

**Wintersemester 2020: 15340 - Techniken der Empirischen Parlamentsanalyse in R**

Donnerstag, 10-12 Uhr

**Zusammenfassung**

Wie können wir die Komplexität moderner politischer Prozesse empirisch erfassen? Das Ziel dieses Kurses ist es, grundlegende Techniken der quantitativen Analyse in der Statistiksoftware R anhand von spezifischen Anwendungsbeispielen zu vermitteln. Am Beispiel Föderalismus und Parlamenten betrachten wir so nicht nur Standardtechniken (Visualisierung, deskriptive und inferentielle Statistik) sondern auch fortgeschrittene Methoden (Text-as-Data, Diskursive Netzwerkanalysen).

*Eine gleichzeitige Teilnahme am Projektkurs „Dynamiken im föderalen System“ wird empfohlen, da dieser Kurs theoretische und empirische Überschneidungen hat.*

Materialien im Kurs sind zum Teil in Englischer Sprache. Die Kurssprache ist Deutsch.

**Organisatorische Hinweise**

**Literatur und Kursmaterial**

Die Pflichtlektüre und weitere Informationen (Videos + R Code + Datensätze) werden auf Blackboard bereitgestellt.

**Sprechstunden**

Dienstag und donnerstags 13:00-14:00 Uhr. Bitte vereinbaren Sie vorab einen Termin unter **Removed**

Leider können in diesem Semester Sprechstunden nur digital angeboten werden. Sie können direkt über **Removed** der Sprechstunde beitreten.

Bitte beachten Sie auch die gesonderten Sprechstunden die exklusiv für diesen Kurs angeboten werden (siehe Seminarplan).

**Kommunikation**

Die Kommunikation im Kurs findet via Blackboard und Email statt. Bitte kontaktieren Sie mich über die E-Mail Adresse **Removed**. Inhaltliche Fragen sollten ausschließlich in den Sprechstunden besprochen werden.

**Online Lehre und Coronavirus**

Die Pandemie stellt uns vor besondere Herausforderungen. Der Ablauf (vgl. Seminarplan) ist entsprechend getaktet, und verbindet synchrone und asynchrone Lehrelemente, um Ihnen eine größere Flexibilität zu gewährleisten. **Bitte beachten Sie auch, dass es nicht in jeder Woche ein synchrones Online-Seminar geben wird.**

Spezifisch hat der Seminarplan die folgenden Elemente:

- **Video:** Vorab aufgezeichnete Videos, in denen technische oder theoretische Inhalte vermittelt werden. Diese Videos sind die Basis der Online-Diskussion **und sollten vor dem jeweiligen Seminartermin bearbeitet werden.**
- **Datacamp:** Digitale Lerneinheiten, die die technischen Fähigkeiten vermitteln direkt testen. Auch diese Inhalte sind Grundlage für die Online-Diskussionen, und sollten deshalb **vor dem jeweiligen Seminartermin abgeschlossen werden.**  
  
Einheiten die als „Extra-Datacamp“ ausgewiesen werden sind als mögliche Erweiterungen gedacht, werden aber nicht zwangsläufig erwartet.
- **Diskussionen** Synchron Online Diskussionen via Webex. Begin um 10:15. Idealerweise werden wir nicht die vollen 90 Minuten ausschöpfen, da die asynchronen Elemente vergleichsweise viel Zeit einnehmen werden. Den Zugang zu unserem „Seminarraum“ finden Sie unter diesem Link: **Removed**
- **Sprechstunde** In Wochen ohne synchrone Online-Diskussionen stehe ich Ihnen für kursspezifische Kleingruppensprechstunden (mit maximal 4 Teilnehmer\*Innen) zur Verfügung. Bitte nutzen Sie für die Anmeldung in dieser Sprechstunde diesen gesonderten Link **Removed**
- **Problem-Sets** Um kursspezifische Anwendungsbeispiele der Methoden zu ermöglichen, werden im Laufe des Semesters drei Problem-Sets angeboten. Geben Sie diese bitte **bis zum Dienstagabend (23:59 Uhr) vor der Deadline ab.** Es geht hier explizit nicht darum, dass Sie schon alle Antworten beantworten können. Vielmehr ist es hilfreich, Verständnisprobleme so schnellstmöglich zu identifizieren und beheben zu können. Weiterhin ist die Kollaboration zwischen Studierenden hier ausdrücklich erwünscht. Sie können in Kleingruppen (Max. 4 Personen) zusammenarbeiten und die Problemsets gemeinsam abgeben.

Wir wissen noch nicht mit Bestimmtheit, ob wir die unten (im Plan) vorgesehene Organisation immer optimal und technisch einwandfrei realisieren können. Manches muss also improvisiert werden, und ggf. müssen wir die unten ausgewählten Formate im Laufe des Semesters an das Mögliche anpassen. Haben Sie bitte Verständnis und helfen Sie mit!

#### *Technische Voraussetzung*

Leider ist aufgrund der Corona-Epidemie der Zugang zum PC-Pool nicht möglich. Für die Teilnahme am Kurs benötigen Sie deshalb einen PC/Mac auf dem das Statistikprogramm R und R-Studio installiert und genutzt werden können. Tablets und ähnliche Endgeräte sind leider nicht ausreichend.

#### *Datacamp Zugang*

Dieser Kurs wird mit Hilfe des Online Dienstes Datacamp angeboten. Diese Unterlagen werden auf Englisch angeboten. Den Zugang zu Datacamp lasse ich Ihnen nach Beginn des Seminars zukommen.

#### **Best Practice - Videoseminare**

Die Online-Seminare finden in WebEx statt. Bitte Installieren Sie die Software auf Ihrem PC oder Mac. Nähere Informationen zur Installation und Anwendung finden sie [hier](#).

Grundsätzlich gelten im Internet die gleichen Regeln wie im normalen Raum: Rücksicht und Respekt helfen, den Austausch produktiv zu gestalten. Trotzdem gibt es einige Dinge, die in Telekonferenzen besonders wichtig sind.

- a. Stellen Sie ihr eigenes Mikrofon stumm, wenn Sie nicht sprechen. Auch wenn moderne Mikrophone sehr gut darin sind, die Hintergrundgeräusche zu filtern, können besonders in großen Veranstaltungen schnell störende Geräusche entstehen.
- b. Nehmen Sie Rücksicht auf die Sprecherin oder den Sprecher und unterbrechen Sie nur im Notfall. Selbst bei relativ schnellen Verbindungen gibt es oft Verzögerungen und Probleme, wenn mehr als eine Person spricht. Es wird empfohlen, die „Handheben“ – Funktion in Webex zu nutzen.
- c. Wenn die Bandbreite für eine stabile Verbindung nicht ausreicht, versuchen Sie zuerst Ihre Videokamera auszuschalten.
- d. Wählen Sie sich 5 bis 10 Minuten vor Beginn des Seminars in den virtuellen Seminarraum ein. Den Link für den virtuellen Seminarraum finden Sie auf Blackboard
- e. Soweit möglich, versuchen Sie einem ruhigen Ort ohne viele Hintergrundgeräusche und mit stabiler Internetverbindung zu finden.
- f. Für die non-verbale Kommunikation ist es sehr hilfreich, sich gegenseitig sehen zu können. Ich würde mich deshalb sehr freuen, wenn Sie Ihre Kamera angeschaltet lassen können. Allerdings ist mir auch bewusst, dass dies nicht für alle Menschen gleichermaßen umsetzbar ist, und ich habe Verständnis dafür, dass Sie aus persönlichen oder technischen Gründen die Kamerafunktion deaktivieren müssen. Falls Sie Ihre Privatsphäre schützen wollen, bietet Webex auch die Möglichkeit, den Hintergrund auszublenden oder nur verschwommen darzustellen.
- g. Leben und Lehre in der Pandemie I: Die besondere Pandemiesituation ist für uns alle eine Herausforderung. Ich werde versuchen den „normalen“ Seminarzustand so weit wie möglich aufrechtzuerhalten. Ich habe aber auch Verständnis, dass Studierende persönliche oder familiäre Verpflichtungen haben. Falls Sie Probleme haben, sagen Sie mir bitte Bescheid, so dass ich versuchen kann, eine gemeinsame Lösung mit Ihnen auszuarbeiten.

## Leistungsanforderungen im Seminar

### *Aktive Teilnahme:*

Abschluss der Aufgaben in Datacamp + Problem Sets 1 - 4

### *Leistungsschein:*

#### **Deadlines:**

28.1.2021: Abgabe der Exposees + Anmeldung für die Hausarbeit bitte digital über Blackboard

30.3.2020: Abgabetermin – bitte digital **Removed** und über Blackboard.

1. Leistungen für die aktive Teilnahme + Vertiefung der Problem Sets / Forschungsdesign (3000 Worte, exkl. Bibliographie)

Oder

2. Leistung für die aktive Teilnahme + Hausarbeit (6000 Worte, exkl. Bibliographie)

## Hintergrundliteratur

Wickham, Hadley. 2016. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media. Kostenfrei online [hier](#).

Healy, Kieran. 2018. *Data Visualization - A Practical Introduction*. Princeton University Press.

## Seminarplan

Sitzung 1: Einführung und Vorstellung (synchron)	5.11.2020
--	-----------

---

Online-Seminar: Vorstellung und Q&A – Ablauf und Logistik

Sitzung 2: Einführung in das Statistikprogram R Sitzung I	12.11.2020
---	------------

---

**Deadline:** ***Registrierung bei Datacamp***

Datacamp: Introduction to R (Beta)

Videos: R Installation + Projektorganisation in R

Onlinesprechstunde: Bitte um vorherige Anmeldung

Extra-Datacamp: Intermediate R

### *Literatur:*

Fortin-Rittberger, J., & Eder, C. (2013). Towards a Gender-Equal Bundestag? The Impact of Electoral Rules on Women's Representation. *West European Politics*, 36(5), 969–985.  
<https://doi.org/10.1080/01402382.2013.796702>

**Deadline:** ***R muss auf dem eigenen Rechner installiert sein***

Datacamp: Introduction to the tidyverse (Data wrangling + Grouping and summarizing)

Videos: Deskriptive Statistik im Tidyverse

Onlinesprechstunde: Bitte um vorherige Anmeldung

*Literatur:*

Bäck, H., & Debus, M. (2020). Personalized versus partisan representation in the speeches of migrant members of parliament in the German Bundestag. *Ethnic and Racial Studies*, 43(9), 1673–1691. <https://doi.org/10.1080/01419870.2019.1654112>

**Deadline:** ***Problem Set I – Deskriptive Daten***

Online-Seminar: Diskussion: Repräsentation in den Parlamenten  
Diskussion: Problem Set I

*Literatur:*

Wenzelburger, G., & Fehrenz, S. (2018). Die Union und die „Ehe für Alle“. Bestimmungsfaktoren des Abstimmungsverhaltens in der CDU/CSU-Fraktion im Bundestag. *Zeitschrift Für Parlamentsfragen*, 49(3), 512–530

Datacamp: Datavisualization with ggplot2 (Part 1 + Part 2)

Video: Daten visualisieren am Beispiel Legislatives Verhalten

Onlinesprechstunde: Bitte um vorherige Anmeldung

Extra-Datacamp: Interactive Data Visualization with plotly in R

*Literatur:*

Stecker, Christian. 2016. “The Effects of Federalism Reform on the Legislative Process in Germany.” *Regional and Federal Studies* 26(5): 603–24.

Lorenz, Astrid. 2010. Verfassungsänderungen in föderalen und unitarischen Demokratien im Vergleich. Befunde einer empirischen Analyse für den Zeitraum von 1945 bis 2004. In: Föderalismusreform in Deutschland. Bilanz und Perspektiven im internationalen Vergleich. Hrsg. Julia von Blumenthal und Stephan Bröckler, 13-36. Wiesbaden: VS Verlag für Sozialwissenschaften.

**Deadline:** ***Problem Set II – Visualisierungen***

Online-Seminar: Diskussion der Visualisierungen + Diskussion Problem Set II

Datacamp: Modeling with Data in the Tidyverse  
 Video: Regressionsanalysen am Beispiel  
 Onlinesprechstunde: Bitte um vorherige Anmeldung

**- Winterpause -**

Sitzung 8: Anwendung 3 – Regressionsanalysen (synchron)

7.1.2021

**Deadline:** ***Problem Set III – Regression***

Online-Seminar: Diskussion der Regressionsanalysen + Diskussion Problem Set III

Sitzung 9: Text-as Data I: Basics

14.2021

Datacamp: Introduction to text analysis in R (Chapter 1-3)  
 Video: Gernaparl und der Deutsche Bundestag

Onlinesprechstunde: Bitte um vorherige Anmeldung

Sitzung 10: Text-as Data II: Gernaparl und Parlamentsprotokolle (synchron)

21.1.2021

**Deadline:** ***Problem Set IV – Gernaparl***

Online-Seminar: Diskussion der Bundestagsdaten + Diskussion Problem Set 4

Sitzung 11: Text-as-Data III: Policy Positionen

28.1.2021

Onlinesprechstunde: Bitte um vorherige Anmeldung

*Literatur:*

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2), 311–331.  
<https://doi.org/10.1017/S0003055403000698>

Bernauer, J., & Bräuninger, T. (2009). Intra-Party Preference Heterogeneity and Faction Membership in the 15th German Bundestag: A Computational Text Analysis of Parliamentary Speeches. *German Politics*, 18(3), 385–402.  
<https://doi.org/10.1080/09644000903055823>

*Weiterführende Literatur*

Proksch, S.-O., & Slapin, J. B. (2009). How to Avoid Pitfalls in Statistical Analysis of Political Texts: The Case of Germany. *German Politics*, 18(3), 323–344.

Sitzung 12: Text-as-Data IV: Sentiment Analyse (synchron)

4.2.2021

Online-Seminar: Vortrag von Tobias Widemann und Vicente Valentim (in Englisch)

Extra-Datacamp: <https://learn.datacamp.com/courses/sentiment-analysis-in-r>

*Literatur:*

Valentim, V., & Widmann, T. (2020). Does Radical-Right Success Make the Political Debate More Negative? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3590343>

*Weiterführende Literatur:*

Widmann, T. (2020). *Beyond Valence , Beyond a Bag of Words : Creating and Comparing Dictionary and Word Embedding Models to Measure Discrete Emotional Appeals in German Political Communication*.

Rauh, C. (2018). Validating a sentiment dictionary for German political language—a work-bench note. *Journal of Information Technology and Politics*, 15(4), 319–343.  
<https://doi.org/10.1080/19331681.2018.1485608>

Sitzung 13: Text-as-Data V – Discursive Network Analysis

11.2.2021

Video: Discursive Network Analysis in R

Onlinesprechstunde: Bitte um vorherige Anmeldung

*Literatur:*

Leifeld, P. (2017). Discourse Network Analysis: Policy Debates as Dynamic Networks. In *The Oxford Handbook of Political Networks* (pp. 301–326). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780190228217.013.25>

Sitzung 14: Offen – Eure Fragen, Themen, Interessen (synchron)

18.2.2021

**Deadline:** *Exposee für die eigene Hausarbeit (1 Seite A4 Max.)*

Sitzung 15: Abschluss und Peer-Feedback (synchron)

25.2.2020

Online-Seminar: Diskussion der Exposees + Peer Feedback Runde