Statistics and the Political Sociology of Quantification

François Briatte
Fall 2018

British author H. G. Wells famously declared that “statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.” Like citizenship and mass education, statistics as we know them today are first and foremost the instruments of statesmen, and a technique of statecraft.

This course examines the political uses of statistics (and, more generally, of numbers) by political authorities. It does so by going through case studies of how quantification – the operation of ‘coming up with numbers’ – is used to serve state interests and to call for political action.

In order to get everyone up to speed on basic statistics, the course begins with a 5-week workshop organised in similar fashion to another course that I teach, Statistical Reasoning and Quantitative Methods. The course then moves to the study of how quantified information feeds into politics, and ends on student presentations on the role of quantification in their own research.

Answers to frequently asked questions

Is this a statistics course? — Yes, but only partly. The first part of the course is basic computing, basic statistics, and empirical research design. The second part is a sociological analysis of quantitative methods in the context of political decision-making.

Do I need to know statistics to take the course? — No. No previous knowledge of either descriptive or inferential statistics is required for taking this course, although motivation to learn about it is crucial to the rest of the course.

How much statistics will I know by the end of the course? — Some, although not enough to independently conduct quantitative research. You will be pointed to the resources that you will need to learn more about quantitative methods.
Course Outline

Part 1 of the course follows parts of Statistical Reasoning and Quantitative Methods: in Weeks 1–5 of this course, we will cover as much as possible from Sections 1–9 of that other course, which is basically a crash course in statistics for social scientists.

Part 2 then switches to traditional lectures for two weeks, in order to convey the most essential components of a political sociology of quantification, using some of the selected readings listed in the last part of this syllabus.

Last, Part 3 closes with students presentations, i.e. very brief talks in which each student will individually present his or her current Masters dissertation project, with particular attention to the role of quantification and statistics in the project.

1. Introduction · Statistical Software

Part 1 – Statistics for Social Scientists

2. Datasets and Variables
3. Descriptive Statistics and Distributions
4. Associations and Correlations
5. Linear and Nonlinear Regression Models

Part 2 – Quantification for Social Scientists

6. Measurement as Representation
7. Visualization as Seduction

Part 3 – Assessment

8. Student Presentations / 1
9. Student Presentations / 2

Grading policy

10% Attendance — Just come to class.

20% Homework — Short readings that will be attached to the course emails.

70% Presentation — See the instructions on the last page of this syllabus.
Course Resources

The next pages list some journals, readings and websites that are useful to get a grasp of the topics covered in the course (quantification and statistics). All recommended material is written either in English or in French, and particularly recommended resources are highlighted in yellow.

Selected Journals

*Annals of Applied Statistics*
*Annals of Statistics*
*Annual Review of Statistics and its Application*
*Biometrika*
*Biostatistics*
*Econometrica*
*European Sociological Review*
*Journal of Business & Economic Statistics*
*Journal of Computational and Graphical Statistics*
*Journal of the Royal Statistical Society, Series A (Statistics in Society)*
*Journal of Statistical Software*
*Journal of the American Statistical Association*
*Radical Statistics*
*Review of Economics and Statistics*
*Scientific Data*
*Scientometrics*
*Significance*
*Social Indicators Research*
*Socio-Economic Review*
*Statistical Journal of the IAOS*
*Statistical Journal of the United Nations Economic Commission for Europe*
*Statistical Methods in Medical Research*
*Statistical Science*
*Statistics and Public Policy*
*Statistics in Medicine*
*Statistique et société · in French*
Selected Readings


Hirschman, Daniel Abramson. 2016. Inventing the Economy, Or: How We Learned to Stop Worrying and Love the GDP. PhD dissertation, University of Michigan.


---

¹ This article is about the Battle Deaths Dataset published by the Peace Research Institute Oslo (PRIO).
² See also the extended reviews of that book by Agnès Labrousse and Boris Samuel (both in French).
³ See also Michael Spagat’s blog for more case studies of (mis)counting battleground deaths.


---

4 The article by Geoffrey Traugh is also available in English per request to the author.


Selected Websites

**ACQUA – Approches Critiques de la QUAntification** · in French

**Additional Data Sources**, a list of social science datasets compiled for my, *Statistical Reasoning and Quantitative Methods* course

**ArXiv** papers from *cs.CY – Computers and Society* and *stat.AP – Statistics: Applications*

**DATACTIVE**, a research project by *Jonathan Gray* and others

**Gapminder**, created by *Hans Rosling*, famous for a *brilliant TED Talk* given in 2006

**LearnOpenData**, by Claire Foulquier-Gazagnes, from *Etalab*

**Open Data Institute**

**Open Knowledge International Blog** and **School of Data Blog**, by *Open Knowledge International*

**Philosophy of Data Series** and **Politics of Data Series**, from the Impact of Social Sciences Blog published by the London School of Economics and Political Science

**Radical Statistics Group**, the group that publishes the *Radical Statistics* journal

**Stats and Stories**, a podcast that interviews practitioners of statistics — *highly recommended*

**Statistiques en société** · in French

**US City Open Data Census**, an example of an open data census by *Open Knowledge International*

**Vital and Health Statistics Series**, by the U.S. *National Center for Health Statistics*
Instructions for student presentations

In what follows, references to readings not previously cited in the syllabus are highlighted in green and are listed at the end of this section. Please make sure to check at least some of them.

You will receive (even) more references as part of the general feedback that I will send by email after having listened to all student presentations.

In brief —

- Your presentation will focus on your research dissertation topic, which should include a research question, a data collection strategy, and an analytical perspective.

  Please refer to the ESPOL Guide to Academic Writing and to the ESPOL Writing Guide for Master Dissertations for general guidance, as well as to the contents and references of your research methods and epistemology courses.

  You might also want to turn to methods handbooks like della Porta and Keating (2008) and Hancké (2009) if you have limited experience with research design or need a refresher on how to conduct empirical social science research.

- You will have 7.5 minutes and a maximum of 6 slides to make a link between the course content and your research dissertation topic.

  Please read the rest of the instructions below to understand how to formulate that link, which might or might not involve conducting quantitative research yourself, and make sure to read Espeland and Stevens (2009) as soon as possible.

- You will have to upload your slides in PDF format, at least 48 hours before presenting. Late submissions will either be harshly penalized or will not be accepted at all.

  You will have to present in the session you are assigned to: absences will be graded 0/20, and there is no ‘catch-up’ session. The address of the Google Drive folder to which to upload your presentation will be sent to you by email, along with some final instructions.

- I will present Espeland and Stevens (2009) and some general ideas about the study of quantification in Week 6.

  If you have any questions regarding the presentations, please prepare them in advance in order to ask them in class during that session.

See the next pages for a longer explanation of what your presentations should aim at examining. Again, please make sure to check the references provided for additional guidance.
Research dissertations and the study of quantification

1. As a student, you have surely heard of the qualitative/quantitative divide, since it still largely structures methods courses, programs and textbooks. I want to briefly argue here that this divide is counter-productive, for at least two reasons:

   i. First, the divide does not explicitly refer back to its epistemological origin, which has to do with concept intension and extension, and therefore with comparability;

   ii. And second, the divide is insufficiently acknowledged as an ideal-typical one: qualitative and quantitative modes of scientific inquiry are, in practice, always mixed together to varying extents.

   The main consequence of (1) is a poor understanding of research methods that manifests itself through various forms of parochialism. The main consequence of (2) is the neglect of (methods of) qualitative inquiry by quantitatively-driven researchers, and conversely, the neglect of (methods of) quantitative inquiry by qualitatively-driven researchers.

2. The latter issue has a simple fix: qualitatively-oriented researchers need to pay attention to the role of numbers (a.k.a. measurements) and (statistical) models in their research topics.

   This is the research agenda offered in Espeland and Stevens (2009), who call it a ‘sociology of quantification,’ and who expand the research agenda followed by historians of statistics (such as Desrosières 1993 and Porter 1995) and by philosophers of science. In recent years, many social scientists have furthered this agenda.

   For your presentations, I am asking that you learn enough about this research agenda in order to assess how it might apply to your research dissertation topics.

   The readings section of this syllabus offer many references related to this agenda, and even more references can be found in its selection of academic journals and elsewhere in the scientific literature. Last, you are of course more than welcome to use your own research skills to find, access, read, use and cite other relevant academic work.

3. The exercise described above requires at least two things:

   i. First, that you have received some training in disciplines like epistemology and social science methodology, which are essential to grasp how research design articulates theory, method and data.

      In that regard, I trust that you have already read the relevant ESPOL guides cited at the beginning of this section, since they are meant to guide you through the steps that you have been following since the beginning of the academic year in order to start working on your research dissertations.

   ii. Second, I also trust that you have developed (usually through lots of practice) your own research skills, and that you know how to use tools like Google Scholar or, for bibliographic management and academic referencing, Zotero.
In this course, you have also given a try to using statistical software, in order to get familiar with an additional research skill: statistical computing. The (limited) goal of this course is only to provide you with an overview of that skill, not to teach you how to use statistical programming in your own research practice.

4. What the exercise above does not require is that you ‘turn’ your research topic into a quantitative analysis for the purpose of presenting it in class, or that you perform some kind of quantitative analysis in your presentation. On the contrary, the research agenda described in my previous points is, very explicitly, qualitative in nature.

Now, if you had already decided to study your research dissertation topic via quantitative methods like the ones that we surveyed in class, you are, of course, most welcome to frame your presentation accordingly, by combining some qualitative and quantitative points.

Be aware, however, that using quantitative methods in your research requires a lot more practice with statistical software than the limited overview offered in this course, and that mixing qualitative and quantitative methods – as does e.g. Lieberman 2009 – is a research method in itself, and therefore requires (even) more research training.

5. The expected benefits of this course are twofold:

i. The first segment of the course intends to show you what tools are available to perform professional-grade quantitative/statistical analysis, beyond the very basic spreadsheet editors – such as Google Sheets, LibreOffice Calc or Microsoft Excel – that you were already familiar with.

ii. The second segment of the course intends to suggest a possible way to frame your research dissertation topic by emphasizing the role of data, numbers and statistical analysis in the formation of (especially contemporary) expertise.

Again, it should be obvious that the two expected benefits above are complementary to each other, but that they can also be used in isolation to each other: one does not need to perform a quantitative analysis in order to produce a qualitative assessment of how such forms of analysis contribute to the ‘rule of experts’ (Mitchell 2002).

Most of the points above will be restated in Week 6 of the course, in which I will offer an illustrated overview of the research agenda offered in Espeland and Stevens (2009), featuring many examples of how to study things ranging from public policy performance indicators (Ogien 2013) to the measurement of human rights violations (Merry 2016) and crime (Martin 2015) to state-level measurements of e.g. democracy and state ‘fragility’ or ‘failure’ (Figueroa Helland and Borg 2014).

All studies mentioned above are related, to various extents, to the study of expertise, on which see e.g. Stampnitzky 2013 on terrorism and Hagan, Schoenfeld and Palloni 2006 on human rights, war crimes and humanitarian interventions, as well as Eyal and Buchholz 2010 and Carr 2010 for more general reviews, and Jasanoff 2004 for a useful theoretical framework.
Additional references


Thanks

This syllabus contains some ideas and references that were discussed with (or provided by) Philippe Bonditti, Brendan Coolsaet, Janis Grzybowski, and other colleagues at ESPOL in Lille.

I also owe some of my knowledge of the scientific literature on quantification to Émilien Ruiz, with whom we organised a research seminar on that topic while I was working on this syllabus.

Last, I also owe some of my knowledge of open data to Joël Gombin and Samuel Goëta, with whom we teach an “Open Data for Urban Research” course at Sciences Po in Paris.