

Übungsblatt 5

Datenanalyse und -visualisierung mit R
Hochschule Bonn-Rhein-Sieg
Wintersemester 2022-23

Pekka Sagner M.Sc.

Mutating Joins - Grundlagen

Aufgabe 1: »Ein Datensatz kommt selten allein.«

- a) Laden Sie die Dateien `studis_*.csv` herunter. Ihnen liegen folgende Informationen zu den Dateien vor.

Dateiname	Beschreibung
<code>studis_insg.csv</code>	Studierende je 1.000 Einwohner, 1995-2017
<code>studis_insg_2017.csv</code>	Studierende je 1.000 Einwohner, 2017
<code>studis_fh.csv</code>	Studierende an Fachhochschulen je 1.000 Einwohner, 1998-2017
<code>studis_fh_2017.csv</code>	Studierende an Fachhochschulen je 1.000 Einwohner, 2017

- b) Laden Sie die Datensätze `studis_insg_2017.csv` und `studis_fh_2017.csv` in R ein. Verschaffen Sie sich einen Überblick über die Daten. In welchem Landkreis ist der Anteil der Studierenden an einer Fachhochschule an allen Studierenden am größten - in welchem am kleinsten?
- c) Laden Sie die Datensätze `studis_insg.csv` und `studis_fh.csv` in R ein. Verschaffen Sie sich einen Überblick über die Daten. Wie hat sich der Anteil der FH-Studierenden an allen Studierenden im Rhein-Sieg-Kreis entwickelt? Stellen Sie diese Entwicklung auch grafisch dar. Was fällt Ihnen auf? Ist Ihr Ergebnis sinnvoll?
- d) Visualisieren Sie die Entwicklung aller Studierenden je 1.000 Einwohner sowie der FH-Studierenden je 1.000 Einwohner im Zeitverlauf für die sieben größten deutschen Städte (Berlin, Düsseldorf, Frankfurt a.M., Hamburg, Köln, München, Stuttgart). Nutzen Sie weiterhin die Daten in `studis_insg.csv` und `studis_fh.csv`.