

# Data Science für die Humangeographie I

## Ein pragmatischer Einstieg mit R

true

Wintersemester 2020/21

## Contents

<b>Terminüberblick</b>	<b>1</b>
<b>Online-Ressourcen</b>	<b>2</b>
R Tutorials und eBooks . . . . .	2
Inspiration für Visualisierungen . . . . .	2
Spezialthemen . . . . .	3
<b>1 Vorbesprechung</b>	<b>3</b>
1.1 Überblick . . . . .	3
1.2 Seminarformat . . . . .	4
1.3 Leistungsnachweise . . . . .	5
1.4 Lehrphilosophie . . . . .	6
<b>2 Erste Schritte</b>	<b>7</b>
2.1 Vorbereitung . . . . .	7

## Terminüberblick

*Alle Sitzungen finden von 13 bis 16h c.t. statt*

Datum	Sitzung	Inhalt
2. November 2020	1	Vorbesprechung
9. November 2020	2	Erste Schritte
16. November 2020	3	[Text] Anderson 2008
23. November 2020	4	Datenstrukturen
30. November 2020	5	Visualisierungen
7. Dezember 2020	6	[Text] Shelton et al. 2014
14. Dezember 2020	7	Geodaten
11. Januar 2021	9	Choroplethen

Datum	Sitzung	Inhalt
18. Januar 2021	10	[Text] Chandra 2014
25. Januar 2021	11	HTML-Tabellen
1. Februar 2021	12	Web Scraping
8. Februar 2021	13	[Text] Straube 2021
15. Februar 2021	13	Präsentationen
31. März 2021		Abgabe Exposé

## Online-Ressourcen

### R Tutorials und eBooks

- **R for Data Science**  
<https://r4ds.had.co.nz/>  
 Ausführliches Handbuch, Fokus auf Data Science
- **RStudio Cloud Primers**  
<https://rstudio.cloud/learn/primers/1>
- **Swirl**  
<https://swirlstats.com/students.html>  
 Interaktives Tutorial als R-Paket, mit verschiedenen Lektionen
- **Quick-R**  
<https://www.statmethods.net/r-tutorial/index.html>  
 Überblickartiges Tutorial, kurz und bündig
- **RStudio Cheat Sheets**  
<https://www.rstudio.com/resources/cheatsheets/>  
 Einseitige Cheat Sheets zu verschiedenen Themen
- **Google's R Style Guide**  
<https://google.github.io/styleguide/Rguide.xml>  
 Regeln für leserlichen R Code

### Inspiration für Visualisierungen

- **R Graph Gallery**  
<https://www.r-graph-gallery.com/>  
 Viele Beispiele für verschiedenste Visualisierungen
- **DDJ Katalog**  
<http://katalog.datenjournalismus.net/#/>  
 Portfolio Datenjournalismus, leider etwas veraltet
- **Subreddits**  
<https://www.reddit.com/r/dataisbeautiful>

<https://www.reddit.com/r/DataArt/>  
<https://www.reddit.com/r/MapPorn/>

## Spezialthemen

- **Tutorial Reguläre Ausdrücke**  
<https://danielfett.de/en/tutorials/tutorial-regulare-ausdrucke/>  
Deutschsprachige Einführung zu regulären Ausdrücken

# 1 Vorbesprechung

## 1.1 Überblick

### 1.1.1 Seminar im Curriculum

- Dieses Seminar ist Bestandteil des Moduls BA3.
- Das Projektseminar besteht aus zwei Teilen über zwei Semester:
  - Konzeption quantitativer Forschung (Wintersemester)
  - Analyse quantitativer Daten (Sommersemester)
- Im Winter gibt es 12 inhaltliche Termine (davon 4x Textarbeit).
- Im Sommer wird das Seminar mit den selben Teilnehmer\*innen fortgeführt.

### 1.1.2 Lernziele für das Wintersemester

Sie können...

- einfache Skripte in R eigenständig erstellen.
- Datensätze in vielfältigen Formaten visualisieren.
- Online-Ressourcen gezielt einsetzen.
- Möglichkeiten der Datenbeschaffung identifizieren.
- epistemologische Verschiebungen durch Data Science wiedergeben.

### 1.1.3 Technische Anforderungen

- Es sind keine Vorkenntnisse in R erforderlich.
- Sie brauchen einen Laptop, mit dem Sie gut arbeiten können.
- Wir benutzen die RStudio Cloud als Plattform.
- Sie brauchen einen ruhigen Arbeitsplatz.

### 1.1.4 Unterstützung im Corona-Semester

- Die Uni bietet einen “Semesterlaptop” an.
- Bei Bedarf kann ich gerne versuchen, Arbeitsplätze im Seminarraum (PEG) anzubieten.
- Bitte kontaktieren Sie mich per E-Mail, falls Sie einen Arbeitsplatz regelmäßig in Anspruch nehmen wollen würden.

## **1.2 Seminarformat**

- Das Seminar findet jede Woche Montags, 13–16h c.t. statt.
- Der Zoom-Link, den Sie per E-Mail erhalten haben, bleibt gleich.
- Wir machen um ca. 14:25h eine zehnminütige Pause.
- Für Textbesprechungen wird die Gruppe zweigeteilt.
- Dieses Seminar findet in verschiedenen Modi statt:

### **1.2.1 Input und Plenum**

- Ich rede oder moderiere (mit Folien oder ohne)
- Sie “melden” sich für Redebeiträge oder Fragen (Zoom-Funktion)
- Sie hören mir und Ihren Kommiliton\*innen aufmerksam zu

### **1.2.2 Think-pair-share**

- Sie bearbeiten eine Fragestellung in zufälligen Zweier-Konstellationen (Breakout-Session)
- Nach einer vorgegebenen Zeitspanne kehren Sie ins Plenum zurück
- Ich fordere Sie ggf. auf, Ergebnisse und offene Fragen mit der Gruppe zu teilen

### **1.2.3 Follow the recipe**

- Ich teile ein unvollständiges Beispielprojekt.
- Wir gehen die Teilschritte nach und nach durch.
- Ich “habe den Plan”, stelle aber immer wieder Fragen ans Plenum.
- Sie vollziehen die Schritte an Ihrer eigenen Kopie des Projekts nach.

### **1.2.4 Hands-on session**

- Sie bearbeiten praktische Aufgabenstellungen alleine.
- Dabei sind sie in zufälligen Dreier-Konstellationen (Breakout-Session).
- Bei Fragen oder Problemen wenden Sie sich zunächst an Ihre Kleingruppe.
- Falls Sie nicht weiterkommen, fordern Sie Hilfe an (Zoom-Funktion).
- Ich reagiere auf Hilfesuche oder schaue in zufälligen Gruppen vorbei.

### **1.2.5 Share your work**

- Ich wähle eine Teilnehmer\*in zufällig aus.
- Die Person teilt ihren Bildschirm und berichtet von ihrer Bearbeitung eines Problems.
- Alle anderen unterstützen solidarisch durch aktives Nachvollziehen, Nachfragen und Hinweise.

## 1.3 Leistungsnachweise

### 1.3.1 Exposé

- Zum Ende des Wintersemesters geben Sie ein Exposé für ein Untersuchungsvorhaben für das Sommersemester ab.
- Sie können sich mit bis zu vier Personen zusammenschließen.
- Die Projektgruppe besteht dann verbindlich für das Sommersemester.
- Damit steigen aber auch die Anforderungen an Umfang, Detail und technischen Anspruch.
- Umfang für das Exposé: max. 15k Zeichen inkl. Leerzeichen, exkl. Literaturverzeichnis
- Als Abgabetermin haben wir den 31. März vereinbart.

#### 1.3.1.1 Inhalte

- Einführung ins Thema
- Forschungsstand / Literaturüberblick
- Herleitung einer klar abgegrenzten (vorläufigen) Forschungsfrage
- Konkrete Datenquellen
- Ideen für Verfahren und Visualisierungen

**1.3.1.2 Bewertungskriterien** Alle Kriterien werden mit einer (runden) Schulnote bewertet. Der gewichtete Schnitt ergibt die Gesamtnote.

Kriterium	Gewichtung	Erläuterung
Zitierweise und Formatierung	10%	Der Text erfüllt formale Anforderungen an Wissenschaftlichkeit.
Ausdruck und Rechtschreibung	10%	Der Text ist sprachlich gelungen.
Roter Faden	10%	Der Text ist übersichtlich strukturiert und die Einzelteile greifen gut ineinander.
Literatur	10%	Die zitierten Quellen sind für eine Einführung ins Thema geeignet und werden gut zusammengefasst.
Theorie	10%	Relevante wissenschaftliche Perspektiven werden anhand von geeigneter Fachliteratur aufgezeigt.
Fragestellung	10%	Die Forschungsfrage ist für das Vorhaben geeignet und wird überzeugend hergeleitet.

Kriterium	Gewichtung	Erläuterung
Datenquellen	20%	Die Datenquellen sind geeignet und detailliert beschrieben.
Design	20%	Das Untersuchungsvorhaben ist nachvollziehbar beschrieben, und der technische Anspruch ist dem Projektseminar angemessen.

### 1.3.2 Anwesenheit

- Es besteht Anwesenheitspflicht.
- Für Ihre ersten zwei Fehltermine pro Semester brauche ich keine Entschuldigung (aber Sie sollten das ggf. mit ihrer Projektgruppe absprechen).
- Sie sind dann selbstständig für die Nacharbeit der behandelten Themen zuständig.
- Im Falle eines zusätzlichen Fehltermins brauche ich ein Attest und einen Nachweis über Nacharbeit.
- Zur Anwesenheit gehört...
  - uneingeschränkte Aufmerksamkeit über die komplette Veranstaltungsdauer,
  - aktive Mitarbeit an Beispielen,
  - Bearbeitung von Übungsaufgaben,
  - aktive Beteiligung an Diskussionen.
- Eine eingeschaltete Kamera macht das allen Beteiligten leichter!

## 1.4 Lehrphilosophie

- Die folgenden vier “Säulen” habe ich mal im Rahmen einer Fortbildung als meine Lehrphilosophie definiert.
- Sie spiegeln meinen eigenen Anspruch an meine Lehre wider und sind als Vorschlag für ein gutes Miteinander zu verstehen.
- Begreifen Sie das gerne auch als Ermunterung, Aspekte hiervon einzufordern, wenn sie in der Veranstaltung zu kurz kommen.

### 1.4.1 Transparenz

- Erforderliche Leistungen und Bewertungskriterien sind vorab bekannt.
- Termine und Regelungen werden in der Vorbereitungssitzung verbindlich vereinbart.
- Aktuelle Lehrmaterialien stehen online durchgängig zur Verfügung.

### **1.4.2 Praktische Übungen**

- Eigenständige Anwendung steht im Vordergrund.
- Verfahren und Techniken werden mit Beispielen und Übungen erarbeitet.
- Die perfekte Aufgabe ist immer ein bisschen “zu schwer”.
- Toleranz für Frustration ist eine wichtige Fähigkeit und lässt sich trainieren.

### **1.4.3 Geschützte Räume**

- Alle können sich im Plenum respektiert und sicher fühlen. Verletzendes Verhalten wird benannt.
- Es gibt einen vertrauensvollen Rahmen für ehrlichen Austausch.
- Frustrationen und Momente des Scheiterns werden ernst genommen und konstruktiv bearbeitet.

### **1.4.4 Kritische Reflexion**

- Auch Teilnehmende, die kein weiterführendes Interesse an der Anwendung quantitativer Verfahren haben, sind im Seminar gut aufgehoben.
- Verfahren werden kontextualisiert, ihre Limitationen werden aufgezeigt.
- Kritische Forschung zu quantitativen Praktiken wird besprochen.

## **2 Erste Schritte**

### **2.1 Vorbereitung**

- Machen Sie sich einen kostenlosen Account auf <https://rstudio.cloud>
- Treten Sie dem Seminar-Workspace bei. (Sie erhalten eine Einladung per E-Mail.)
- Optional/alternativ: installieren Sie R und RStudio auf Ihrem Computer.