



Sequence Analysis Workshop

(digital)

SPS Third-term Workshop 2019-2020

Dates: Monday, 11 May 10:00-16:30

Tuesday, 12 May 10:00-16:30

Location: through zoom: registered participants will receive a link

[Register for this workshop](#)

[Organising administrative assistant: Maureen Lechleitner](#)

Instructor: Anette Fasang, Humboldt-University of Berlin & WZB Berlin Social Science Center

COURSE DESCRIPTION

This workshop introduces sequence analysis for social science research. Sequence analysis, originally developed in biology to analyze strings of DNA, is increasingly established for analyzing sequences of categorical states on the individual level, including family formation and employment careers. More recently, it is also applied in history, economics and political science, for example to historical sequences of social policy development or types of government using countries as the unit of analysis. This workshop covers basic techniques of sequence analysis, as well as recent methodological developments tailored at social science research questions. Topics include different ways of calculating similarity among sequences, cluster analysis after sequence analysis to identify typologies of trajectories, and sequence visualization with hands-on applications. We will also cover a basic overview of techniques for analyzing multidimensional and dyadic sequences and the sequence analysis multistate model that combines sequence analysis with event history modeling to estimate the impact of time-varying covariates on subsequent trajectories. All hands-on examples will use R's TraMineR and Weighted Cluster packages.

All readings and course materials will be available on the online platform.

This workshop is worth 10 credits. However, you are welcome to audit the course without credit. You are welcome to audit the course without credit. Requirements for 10 credits are:

- active participation in class
- five to seven page paper (12pt, 1.5 line spacing), in which you specify a research question and present sequence analysis results using the methods applied in class. You can use example data from class, or your own data. The paper is due by May 31st, 2020. I will be available for individual digital office hour consultations on May 14th.

Schedule:

Monday, 11 May:

10:00 - 11:30: Introduction to Sequence Analysis

- Basic introduction and history of sequence analysis in the social sciences
- Sequence analysis in the context of other longitudinal techniques
- Classic and recent applications

12.30 -14.00: Hand- on: Sequence Visualization and Optimal Matching

- Basic introduction to R
- Describing and visualizing sequences
- Application of optimal matching analysis

14.30-16.30: Hands-on: Independent application

- Repetition of applications with other example data, alone or in pairs/small groups of students

Tuesday, 12 May

10.00-11.30: Troubleshooting and clustering after sequence analysis

- Discussion of questions from independent application
- Cluster analysis and cluster quality criteria for typologies of trajectories

12.30-14.00: Hands-on: Cluster Analysis after Sequence Analysis

- Cluster analysis after sequence analysis
- Visualizing cluster-cut off criteria and clusters

15.30-16.30: Outlook and Frequently Asked Questions in applied sequence analysis

- Overview of advanced sequence analysis: Multidimensional and dyadic sequences, sequence analysis multistate model
- Challenges in applied sequence analysis and what you can (and can't) do about them
- Final discussion

Required reading

Abbott, Andrew. 1992. From Causes to Events: "Notes on Narrative Positivism." *Sociological Methods & Research* 20: 428-455.

Aisenbrey, S., & Fasang, A. E. (2010). New life for old ideas: The "second wave" of sequence analysis. Bringing the "course" back into the life course. *Sociological methods & research*, 38(3), 420-462.

Gabardinho, Alexis, Gilbert Ritschard, Nicolas Séverin Mueller, and Matthias Studer. 2011. "Analyzing and visualizing state sequences in R with TraMineR." *Journal of Statistical Software* 40: 1-37.

Studer, Matthias. 2013. "WeightedCluster Library Manual. A practical guide to creating typologies of trajectories in the social sciences with R." *LIVES Working Papers* 24. DOI: <http://dx.doi.org/10.12682/lives.2296-1658>. 2013.24.

Studer, M., & Ritschard, G. (2016). What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(2), 481-511.

Studer, M., Struffolino, E., & Fasang, A. E. (2018). Estimating the relationship between time-varying covariates and trajectories: The sequence analysis multistate model procedure. *Sociological Methodology*, 48(1), 103-135.

Recent applications (chose one you are particularly interested in)

Democratization and Electoral Participation (Basics of Sequence Analysis)

Buton, F., Lemerancier, C., & Mariot, N. (2014). A contextual analysis of electoral participation sequences. In *Advances in sequence analysis: Theory, method, applications* (pp. 191-211). Springer, Cham.

Casper, G., & Wilson, M. (2015). Using sequences to model crises. *Political Science Research and Methods*, 3(2), 381-397.

Family Demography (Multichannel sequence analysis)

Aisenbrey, S., & Fasang, A. (2017). The interplay of work and family trajectories over the life course: Germany and the United States in comparison. *American Journal of Sociology*, 122(5), 1448-1484.

Stratification (Dyadic sequence analysis)

Karhula, A., Erola, J., Raab, M., & Fasang, A. (2019). Destination as a process: Sibling similarity in early socioeconomic trajectories. *Advances in Life Course Research*, 40, 85-98.

Poverty (Sequence analysis multistate model)

Struffolino, E., für Sozialforschung, W. B., & Van Winkle, Z. (2019). *Is There Only One Way Out of In-work Poverty?: Difference by Gender and Race in the US*. WZB Discussion Paper 2019-601.

<https://pdfs.semanticscholar.org/abd0/0c8eed32c50dfdeef4515e49823223672f3f.pdf>