

Syllabus “Einführung in die Netzwerkanalyse” (33335f)

Felix Lennert

Sommersemester 2020

Kontakt

Da der komplette Kurs online stattfinden wird, werden wir nur digital kommunizieren.

- E-mail: felix.lennert@liu.se – bitte benutzt meine LiU Mail Adresse. Ich schaffe es leider nicht, den Regensburger Mail-Account in mein Mailprogramm so einzubinden, dass alles zuverlässig funktioniert.
- Webseite: **TBA** – meine Webseite ist gerade “work in progress”. Sobald sie fertig ist, kannst Du dort alle Materialien finden. Weiterhin werde ich auch alles bei GRIPS einstellen.
- Sprechstunden: freitag, 10-11. Bitte trage Dich hier ein. Wenn niemand angemeldet ist, werde ich den Zoom Raum auch nicht öffnen. Wenn Du während der Sprechzeiten nicht kannst, schreib mir eine Mail. Zoom Meeting ID: 911 0279 5820, Passwort: office. Link für meeting hier
- Veranstaltungen finden montags und donnerstags von 16-18 Uhr statt. Zoom Meeting Link – Meeting ID: 972 9681 0305, Passwort: bigdata

Kursbeschreibung

”Social network analysis takes as its starting point the premise that social life is created primarily and most importantly by relations and the patterns formed by these relations” (Marin/Wellman 2011: 11).

Die soziale Netzwerkanalyse (SNA) ist eine Methode der empirischen Sozialforschung zur Erfassung und Analyse sozialer Beziehungen und sozialer Netzwerke. Die SNA propagiert eine bestimmte Sichtweise auf soziale Phänomene, die deren relationalen Charakter betont. Verbindungen und Interdependenzen zwischen Einheiten (z. B. Personen oder Organisationen) stehen im Vordergrund, nicht deren individuelle Attribute und Eigenschaften. Soziale Beziehungen und deren Struktur werden so zur Analyseeinheit selbst.

R ist eine Programmiersprache, welche sich seit ihren Anfängen in den 1990er-Jahren wachsender Beliebtheit unter Sozialwissenschaftlern erfreut. R ist Open Source, somit stehen viele verschiedene packages für mannigfaltige Anwendungen zur Verfügung, und überdies hinaus kostenlos. Zwar liegt der Fokus des Kurses auf dem Kennenlernen und der praktischen Anwendung von Methoden aus dem Bereich der Netzwerkanalyse, den größten Teil ihrer Zeit werden die Studierenden allerdings für das Erlernen von R aufwenden. Hier wird nicht nur eine kurze Einführung in die Grundbefehle gegeben, welche man zur grundlegenden Analyse von Netzwerken benötigt. Vielmehr wird ein ganzheitlicher Ansatz verfolgt. Zu Beginn werden die Basics erklärt – wie sieht die Benutzeroberfläche aus, welche grundlegenden Operationen können durchgeführt werden etc. Anschließend werden die Studierenden in RStudio, RMarkdown und das sogenannte tidyverse eingeführt. Während ersteres ein sogenanntes GUI darstellt, welches die Funktionalität von R deutlich erweitert, stellt zweiteres einen Texteditor dar, welcher universal einsetzbar ist. Das tidyverse hingegen ist eine Sammlung von Programmen, welche eine gemeinsame Philosophie und Befehlsstruktur teilen: Daten müssen tidy sein. Diese wird erläutert, die Vor- und Nachteile diskutiert und anschließend die Handhabung der verschiedenen Pakete eingeübt. Am Ende des Kurses werden die Studierenden fähig sein, Netzwerkdaten für ihre Forschungszwecke fruchtbar zu machen und die Ergebnisse in einer ansprechenden Form zu präsentieren.

Da die Veranstaltungen zwei mal pro Woche stattfinden, können die Inhalte grob aufgeteilt werden: Montags geht es um "praktische" Dinge in R, in den Donnerstagssitzungen können anfangs Fragen hierzu gestellt werden, anschließend gibt es theoretischen Input zu sozialer Netzwerkanalyse. Vorkenntnisse in R, SPSS, Python oder Stata sind von Vorteil, aber definitiv kein Muss. Wichtige Voraussetzung ist vielmehr die Bereitschaft, sich mit fortgeschrittenen sozialwissenschaftlichen Forschungsmethoden zu beschäftigen. Abgesehen davon ist ein eigener Rechner erforderlich.

Ziele

Das Ziel dieses Kurses ist in erster Linie, den Studierenden eine Einführung in die Netzwerkanalyse zu bieten. Um derartige Analysen durchzuführen, werden die Programmiersprache R und die IDE RStudio verwendet. Somit kann der Kurs in zwei Teile aufgespalten werden: Einen theoretischen, in dem die Hintergründe sozialer Netzwerkanalyse besprochen werden, und einen praktischen, in dem zuerst Grundkenntnisse in R und anschließend praktische Netzwerkanalyse in R vermittelt werden.

R

Einführung in R

Dieser Kurs bietet Dir eine Einführung in die R-Sprache. R ist eine Programmierumgebung für statistische Analysen und Grafiken. Bonus: es ist kostenlos. Wir werden in R unter Verwendung der IDE RStudio programmieren, die ebenfalls kostenlos ist. R und RStudio haben einige nette Funktionen, wie RMarkdown¹, RStudio Projects oder eine GitHub-GUI. Sie werden ebenfalls vorgestellt werden.

Einführung in das tidyverse

Oft sind die Daten, die wir für Soziale Netzwerkanalyse brauchen, nicht bereits in der richtigen Form. Somit ist es essentiell, dass Du weißt, wie man die Daten in die richtige Form bringt. Hierbei kannst Du die Pakete des sogenannten tidyverse benutzen.

Visualisierung mit ggraph

Ein großer Bestandteil von Netzwerkanalysen ist die graphische Darstellung. Du wirst lernen, wie Du das mit ggraph machen kannst.

“Functional programming” in R

Sobald Du Code kopierst, läufst Du Gefahr, Fehler zu machen. Hierbei hilft “Functional programming”: Alles, was Du öfter als einmal machen musst, kann zu einer Funktion werden.

Einführung in SNA

- Geschichte der SNA
- Deskriptive Maßzahlen

Einige Regeln

Hier findest Du ein paar allgemeine Richtlinien.

¹mit welchem dieser Syllabus erstellt wurde.

Grundlegende Verhaltensregeln

- Einerseits: Schalte Dein Mikrofon während der Sitzungen immer stumm.
- Andererseits: Unterbreche mich, wann immer Du etwas nicht verstanden hast.
- Mach Dich über Deine Kommilitonen nicht lustig.
- Wenn Dich etwas stört, schreib mir eine Mail.
- Bitte kopiere Code aus dem Internet. Ich sehe das nicht als Plagiat an. Sollte jemand das ernsthaft als Plagiat ansehen, hat er oder sie wahrscheinlich noch nie programmiert.

Teilnahme

In diesem Sommer ist die Teilnahme am Kurs wahrscheinlich einfacher denn je – Du kannst theoretisch einfach im Bett bleiben und trotzdem am Kurs teilnehmen. Wenn Du aufgrund technischer Probleme nicht teilnehmen kannst, schicken mir einfach eine E-Mail. Ich kann Dir dann Informationen darüber geben, wie Du Dich über Dein Mobiltelefon einwählen kannst

Die Sitzungen, die mit R gekennzeichnet sind, sind in der Regel für R vorgesehen. Ich zeige Euch einige Sachen auf meinem Laptop und beantworte gerne Fragen dazu. Im Grunde genommen gehe ich nur das Skript durch, welches im Vorfeld hochgeladen wird. Insofern kannst Du es Dir gerne vorher bereits durchlesen und dann gezielt Fragen stellen. Ich gebe während der Sitzung keine Zeit, die Sachen selbst auszuprobieren – mach das doch einfach danach. Im Gegenzug kann es sein, dass einige Sitzungen nicht die gesamten 1,5 Stunden dauern. Am Ende der Skripte findest Du Übungen, die sich auf die Dinge beziehen, die ich behandelt habe. Gerne schaue ich mir Deine Lösungen an und spreche während der Sprechstunde (die natürlich auch verlängert werden kann) darüber. Bitte kontaktiere mich vorher und füge auch den Code an.

Solltest Du an der “R-Sitzung” nicht teilnehmen (können), dann bist Du verpflichtet, die angefügten Übungen der verpassten Sitzung zu machen und mir bis zur darauffolgenden Sitzung zuzusenden.

Es ist mir prinzipiell egal, ob Du mir jede Woche etwas schickst – solange Du an den R-Sitzungen teilnimmst. Es wird allerdings ein “take-home assignment” geben

E-mail Regeln

E-mails werden wohl unser bevorzugter Kommunikationsweg sein. Bevor Du mir eine E-mail mit einem Problem schickst, versuche bitte zuerst folgende Strategien, um Deine Fragen zu beantworten/Probleme zu lösen:

- Schau Dir den Syllabus nochmal an.
- Google Dein Problem. Versuche mindestens drei Lösungen von Stackoverflow.
- Schau Dir das Skript nochmal an. In der Regel ist hier alles enthalten.

Wenn das alles nicht geholfen hat, schicke mir doch gerne eine E-mail. Beachte dabei bitte Folgendes:

- Beginne den Betreff mit *snamr* – so weiß ich, wem ich zuerst zurückschreiben muss.
- Wenn das Problem mit R zu tun hat, füge bitte ein *reprer* bei.
- Versuche, Dein Problem genau zu beschreiben.

Prüfungsleistungen

Du musst drei Dinge für den Kurs machen

- Eine “data challenge”. Sobald die Einführung in Data Wrangling abgeschlossen ist (sprich, am 22.06.2020), werde ich Euch einen Datensatz geben, den Ihr dann innerhalb einer Woche in Form für eine Netzwerkanalyse bringen müsst. Ihr könnt gerne Zweiergruppen bilden.

- Ein Referat. Themen: Dichte, Transitivität, Reziprozität, Degree, Zentralität, PageRank, Homophilie, HITS-Algorithmus. Aufbau des Referats: Vorstellung der Maßzahl; Vorstellung einer Studie, die sie anwendet; Beispiel in R – Dauer: ~15 Minuten, PowerPoint-Folien sind Pflicht; Bitte sende mir die Präsentation und das R-Skript spätestens am Vorabend zu!
- Ein kurzes Paper (~5 Seiten). Such Dir einen Datensatz, denk Dir eine Forschungsfrage aus und wende das Gelernte an. Kann in Zweiergruppen geschrieben werden. Deadline ist der letzte Tag des Semesters (30.09.2020). Schreib mir eine Mail mit “Ich schaffe es nicht” und Du bekommst eine Woche obendrauf. Weitere Verlängerungen gibt es nur mit einem “ernsthaften” Grund (z.B. Praktika, Krankheit).

Übersicht

2020-06-08: R

Let's get it started!

- Einführung in R, Rstudio, RStudio Projects, RMarkdown, GitHub
- R als Taschenrechner
- Vektoren
- Das `tibble`

2020-06-11

Fronleichnam. Hab' einen schönen Tag!

2020-06-15: R

Let's get into (tidy) data!

- Das Konzept (Wickham, Hadley 2014: Tidy Data. In *Journal of Statistical Software* 59(10). Verfügbar hier)
- Wie sieht das für SNA aus?
- Datenimport: `readr`, `haven`, `readxl`, `vroom`
- Datenspeicherung
- Die `pipe`
- `tidyr`
- Die verschiedenen `tidyverse` Pakete

Sitzung wird wohl kürzer sein. Sollten bereits Fragen aufkommen, kannst Du sie gerne nach der Sitzung stellen.

2020-06-18

Einführung in die soziale Netzwerkanalyse

- Geschichte
- Prinzipien
- Repräsentation

Literatur

- Borgatti, Stephen/Mehra, Ajay/Brass, Daniel/Labianca, Giuseppe, 2009: Network Analysis in the Social Sciences. In *Science* (323), S. 892-895.
- Freeman, Linton 2004: *The Development of Social Network Analysis. A Study in the Sociology of Science*. Vancouver: Empirical Press.

2020-06-22: R

Data Wrangling

- `dplyr`
- `lubridate`
- `forcats`

Die Sitzung wird wohl ziemlich umfangreich und intensiv werden. Sollte Bedarf bestehen, können am darauffolgenden Donnerstag nach der Sitzung gerne Fragen gestellt werden.

2020-06-25: R

Einführung in Netzwerkanalyse in R

- `tidygraph`

2020-06-29: R

Visualisierung + take-home assignment

- `ggraph`
- take-home assignment wird ausgegeben

2020-07-02: R

“Functional programming”

- Grundprinzipien
- Funktionen
- Loops
- Flow-control

2020-07-06

Referate:

Thema	Referent*in
Dichte	Nico
Transitivität	Nariman
Reziprozität	Elfi
Degree	Jakob

Relevante Literatur:

- Newman, Mark 2010: Networks. An Introduction. New York: Oxford University Press Inc.

2020-07-09

Referate:

Thema	Referent*in
Zentralität	Sebastian
Homophilie	Janine
PageRank	Theresa
HITS-Algorithmus	Nele

Relevante Literatur:

- Newman, Mark 2010: Networks. An Introduction. New York: Oxford University Press Inc.

Hausaufgabe

- Brainstorming für Paper
- Gerne in Zweiergruppen

2020-07-13

Deadline take-home assignment

Lass uns über Dein Paper sprechen

- Ideen?
- Gruppen?

2020-07-16

Wrap-up

- Feedback
- Ausblick