



## INTRODUCTION TO SEQUENCE ANALYSIS

### Profesorado:

Dr. Nicola Barban  
University of Oxford

### Objetivos

This course gives an introduction to the theoretical and practical concepts of sequence analysis. Sequence analysis, originally developed in biology to analyze strings of DNA, has attracted increasing attention in the social sciences for the **analysis of longitudinal data**. Most applications in the social sciences study life course processes, such as labor market careers, educational careers, or family formation. During the short course, we will discuss the usefulness of sequence analysis in applied social sciences as an holistic approach to investigate timing, quantum and sequencing of life course events. We will consider the practical implementation of these methods using available data. Concepts covered include the statistical representations of categorical time series, measures of sequence dissimilarity (i.e., Optimal Matching Algorithm, Longest Common Subsequence, Hamming distance); patterns identification in life course trajectories; classification techniques; criticisms to sequence analysis and new developments. Whereas the instructor will provide data sets, students are encouraged to bring their own data material.

The lab session will be conducted using R's TraMineR Package. Previous knowledge of R statistical software is appreciated but not required. The first lab session will be dedicated to introduce R environment. If you plan to use your laptop, please install the R software by going to <http://cran.r-project.org/>.

### Programa

Thursday 28 May: Introduction to life course analysis; Different approach for life course studies; Representation of trajectories as sequences; Introduction to sequence analysis; Example of applications; Sequence analysis and demography; Data; Optimal Matching; Other measures; Description of trajectories, indicators; Graphical representations.

Friday 29 May: Clustering and analysis of clusters; Alternative metrics for sequence analysis; Determinants of life course trajectories; Analysis of consequences of life course trajectories; Analysis of heterogeneity in life course. Advanced topics; Sequence analysis for causal inference.

### Observaciones:

Duración: 10 horas.

Fecha: 28 y 29 de mayo de 2015.

Horario: De 9:00 a 14:30 horas (con descanso de 30 minutos).

Lugar de impartición: Aula de formación del Instituto de Estadística y Cartografía de Andalucía. C/ Leonardo da Vinci, nº 21. Isla de la Cartuja. 41071-Sevilla.

Idioma de impartición: **Inglés**