



Einführung in die Datenanalyse mit

Data Science - die rechnergestützte Analyse von Daten mit grafischen und statistischen Methoden - gehört derzeit zu den gefragtesten Feldern der Informatik. Zwei Sprachen sollten sich momentan unbedingt im Werkzeugkasten eines jeden "Data Scientisten" befinden - Python und R. R ist nicht nur eine statistik-orientierte Programmiersprache sondern bietet rund um seinen Kern ein ganzes Ökosystem an Paketen für die verschiedensten Analyse-Methoden und Visualisierungen. Gerade in dieser großen Vielfalt und Menge an Verfahren, sowie der schnellen, interaktiven Realisierung von Analysen liegt auch der Vorteil, der R - neben Python - für den Daten-Analysten unverzichtbar macht.

Ziel:

Die Vorlesung soll dem Studenten das nötige Rüstzeug vermitteln, selbstständig grafische und numerische Analysen in R durchführen zu können. Das dafür notwendige Wissen über grundlegenden Elemente der Kernsprache (einfache Datenstrukturen, Kontrollstrukturen, S3-Klassen, ...) sowie wichtige und hilfreiche Erweiterungspakete werden anhand von Beispielen erarbeitet. Ein Ausblick soll auch auf fortgeschrittenere Themen, wie S4 Klassen und das Erstellen eigener Packages gegeben werden.

Themen:

- Die Arbeitsumgebung, R, Editor/IDE, ...
- Die grundlegenden Datenstrukturen von R - Vector, Matrix, Data Frames, ...
- R als Sprache
- Daten-Import (Daten-Quellen)
- Umformen von Daten (Data-Munging/Data-Wrangling)
- Visualisierung von Daten - "*The Grammar of Graphics*" (ggplot2)
- S3-Klassen, Ausblick auf S4-Klassen
- Moderne Ansätze für die Daten-Analyse - das "*Tidyverse*"
- Erstellen von R-Notebooks
- Erstellen von R-Packages

Literatur:

- [John M. Chambers: "Software for Data Analysis - Programming with R", Springer](#)
- [Hadley Wickham, Garrett Grolemund: "R for Data Science - Visualize, Model, Transform, Tidy, and Import Data", O'Reilly](#)
- [Uwe Ligges: "Programmieren mit R", Springer](#)

Links:

- [R Projekt Seite](#)
- [R Projekt Seite - Contributed Documentation](#)
- [R-bloggers](#)