lernen gegeben.

Die Idee des R-Lab ist, Ihnen einen leichten Einstieg in R zu ermöglichen. Im geschützten Raum und in kleinen Gruppen nähern wir uns der Arbeit mit R in den Sozialwissenschaften an. Es wird sowohl eine angeleitete Einführung als auch freie Phasen zum Ausprobieren, Üben und Knobeln geben. Studierende, die bereits Erfahrungen mit der Arbeit in R haben, können gern zu diesen Phasen dazustoßen. Die Mitarbeiter:innen des Lehrgebiets werden in den freien Phasen vor Ort sein und bei Problemen unterstützen. Begleitend werden kürzere Intensivkurse angeboten, welche sich spezifischen Themen widmen (Liste s.u.). Ziel des R-Labs ist, dass Sie R grundständig kennengelernt haben und eine Grundlage für selbstständiges vertiefendes Lernen in der R-Umgebung besitzen.

## Voraussetzungen für die Teilnahme:

Sie benötigen keine Programmierkenntnisse, um beim R-Lab teilzunehmen. Da wir uns auf die technische Umsetzung der Datenanalyse konzentrieren, sind methodische Grundlagen des quantitativen Arbeitens sowie erste praktische Erfahrungen von Vorteil. Eine eigene Fragestellung mit entsprechender Datengrundlage ist für die Übungsphasen hilfreich, aber nicht notwendig. Alternativ zu eigenen Projekten steht entsprechend des E-Books (s.u.) der Allbus 2018 bzw. 2021 zur Verfügung.

**Wichtig:** Bitte bringen Sie Ihren eigenen Laptop mit, da keine Geräte zur Verfügung gestellt werden können.

Text-Grundlage ist das E-Book von Diehl, Yannick und Moosdorf, Daniel (2022) R Workshop – Sozialwissenschaften; abrufbar unter: <https://rloesung.github.io/RWorkshop/>

## Anmeldung:

Bitte melden Sie sich bis zum 17.11.2023 per E-Mail an [moosdorf@staff.uni-marburg.de](mailto:moosdorf@staff.uni-marburg.de) (Betreff: R-Lab) an. Bitte geben Sie an, ob Sie zur Einführung oder nur zu den freien Phasen (auch stundenweise möglich) kommen möchten. Gerne können Sie sich auch zu einem angebotenen Intensivkurs anmelden:

## Intensivkurse:

Die Intensivkurse sind dafür gedacht, zu den jeweiligen Themen “deep dives” für die eigene tägliche Anwendung zu ermöglichen. Zudem werden den Teilnehmenden konkrete workarounds und shortcuts an die Hand gegeben, um selbstständig weiter arbeiten zu können. Details zu den Kursen können Sie dem Wochenprogramm entnehmen.

- Montag, 20.11.2023, 14-16 Uhr: RStudio/Quarto
- Mittwoch, 22.11.2023, 14-16 Uhr: Erstellen von Grafiken mit ggplot
- Donnerstag, 23.11.2023, 14-16 Uhr: Analyse von Textdaten mit GermaParl
- Freitag, 24.11.2023, 14-16 Uhr: Mit iGraph Netzwerke visualisieren

## Wochenprogramm:

### Montag, 20.11.2023: Installation von R, RStudio und den notwendigen Paketen

- 10-14 Uhr: Einführung in die Arbeit mit R und R-Studio – Teil 1

Die erste Schwierigkeit besteht bei der Arbeit mit R bei der Installation. Vor allem, wenn Sie noch nie mit R gearbeitet haben, können das Einrichten und die Arbeit mit packages zu Irritationen führen. Ziel ist daher, dass alle Interessierten R, RStudio sowie die für statistische Analysen und wissenschaftliches Arbeiten notwendigen Pakete (sog.

packages) installiert haben, sodass in der Gruppe mit den gleichen Funktionen im Verlauf der Woche gearbeitet werden kann.

Gemeinsam mit den Dozierenden werden anschließend die ersten zwei Kapitel des E-Books durchgearbeitet. Es soll in die Arbeit in R eingeführt, Grundlagen der Programmiersprache sowie erste Schritte der Datenverarbeitung (Beginn Kapitel 3) vermittelt werden.

- 14-16 Uhr: Freie Übungs- und Praxisphase
- 14-16 Uhr: **RStudio/Quarto-Intensiv**: Wie schreibe ich mit Quarto Hausarbeiten und Präsentationen? Was bringt mir Rmarkdown? Und wie integriere ich einen Arbeitswork-flow darin? Diese Fragen und Fragen, die am ersten Tag aufgetreten sind, werden hier vertieft und die Arbeit mit codebasierten Formaten gemeinsam erschlossen.

### **Dienstag, 21.11.2023: Angeleitet durch die ersten Schritte**

- 10-14 Uhr: Einführung in die Arbeit mit R und R-Studio – Teil 2  
Am zweiten Tag wollen wir die Einführung in die Datenverarbeitung (Kapitel 3 des e-Books) abschließen und erste Schritte in der uni- und bivariate Datenanalyse gemeinsam gehen (Kapitel 4 des e-Books). Auch hierbei sollen Ihre persönlichen datenanalytischen Fragestellungen im Vordergrund stehen und die einzelnen Schritte individuell an ihren Daten bearbeitet werden.
- 14-16 Uhr: Freie Übungs- und Praxisphase

### **Mittwoch, 22.11.2023: Eigene Analysen und erste Grafiken**

- 10-14 Uhr: Einführung in die Arbeit mit R und R-Studio – Teil 3  
Zu Beginn des dritten Tages sollen gelernte Inhalte wiederholt und erste Grafiken erstellt werden (Abschluss Kapitel 4). Mit den einfachen Funktionen des sjPlot-Pakets können wir Ergebnisse unserer uni- und bivariaten Analysen schnell und ansprechend grafisch aufbereiten. Anschließend steigen wir in der verbleibenden Zeit in die multivariate Datenanalyse (Kapitel 5) ein.  
Die weitere Vertiefung in R soll individuell erfolgen. Dabei soll ihr persönliches Interesse im Vordergrund stehen. Wenn Sie aktuell kein eigenes datenanalytisches Projekt haben, können Sie sich an der Struktur und den Analysen des e-Books orientieren. Die anwesenden Dozierenden werden Sie bei jedem Arbeitsschritt bei Bedarf unterstützen.
- 14-16 Uhr: Freie Übungs- und Praxisphase

- 14-16 Uhr: Erstellen von Grafiken mit **ggplot**: Hier wollen wir uns ganz der Visualisierungspower mit R widmen und die Formate und Formen der verschiedenen Arten zur Darstellung von Daten anschauen. Ziel ist es, mit der spezifischen Syntax und Notation von ggplot vertraut zu werden und anhand von Beispieldaten erste Gehversuche mit diesem Paket zu unternehmen.

### **Donnerstag, 23.11.2023 und Freitag, 24.11.2023: Individuelle Vertiefung**

- 10-16 Uhr: Freie Übungs- und Praxisphase

Die verbleibenden zwei Tage des R-Labs sind inhaltlich offen. Sie können genutzt werden, um Erlerntes zu wiederholen oder zu vertiefen. Gern können Sie hier auch eigene Projekte anfangen oder weiterführen. Die Dozierenden nehmen an diesen Tagen eine eher passive Funktion ein, stehen Ihnen aber bei Fragen und Problemen zur Seite. So soll ein Selbstlernort geschaffen werden, in dem wir unsere eigenen Projekte voranbringen und uns gegenseitig unterstützen. Zudem bieten wir noch zwei weitere Intensiv-Sessions an.

- Do., 14-16 Uhr: Analyse von **Textdaten mit GermaParl**: Cooking with GermaParl lädt uns ein mit dabei zu sein, live und in Farbe zu sehen, was Textdaten aus einem Textdatenkorpus leisten können, was sie so besonders macht und welche Analysemethoden damit angewandt werden können.
- Fr., 14-16 Uhr: Mit **iGraph Netzwerke visualisieren**: igraph ist das ggplot für Netzwerkdaten. Also nicht nur Daten aus Social Media, auch andere Daten sind mittels sozialer Netzwerkanalyse analysierbar und somit erschließbar. Wie wir Netzwerke dann sichtbar und somit greifbar machen können, ist Ziel dieser Intensivstunde.