

Seminar

Empirische Studien mit R (M)

VORBESPRECHUNG, DATUM: 21.07.2022

PROF. DR. SEBASTIAN KRANZ,
THERESA GRÄFE, SIMON MAIER

Wichtige Termine

- Frist Abgabe der Präferenz zu Forschungspapieren: 28.07.2022
- Betreuung durch Theresa Gräfe
- Formale Anmeldung im Hochschuldiensteportal: **25.11.2022**
 - Anmeldefristen beachten! 4 Tage zuvor anmelden
- Voraussichtliche Präsentation: Ende Januar / Anfang Februar, genauer Termin wird noch bekannt gegeben

Ablauf

- Vorbesprechung
- Abgabe eurer Präferenzen über die Forschungspapiere / Vorschlag eines Forschungspapiers
- Zuteilung eines Forschungspapiers
- Paper gut durchlesen, Daten herunter laden (alle Daten sind frei zugänglich)
- Termin mit Betreuer vereinbaren (spätestens September / Anfang Oktober gerne früher)
- Abgabe der Vortragsfolien und des gut kommentierten RMarkdown Files 2 Tage vor Seminar
- Vorstellung im Seminar

Anforderungen

- Replikation oder Erstellung einer eigenen Graphik mit R (gerne auch mehrere)
- Abgabe des gut kommentierten R Codes in einem RMarkdown Dokument mit der Replikation der Graphen und Regressionen. Der Code sollte so geschrieben sein, dass wir die Ergebnisse replizieren können, wenn wir die Rohdaten verwenden, d. h. inkl. Data Cleaning
- Präsentation der Kernergebnisse des Forschungspapiers inkl. der selbst erstellten Ergebnisse im Seminar
- Replikation oder Erstellung einer eigenen Regression mit R (gerne auch mehrere)

Kerninhalte

Fragestellung	Welche Kernfrage untersucht die Studie? Kurzer Einstieg ins Thema. Was ist der genaue kausale Effekt, der geschätzt werden soll?
Daten	Was sind die Daten? Woher stammen sie? Ort und Zeit. Zeigen Sie ein paar Zeilen des wichtigsten Datensatzes. Zeigen Sie auch deskriptive Statistiken und Diagramme die einen Überblick / Beispiel der Daten geben.
Methodik	Wie schätzt die Studie den Effekt? Warum ist ein einfacherer Ansatz ggf. problematisch (Endogenitätsprobleme)?
Kernergebnisse	Zeigen und interpretieren Sie das Kernergebnis. Meist eine Regression, manchmal auch grafische Darstellung. Artikel haben oft sehr volle Tabellen mit verschiedenen Regressionen. Konzentrieren Sie sich auf das Kernergebnis. Erstellen Sie eine eigene Tabelle in der man die wichtigen Koeffizienten klar und groß erkennen kann. Interpretieren Sie die Effektgröße und auch die Genauigkeit des Schätzers. Wenn Zeit, können Sie auch weitere Ergebnisse und Robustheitschecks ansprechen.
Diskussion	Gibt es Probleme bei der Studie? Wie sähe idealer Datensatz für weitere Studie aus? Welche Fragen sind noch offen?

Weitere Aspekte

- Quellenangaben (siehe Beispielpräsentation)
- Geben Sie genau an welche Grafiken / Analysen entnommen wurden und welche Sie selbstständig repliziert haben.
- Gerne eine Diskussionsfrage am Ende des Vortrages
- Literaturverzeichnis
- Ggf. Anhang für wichtige weiterführende Analysen die zeitlich nicht in den Hauptvortrag passen.

Rahmenbedingungen Präsentation

- Gesamtvortragszeit 40 Minuten (25min Redezeit, 15min Diskussion)
- Gängige Dateiformate (PPTX/PDF) verwenden
 - Wenn Sie Freude an RMarkdown haben, können Sie auch gerne das R Package *xaringan* nutzen, um mit RStudio HTML Folien zu erstellen (z. B. hat Alexander Rieber alle Folien des Datascience Projektkurses damit erstellt.)
- Bis spätestens 2 Tage vor Seminartermin schicken Sie die Präsentation und das Rmarkdown-File per Email an uns (genaue Frist wird per Moodle bekannt gegeben)

Auswahl des Forschungspapiers

Zur Auswahl des Forschungspapiers gibt es zwei Möglichkeiten:

1) Auswahl aus der vorgeschlagenen Liste: Link zur Themenwahl:
<http://econ.mathematik.uni-ulm.de:9595/semmel/?rank=tltcvf>

2) Eigener Vorschlag (siehe auch Email)

Selbst ausgewählter Artikel muss folgende Kriterien erfüllen:

- Daten müssen verfügbar sein
 - Suche mit ejd.econ.mathematik.uni-ulm.de wird empfohlen
- Es gibt noch kein RTutor Problem Set zu dem Paper
 - Übersicht zu RTutor Problem Sets hier: <https://github.com/skranz/RTutor>
 - Optional: Es gibt Graphen im Paper, die gut replizierbar sind (ansonsten selbst neuen Graphen erstellen) und es muss mindestens eine Regression replizierbar sein.

Bewertungskriterien

Darstellung Forschungspapier

- Fakten sind korrekt
- Methodik und Ergebnisse sind verständlich dargestellt

Ergebnisse in R

- Ergebnisse sind korrekt
- R Code ist gut kommentiert
- R Code ist gut verständlich / gut lesbar, Empfehlung *dplyr*
- Anspruchsniveau der Analysen

Diskussion

- Richtige Antworten
- Gut vorbereitet und souverän antworten

Foliensatz

- Übersichtliches Layout
- Sinnvolle Grafiken und Tabellen
- Quellen (min. -0.3 sonst)
- Gute Rechtschreibung

Vortrag

- Freies Sprechen
- Selbstbewusstes Auftreten
- Körpersprache und Stimme

Formales

- Zeitvorgabe eingehalten
- Präsentation rechtzeitig eingereicht
- R Code rechtzeitig eingereicht
- Angaben auf Titelfolie vollständig

Fragen?
