Instructor details
Dr. Philippe Blanchard
PAIS Department
University of Warwick
p.blanchard@warwick.ac.uk

Short Biography
Philippe Blanchard, PhD in political science, is an Associate professor in political science at the University of Warwick, UK. He works on green politics, political communication, and methods for social and political sciences: multivariate statistics, longitudinal methods, interviewing, content analysis and digital data. He has taught methods and techniques for social and political sciences in France, Switzerland, Austria, Denmark, Singapore, the USA and the UK. He co-chairs the ECPR Standing group on Political Methodology and is currently co-director of a BA in Politics, International Studies and Quantitative Methods and a MA in Politics and International Studies: Big Data & Quantitative Methods.

Prerequisite knowledge
Two prior knowledges are requested:
- Basic applied statistics and use of data sets, through either an Excel-like software, statistics software like SPSS or Stata, or R. Participants who have never practiced any statistics should attend an introductory course or at least read an introductory textbook and practice by themselves.
- Basic knowledge of R: data manipulation and basic descriptive statistics. Most of the hands-on work will be done in R. R-beginners may prepare themselves on their own, or, ideally, by attending beforehand an introductory course prior to the course, such as the one offered in Budapest.

Short course outline
Sequence analysis is the systematic descriptive and causal study of sequences, that is, successions of standard categorical states or events. Sequence analysis is a unique method for representing, comparing and clustering sequences, for extracting prototypical sequences and for mining sequence populations. Its core tool, the optimal matching algorithm, imported from genetics and bio-computing, screens and discriminates longitudinal processes according to the nature of events, their duration and their order.
Numerous fields in the social and political sciences are concerned with sequences, for example: life course analysis (e.g. family and residential transition from youth to adulthood), sociology of professional careers (such as gendered careers or transition to retirement), political sociology (elite and activist careers), evolution of regimes (stages of development or transition to democracy), speech

---

1 Disclaimer: the information contained in this course description form may be subject to subsequent adaptations (e.g. taking into account new developments in the field, specific participant demands, group size etc.). Registered participants will be informed in due time in case of adaptations.
analysis (rhetorical strategies), geopolitics (stages in crises or democratic transition), comparative studies (stages of diffusion of reforms or mobilizations), elections (stages of public opinion formation during campaigns), human geography (distribution of space occupation from city centre) and ethnographic practices (rituals).

Long course outline
Sequence analysis (SA) is the systematic descriptive and causal study of sequences, that is, successions of standard states or events. It is a unique method for representing, comparing and clustering sequences, for extracting prototypical sequences, for describing groups of sequences and mining sequence populations. Its core tool, the optimal matching algorithm, imported from genetics and bio-computing, screens and discriminates longitudinal processes according to the nature of events, their duration and their order.

The course will be both theoretical and practical. A project will be provided by the instructor as a running example throughout the course. For participants who come with their own data and/or who have a precise research project in mind, the course will provide some help to set up a treatment strategy and to articulate SA with other methods.

The course is mainly designed for beginners in sequences, but researchers who have already used R or Stata on sequences may attend so as to improve their theoretical and/or practical knowledge, as well as take advantage of the confrontation with varied fieldworks from students and instructor.

Participants will be invited to do exercises on their own or by pairs if level or fieldwork match. Students are invited to bring their own laptop and to install R prior to the course (http://cran.r-project.org/), as well as the following packages: boot, cluster, colorspace, foreign, graphics, RcolorBrewer, rgrs and TraMineR. All kinds of hardware and operating systems fit.

The course is made of the following sections. Each might be developed more or less, according to the needs of the audience.

1. What is a sequence? This first part proposes definitions of sequences and subsequences, which are very common objects for social scientists at large, but seldom considered by social statisticians.

2. What are the principals of sequence analysis? After a short retrospective glance at the origins of SA, the second part develops the main statistical principles of the method.

3. What is the specificity of sequence analysis? Other longitudinal methods, popular in the social and political science, may attempt at treating longitudinal objects in the social sciences. Yet, none of these approaches can treat categorical time series in a comprehensive approach to time-related phenomena.

4. What data can be treated by means of sequence analysis? Several concrete examples are given, from numerous scientific fields. I also show how different kinds of data may best be investigated. Reference data from personal questionnaire surveys are given as examples, as well as large national database such as the Swiss Household Panel. Students are also encouraged to bring their own data, so that their suitability may be discussed collectively.

5. How should the data be prepared and coded? Due to the diverse origins of sequence data, they need to be cleaned and formatted so as to be properly processed.

6. How are the data processed? This part presents the core tools of sequence analysis, the optimal matching algorithm, its variants and its alternatives.

Sequence analysis is still an emerging method in the field of political science and social sciences at large. If you wonder about its usefulness for your own research project, do not hesitate to contact the instructor: p.blanchard@warwick.ac.uk.
Day-to-day schedule (Monday 1st to Friday 5th of August)

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>Definitions and principles</td>
<td></td>
</tr>
<tr>
<td>Day 2</td>
<td>Review and preparation of sequence data</td>
<td>Hands-on session 1</td>
</tr>
<tr>
<td>Day 3</td>
<td>Case study. Optimal matching (1)</td>
<td>Hands-on session 2</td>
</tr>
<tr>
<td>Day 4</td>
<td>How to make the best use of optimal matching</td>
<td>Hands-on session 3</td>
</tr>
<tr>
<td>Day 5</td>
<td>Sequence graphs and sequence statistics</td>
<td>If time allows and on request: Advanced issues</td>
</tr>
</tbody>
</table>

Day-to-day reading list

Participants should choose their priority readings in the list below according to their disciplinary interests. However all recommended readings are representative of some aspects of the method that are useful for all disciplines. It is strongly recommended to start reading before the course kicks off, so that you may focus during the course on exercises and in-class discussions.

Day 1

**Compulsory**


**Optional**


Day 2

**Compulsory**


**Optional**


8. An introductory course in R, for example:
   - see http://www.r-project.org/doc/bib/R-books.html for more possibilities, in several languages.
## Day 3

**Compulsory**


**Optional**


## Day 4

**Compulsory**


**Optional**


## Day 5

**Compulsory**


**Optional**


### Software requirements

R + following packages (to be installed prior to the course): boot, cluster, colorspace, foreign, graphics, RcolorBrewer, rgrs and TraMineR.

### Hardware requirements

Personal laptop.

One easily accessible power outlet per student.
Literature
See "Optional readings" in section 8.

Recommended training tracks
The following other ECPR Methods School courses could be useful in combination with this one in a ‘training track’. NB this is an indicative list.

Before this course

<table>
<thead>
<tr>
<th>Course title</th>
<th>Summer School</th>
<th>Winter School</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Research design</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2 Introduction to R</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3 Event History and Survival Analysis</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4 Introduction to Statistics for Political and Social Scientists</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After this course

<table>
<thead>
<tr>
<th>Course title</th>
<th>Summer School</th>
<th>Winter School</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Event History and Survival Analysis</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Disclaimer
The information contained in this course description form may be subject to subsequent adaptations (e.g. taking into account new developments in the field, specific participant demands, group size etc.).
Registered participants will be informed in due time in case of adaptations.

Note from the Academic Convenors
By registering to this course, you certify that you possess the prerequisite knowledge that is requested to be able to follow this course. The instructor will not teach these prerequisite items. If you are not sure if you possess this knowledge to a sufficient level, we suggest you contact the instructor before you proceed with your registration.