GESIS Fall Seminar in Computational Social Science 2024
“Introduction to Social Network Analysis with R”

Lecturer: Philip Leifeld  
Affiliation: University of Manchester, Department of Social Statistics  
Email: philip@philipleifeld.com

Date: September 16-20, 2024  
Time: 09:30-16:30

About the Lecturer
Philip Leifeld is a professor in the Department of Social Statistics at the University of Manchester. He was previously a professor at the University of Glasgow and in the Department of Government at the University of Essex. Philip is best known for his work on discourse network analysis, which combines content analysis with network analysis to model policy debates as complex social systems. Philip has also published on statistical models for social networks, particularly longitudinal models, and applications of network science to political institutions and behaviour and public policy. He has also worked with the finance sector in London to implement network analysis for a better understanding and prediction of investment relations between limited and general partners. Philip is the author and maintainer of several software packages including the Java software Discourse Network Analyzer and the R packages btergm and texreg.

Course Description
Social network analysis focuses on the analysis of relations between nodes. Network analysis can be used to model relationships between individuals (e.g., friendship, co-authorship, social contagion of behaviour), organisations (e.g., policy networks, inter-firm relations), countries (e.g., alliances and conflict, migration, trade), variables (e.g., correlation), and other entities. While network analysis has a long history in the social sciences, it has also become a central method for the study of complex systems in science more widely over the last two decades. Social network analysis is a key method for understanding the transmission of behaviour between entities (“social influence”) and how those who share similarities tend to develop ties (“homophily”).

In this course, we will cover the foundations of social network analysis, including the description and visualisation of networks, different types of networks and data formats, measuring centrality in networks, identification of communities or subgroups within networks, and the basic foundations of statistical modelling of networks, which will also enable participants to confidently take an advanced follow-up course on inferential network analysis.

The course is beginner-friendly and requires basic R skills but no prior training on networks or in statistics. We will cover both the methods and their software implementation and will also schedule applied practical sessions where participants can apply their new knowledge to empirical datasets. Those looking for a more advanced course should consider taking “Advanced Social Network Analysis” (23-27 September).

The morning sessions are delivered as lectures and introduce key concepts and methods. The afternoon sessions are a mix of guided software lab sessions using R and practical application sessions with supervised group work. The R sessions rely on several packages, including the statnet suite of packages, btergm, and igraph.

Organizational Structure of the Course
The three-hour morning sessions will introduce key concepts and methods in a lecture format. The three-hour afternoon sessions will be a mix of R software tutorials and group-based practical work on exercises using empirical datasets provided by the instructor. Both the morning and afternoon sessions will be led and supervised by the course instructor. Participants will also be given the opportunity to discuss their own research applications in class,
and there will be time to schedule a selection of voluntary mini-presentations by participants on applications of network analysis to their own research.

Keywords
social network analysis; network science; network visualisation; complex social systems; relational data

Target Group
You will find the course useful if:
▪ you need to analyse relational data or network data of any kind.
▪ you wish to expand your methodological knowledge by learning a set of broadly applicable data science techniques.
▪ you have experience in conducting social network analysis research using other tools but wish to learn doing SNA in R.

Course and Learning Objectives
By the end of the course you will:
▪ be able to format your data in the right way to load them into R and conduct network analysis with them.
▪ have an overview of network analysis methods and their implementation in R.
▪ competently choose between different analytical techniques for analysing empirical network data.
▪ have acquired the necessary knowledge and skills to conduct basic network analysis projects on your own, including conceptualisation of the research and execution and interpretation of the analysis in R.

Course Prerequisites
▪ Essential: Basic familiarity with R, including vectors, data frames, and lists.
▪ Desirable: Basic familiarity with RStudio.
▪ Desirable: Basic familiarity with research methods, such as measurement scales, descriptive statistics, and/or linear regression.

Software and Hardware Requirements
Participants should bring their own laptops and have R, RStudio (or an IDE of their choice), and the following R packages installed and ready for use: statnet; igraph; ggplot2; ggraph; graphlayouts; btergm.

Recommended Literature to Look at in Advance

Day-to-day Schedule and Literature
Day 1: Introduction
▪ Types of networks.
▪ Representations of networks.
▪ Applications.
▪ Research design considerations.
▪ Data collection and measurement.
Day 2: Description and visualisation of networks
- Managing network data.
- Descriptive statistics for networks.
- Dynamic visualisation of longitudinally changing networks.

Day 3: Node-level analysis and higher-order networks
- Node centrality.
- Two-mode networks.
- Combinations and affiliations.
- Multiplex, multi-layer, and multi-level networks.
- Weighted, temporal, and hyper-networks.

Day 4: Subgroup analysis
- Clique analysis.
- Blockmodelling.
- Structural similarity, graph clustering, and dimension reduction.
- Community detection.

Day 5: Introduction to Statistical Models for Networks
- Testing hypotheses about network structure: conditional uniform graph tests and the quadratic assignment procedure.
- Introduction to Exponential Random Graph Models.
- Overview of the variety of different statistical network models and implementations and their use cases.

Additional Recommended Literature