

# Data Scientist of tomorrow? – Einführung in die sozialwissenschaftliche Datenanalyse mit R

Die fortschreitenden Digitalisierung sozialer Interaktionsräume und die vielfältigen Bezugsmöglichkeiten verschiedenster standardisierter/unstandardisierter Daten durch exponentiell wachsende Datenquellen erlaubt es auch den Sozialwissenschaften mit größeren und kleineren Datenmengen umzugehen und diese für spannende Analysen aufzubereiten (Stw. Data Science). Dies wird aktuell noch vermehrt anhand lizensierter Software (wie SPSS oder Stata) vermittelt. Damit sind Einschränkungen verbunden, welche zum einen die methodischen Möglichkeiten dieser Programme betreffen, als auch bezüglich des Zugangs, welcher für Studierende hauptsächlich über die universitären PC Pools oder kostenpflichtigen Lizenzen erfolgt.

Mit der Open Source Programmiersprache R und der integrierten Entwicklungsumgebung RStudio können diese Einschränkungen überwunden werden. Wenn der Einstieg einmal geschafft ist, kann R durch seine 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.

Das Ziel dieses Kurses ist es, den Einstieg in R auch Studierenden und Anderen ohne größere Vorkenntnisse im Programmieren zu ermöglichen und eine Anleitung zu geben, wie sich innerhalb der R Community selbst geholfen werden kann. Um dies zu leisten haben wir einen zweitägigen Workshop geplant, welche folgenden Themen mit den gängigsten Paketen und Ökosystemen (u.a. tidyverse, strengjacked, Easystats) behandelt:

1. Einführung in R
  - Programm kennenlernen
  - Objekte, Pakete
2. Datenaufbereitung
  - Daten importieren und modifizieren
3. Erste Datenanalyse
  - Univariate Datenanalyse
  - Bivariate Datenanalyse
4. Multivariate Datenanalyse
5. Grafiken erstellen

R findet neben der universitären Verwendung zunehmend Einzug in andere Bereiche, in denen mit größeren Datenmengen gearbeitet wird. Mit R arbeiten zu können, kann damit eine Schlüsselqualifikation für verschiedene Tätigkeitsfelder darstellen.

**Kurszeitraum:** 2-tägiger Workshop – 07.06. von 14:00 bis 18:00Uhr/08.06. von 9:00 bis 15:00Uhr.

**Kurzzugang:** Wir bitten alle Interessierten sich bei [daniel.moosdorf@staff.uni-marburg.de](mailto:daniel.moosdorf@staff.uni-marburg.de) zu melden. Um abschätzen zu können, ob wir den Kurs in Präsenz, online oder hybrid anbieten, bitten wir Sie, dass Sie einfach eine E-Mail schreiben, dass Sie Interesse an dem Kurs haben und ob Sie online oder in Präsenz (eigener Laptop) teilnehmen möchten. Über diese formlose Anmeldung erhalten Sie weitere Informationen zum Ablauf des Kurses.

**Zielgruppe:** Der Kurs richtet sich besonders an Studierende vers. Sozialwissenschaften mit Interesse an quantitativer Forschung. Darüber hinaus steht der Kurs aber auch allen offen, die Grundkenntnisse in R erlernen wollen.

**Kursziele:** Selbstständiges Arbeiten in R auf grundständigem Niveau; wesentliche analytische Verfahren auf sozialwissenschaftlichen BA-Niveau in R umsetzen können

**Voraussetzungen:** grundlegende Kenntnisse statistischer Datenauswertung. Installation von R und RStudio (bitte in dieser Reihenfolge!):

- Download von R: <https://cran.r-project.org/>
- Download von RStudio (Desktop): <https://www.rstudio.com/products/rstudio/download/>

**Dozenten:** Diehl, Yannick; Moosdorf, Daniel; Thullen, Kilian – Bei Fragen zum Kurs schreiben Sie uns einfach eine Mail an: [daniel.moosdorf@staff.uni-marburg.de](mailto:daniel.moosdorf@staff.uni-marburg.de)