

Seminar

Empirische Analysen/Studien mit R

VORBESPRECHUNG, DATUM: 15.02.2024

PROF. DR. SEBASTIAN KRANZ,
SIMON MAIER, FELIX WALDHANS

Wichtige Termine

- Erste Frist Abgabe der Präferenz zu Forschungspapieren: **16.02.2024 um 12 Uhr mittags**
- Zweite Frist Abgabe der Präferenzen: **10.03.2024 um 12 Uhr mittags**
- Formale Anmeldung im Portal Studium und Lehre: **30.04.2024**
- Abgabe der Folien für die Präsentation, der RMarkdown und der HTML Datei: bis spätestens **07.06.2024 um 13 Uhr**
- Präsentation: **12.06.2024 und 19.06.2024 in O29 – Raum 1003**

Ablauf

- Vorbesprechung
- Abgabe eurer Präferenzen über die Forschungspapiere / Vorschlag eines Forschungspapiers
- Zuteilung eines Forschungspapiers
- Papier gut durchlesen, Daten herunterladen (alle Daten sind frei zugänglich)
- Vereinbarung eines ersten Besprechungstermins mit dem Betreuer mit Gliederungsvorschlag, evtl. erster replizierter Grafik/Regression (spätestens Ende April, gerne früher)
- Abgabe Vortragsfolien, RMarkdown und HTML
- Vorstellung im Seminar

Anforderungen

- Replikation oder Erstellung einer eigenen Grafik und einer Regression mit R (zum Bestehen des schriftlichen Teils - typischerweise aber mehrere, s.u.)
- 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 Kerninhalte des Forschungspapiers einschließlich der selbst erstellten Ergebnisse im Seminar
- Beispiele für gelungene Abgaben finden Sie auf unserer Seminareseite:

<https://www.uni-ulm.de/mawi/mawi-wiwi/forschung-und-lehre/seminare/empirische-analysen-mit-r-b-und-empirische-studien-mit-r-m/>

Kerninhalte

<u>Fragestellung</u>	Welche Kernfrage untersucht die Studie? Kurzer Einstieg ins Thema. Was ist der genaue kausale Effekt , der geschätzt werden soll?
<u>Daten</u>	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.
<u>Methodik</u>	Wie schätzt die Studie den Effekt? Warum ist ein einfacherer Ansatz ggf. problematisch (Endogenität)? Warum gehen die Autoren wie vor?
<u>Kernergebnisse</u>	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.
<u>Diskussion</u>	Gibt es Schwachpunkte oder Probleme bei der Studie? Wie sähe ein idealer Datensatz für weitere Studien aus? Welche Fragen sind noch offen?

Weitere Aspekte

- Quellenangaben (siehe Beispiel-Präsentationen)
- Geben Sie sowohl in der HTML als auch im Foliensatz genau an welche Grafiken / Analysen aus dem Papier entnommen wurden und welche Sie selbstständig repliziert haben (Quelle)
- Diskussionsfrage am Ende des Vortrages
- Literaturverzeichnis
- Ggf. Anhang für wichtige weiterführende Analysen die zeitlich nicht in den Hauptvortrag passen (Vorbereitung auf mögliche Rückfragen)
- Benutzung von Chat GPT, Bing Chat, etc.: unter Hilfsmittel aufzählen

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 (Beispiel: Folien des Projektkurses Data Science & Business Analytics)
- In der Regel: ca. **eine Woche** (genaue Frist wird bekannt gegeben) **vor dem Seminartermin** schicken Sie die Präsentationsdatei, die RMarkdown-Datei und die zugehörige HTML-Datei **per Email** an uns

Auswahl des Forschungspapiers

Zur Auswahl des Forschungspapiers gibt es zwei Möglichkeiten:

- 1) Auswahl aus der vorgeschlagenen Liste (Link per E-Mail)
- 2) Eigener Vorschlag in Absprache mit dem Betreuer

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
- Eigene Diskussionsfrage vorbereiten

Foliensatz

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

Vortrag

- Freies Sprechen
- Selbstbewusstes Auftreten
- Körpersprache und Stimme

Formales

- Zeitvorgabe eingehalten (Abgaben)
- Angaben auf Titelfolie vollständig

Fragen?
