FULogo_Ausdruck_WORD_RGB.wmfDr. Christoph Nguyen

Otto-Suhr-Institut für Politikwissenschaft

Politisches System Deutschlands

Ihnestraße 22, 14195 Berlin

Mail: [christoph.nguyen@fu-berlin.de](mailto:christoph.nguyen@fu-berlin.de)

# Sommersemester 2020: 15340 -  Techniken der Empirischen Parlamentsanalyse in R

Montag, 14-16 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.

### Allgemeine Sprechstunde

Freitag 17:00-19:00 via Webex. Leider können in diesem Semester Sprechstunden nur digital angeboten werden. Bitte vereinbaren Sie vorab einen Termin unter <https://calendly.com/cgnguyen/sprechstunde>.

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

### Kommunikation

Die Kommunikation im Kurs findet via Blackboard, Email und Webex Teams statt. 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 Synchrone Online Diskussionen via Webex. Begin um 14:15.

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: <https://fu-berlin.webex.com/fu-berlin-en/j.php?MTID=m08af10b25b1bfde1936aa6716141a56f>

* Sprechstunde In Wochen ohne synchrone Online-Diskussionen stehe ich Ihnen für

kurspezifische Kleingruppensprechstunden (mit maximal 4 Teilnehmer\*Innen) zur Verfügung. Bitte nutzen Sie für die Anmeldung in dieser Sprechstunde diesen gesonderten Link <https://calendly.com/cgnguyen/parlament>.

* Problem-Sets Um kursspezifische Anwendungsbeispiele der Methoden zu

ermöglichen, werden im Laufe des Semesters drei Problem-Sets angeboten. Geben Sie diese bitte **bis Sonntagabend (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. Um sich für diesen Dienst zu registrieren, schicken Sie mir bitte eine Email.

### 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](https://wikis.fu-berlin.de/display/webexwiki/Zugang+zu+Webex+Meetings%2C+Training+und+Events).

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.

* 1. 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.
  2. Nehmen Sie Rücksicht auf die Sprecherin oder den Sprecher und unterbrechen Sie nur im Notfall. Selbst bei relativ schnelle Verbindungen gibt es oft Verzögerungen und Probleme, wenn mehr als eine Person spricht. Es wird empfohlen, die „Handheben“ – Funktion in Webex zu nutzen.
  3. Wenn die Bandbreite für eine stabile Verbindung nicht ausreicht, versuchen Sie zuerst Ihre Videokamera auszuschalten.
  4. Wählen Sie sich 5 bis 10 Minuten vor Begin des Seminars in den virtuellen Seminarraum ein. Den Link für den virtuellen Seminarraum finden Sie auf Blackboard
  5. Stellen Sie sicher, dass Sie an einem ruhigen Ort ohne viele Hintergrundgeräusche und mit stabiler Internetverbindung sind.

## Leistungsanforderungen im Seminar

### Aktive Teilnahme:

Abschluss der Aufgaben in Datacamp + Problem Sets 1 - 3

### Leistungsschein:

**Deadlines:**

29.6.2020: Abgabe der Exposees + Anmeldung für die Hausarbeit bitte digital an [christoph.nguyen@fu-berlin.de](mailto:christoph.nguyen@fu-berlin.de)

30.9.2020: Abgabetermin – bitte digital an [christoph.nguyen@fu-berlin.de](mailto:christoph.nguyen@fu-berlin.de)

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

Oder

1. 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’Reily Media. Kostenfrei online [hier](https://r4ds.had.co.nz/).

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

## Seminarplan

Sitzung 1: Einführung und Vorstellung 20.4.2020

Video: Willkommen im Kurs (Ablauf, Logistik)

Online-Seminar: Vorstellung und Q&A.

Sitzung 2: Einführung in das Statistikprogram R Sitzung I 27.4.2020

**Deadline: *Registrierung bei Datacamp – bitte schicken Sie mir eine E-Mail damit ich Sie registrieren kann***

Datacamp: Introduction to R (Beta)

Video: R Installation + Projektorganisation in R

Onlinesprechstunde: Bitte um vorherige [Anmeldung](https://calendly.com/cgnguyen/parlament)

Extra-Datacamp: Intermediate R

Sitzung 3: Einführung in das Statistikprogram R Sitzung II: Tidyverse 4.5.2020

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

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

Video: Deskriptive Daten am Beispiel Föderalismus, Dezentralisierung und

Performanz im internationalen Vergleich

Onlinesprechstunde: Bitte um vorherige [Anmeldung](https://calendly.com/cgnguyen/parlament)

*Literatur*

Ehlert, Niek, Annika Hennl, and André Kaiser. 2007. “Föderalismus, Dezentralisierung Und Performanz. Eine Makroquantitative Analyse Der Leistungsfähigkeit Territorialer Politikorganisation in Entwickelten Demokratien.” *Politische Vierteljahresschrift* 48(2): 243–68.

Sitzung 4: Anwendung I: Deskriptive Daten 11.5.2020

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

Online-Seminar: Diskussion der deskriptiven Ergebnisse

Diskussion Problem Set I

Sitzung 5: Daten Visualisieren mit ggplot2 18.4.2020

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

Onlinesprechstunde: Bitte um vorherige [Anmeldung](https://calendly.com/cgnguyen/parlament)

Extra-Datacamp: Interactive Data Visualization with plotly in R

Extra-Datacamp: Visualizing Geospatial Data in R

Sitzung 6: Fragestellungen und Theorien 25.5.2020

Video: Daten visualisieren am Beispiel Legislatives Verhalten

Onlinesprechstunde: Bitte um vorherige [Anmeldung](https://calendly.com/cgnguyen/parlament)

*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 Bröchler, 13-36. Wiesbaden: VS Verlag für Sozialwissenschaften.

Sitzung 7: Anwendung 2: Visualisierungen 8.6.2020

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

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

Sitzung 8: Regressionsanalyse 15.6.2020

Datacamp: Modeling with Data in the Tidyverse

Video: Regressionsanalysen am Beispiel

Onlinesprechstunde: Bitte um vorherige [Anmeldung](https://calendly.com/cgnguyen/parlament)

Sitzung 9: Anwendung 3 – Regressionsanalysen 22.6.2020

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

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

Sitzung 10: Text-as Data I 29.6.2020

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

Datacamp: Introduction to text analysis in R (Chapter 1-3)

Online-Seminar: Offen

Sitzung 11: Text-as-Data II 6.7.2020

Video: Germaparl und der Deutsche Bundestag

Online-Seminar Diskussion der Exposees

Sitzung 12: Text-as-Data III – Discursive Network Analysis 13.7.2020

Video: Discursive Network Analysis in R

Online-Seminar Offen