FROM ARCHIVES TO ALGORITHMS
## Quantitative Description for Social Scientists
Étienne Ollion (CNRS - Strasbourg)

In the course of their research, social scientists are often faced with quantitative data, or with qualitative data that could be quantified. Organization and administrative registers, bio- and bibliographic information, data about practices or responses to *ad hoc* surveys have long been used by researchers, and the ongoing digitization of everyday life makes such data, big or small, much more readily available for social scientific analysis.

The goal of this course is to offer an introduction to a form of social scientific investigation currently on the rise, quantitative description. Located between inferential statistics and qualitative approaches, these methods of quantification are increasingly used, for exploratory or confirmatory purposes.

The course will offer an introduction to some of the most important techniques in this area. It will show how to operationalize a research question and create a new dataset, and how to analyze this dataset through a range of techniques. We will examine techniques such as basic descriptive statistics, dimensionality reduction approaches (geometric data analysis), clustering, social network analysis, or sequence analysis. The course will also offer a brief overview of machine learning, in order to expound the main principles of this new class of statistical methods.

Each session will present both the mathematical and the ontological assumptions involved in the use of each technique. The approach will nonetheless be hands-on, as each class will feature a practice session meant to introduce the tools that concretely help researchers carry out their own project.
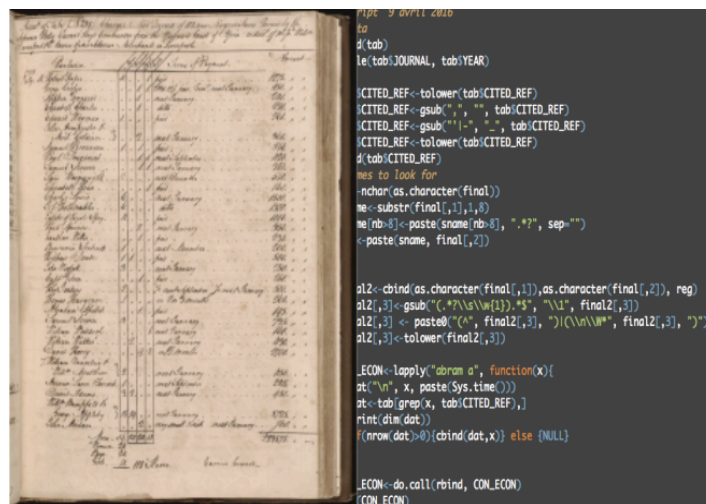
REQUIREMENTS
The course requires no preexisting statistical, skills. All the methods will be presented in plain English with limited mathematical excursus. The analyses will be carried out using the R software.

Sessions will take place every Monday and Friday from 8-10AM, room TBA
Mon, Oct 24th 2016: Quantitative Description for the Social Sciences
This session will present the history, merits and challenges of quantitative description in the social sciences. We will also go through a gentle introduction to R.

Fri, Oct 28th 2016: Principles of Descriptive Statistics
This session will introduce some classic tools for descriptive statistics (measures of spread, measures of central tendency) and their implementation in R. We will also reflect on visualization matters for data exploration.

Mon, Oct 31st 2016: Finding Relations: Dimensionality Reduction (1)
This session will offer an introduction to geometric data analysis. Born half a century ago, it has since then gained international recognition following its use by Pierre Bourdieu to map out social spaces. In addition to this possible use, we will see two other applications of the method: data exploration, and feature extraction.

Fri, Nov 4th 2016: Mapping out the Social Space: Dimensionality Reduction (2)
This session will pursue the presentation of dimensionality reduction techniques by looking at alternatives approaches rooted in machine learning. This will generate a discussion about the possible ways to construct the social space.

Mon, Nov 7th 2016: Making Classes: Clustering
How can observations be grouped together, and according to which principles? This session will introduce various clustering techniques, classic and more novel ones.

Fri, Nov 11th 2016: Capturing Explicit Ties: Network analysis (1)
A well-established area of social scientific inquiry, social network analysis has often been used for descriptive purposes. This session will serve as an introduction to this other form of relational analysis. It will offer a survey of this technique, describe the main strands in the field, and reflect on the merits and perils of visualization in SNA.

Mon, Nov 14th 2016: Capturing Explicit Ties: Network Analysis (2)
The second part of the SNA course will focus on two different though central aspects of SNA: i) various forms of measures and ii) community detection.

Fri, Nov 18th 2016: Quantifying Trajectories: Sequence Analysis
Initially developed for DNA decoding, sequence analysis examines the trajectories of entities (individuals, organizations, texts) in order to discover similarities between them. It will show how the technique can be used to uncover patterns in the trajectories (pathways leading to a certain state, within-group variations, etc.), hence becoming a very powerful descriptive tool.