STATISTICAL REASONING AND QUANTITATIVE METHODS

MSc in International Affairs
MSc Governing the Large Metropolis

François Briatte & Ivaylo Petev
francois.briatte@sciences-po.org & ivaylo.petev@sciences-po.org

Fall and Spring 2010–13
13, rue de l'Université, 75007 Paris
Rooms J306–7

This course is about the **core notions** of quantitative research for the social sciences, based on three fundamental blocks of knowledge: essential statistical concepts, survey data, and various forms of regression analysis.

By design, this course will approach quantitative analysis through methods and examples taken from various branches of the social sciences, with some specific applications to international relations. We will focus on **research design**, as to make sure that we ask valid questions, based on sound hypotheses as well as reliable data, and draw correct inferences. Throughout the course, we will introduce and explain some essential **statistical operations** that can be used to that end. Finally, we will introduce **statistical software** and work through the procedures to produce statistical tests and visualizations of quantitative data.

The emphasis of the course is set on conceptual understanding and statistical reasoning, and each session will apply statistical procedures to real data. Handbook chapters will be used to cover the statistical side of the course, while class sessions will focus on practical experience.

No previous knowledge in any of these topics is required for taking the course, but some computer and Internet skills as well as a genuine interest in understanding why and how we use quantitative information to understand society will prove useful.

COURSE REQUIREMENTS

Students are invited to be regular participants in the course and to complete required readings prior to class meetings. Course sessions start with a theoretical and practical introduction and end with a lab session. In order to learn Stata during the semester, students are required to train as much as needed with the software and to find additional help online if required.

Students are assessed on the basis of two **draft papers** and one **final paper**, for which they should provide replication material. The drafts and final paper focus on a single dataset and research question that students examine in pairs throughout the semester. Expectations about coursework will be outlined in the first class and further detailed at several points.

Feel free to ask for additional guidance on what to read and how to structure your papers, yet *do not wait for the last minute to do so, and read the course documentation first.* The grading policy for the course is 25 points for each draft and 50 points for the final paper. Attendance to all sessions, which are all computer-based, is crucial to the course. Students are also asked to elect a student representative and to provide regular feedback on the course.

COURSE SUMMARY

Data

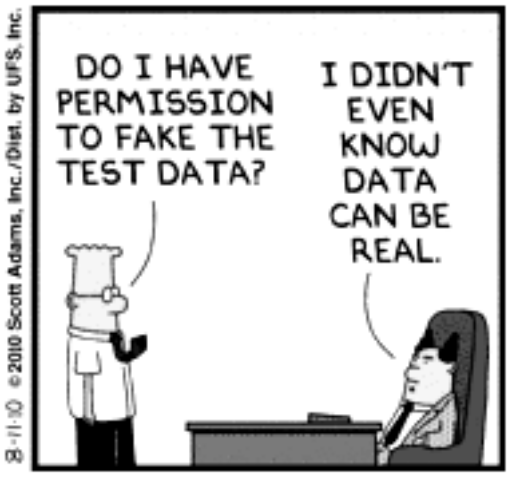
1. Introduction
2. Datasets
3. Variables
4. Distributions

Hypotheses

5. Estimation
6. Comparison
7. Correlation
8. Regression

Models

9. Interactions
10. Linear regression
11. Logistic regression
12. Specification



HANDBOOKS

Briatte, F. and Petev, I. 2012. *Stata Guide. A Student Guide to Statistics With Stata*.

Feinstein, C. H. and Thomas, M. 2002. *Making History Count*. Cambridge University Press.

Urdan, T. 2010. *Statistics in Plain English.* 3rd ed. Routledge.

> ***Reading guide**: The 'Stata Guide' is a draft handbook that covers (most of) the course requirements; Feinstein and Thomas is an accessible introduction to quantitative social science; and Urdan further clarifies the statistical notions covered in the other readings. Read responsibly from all sources.*

ADDITIONAL READINGS

Booth, W. *et al.* 2003. *The Craft of Research*. 2nd ed. University of Chicago Press.

Mitchell, M. 2012. *Interpreting and Visualizing Regression Models Using Stata*. Stata Press.

Tufte, E. 2001. *The Visual Display of Quantitative Information.* Graphics Press.

> ***Reading guide**: Booth* et al. *is an introduction to research practice and research writing, and will be particularly helpful to students with limited training in that area; Mitchell's book is an excellent guide to applied regression in Stata; and Tufte is a treaty on visual design and data visualization.*

EXAMPLE PAPERS

To complete your coursework, you will need to form a group and write an empirical research paper based on your work throughout the semester. If this is your first research paper based on empirical data, see Lynn White, "Writes of Passage: Writing an Empirical Journal Article," *Journal of Marriage and Family* 67 (2005): 791–8, for essential instructions.

Examples of empirical papers using survey or country-level data will be provided in class. You might want to find more examples on Google Scholar, using the exact name of the dataset that you want to use as keywords. You do not need to understand the full statistical methodology of the papers to study how they are structured and written.

LINKS

Course website:
http://f.briatte.org/teaching/quanti/

Course blog:
http://srqm.tumblr.com/

LSE Methodology video tutorials:
https://www.youtube.com/user/MethodologyLSE/videos?query=stata

StataCorp:
http://stata.com/

COURSE OUTLINE

The course is made of three teaching segments: a general section on descriptive statistics and data preparation (sessions 1–5), a focused section on bivariate association tests (sessions 6–8), and a final section on linear and logistic regression models (sections 9–12). Each segment of the course corresponds to a section of the final paper that is due at the end of the semester. Read handbook chapters *before* class and the Stata Guide *after* class.

SESSION 1 INTRODUCTION

Readings

- **Feinstein & Thomas** ch. 1
- **Stata Guide** s. 1–4
- **Urdan** ch. 1

SESSION 2 DATASETS

Readings

- **Data documentation** *see data folder*
- **Stata Guide** s. 5–6
- **Stata tutorials** *see course website*

SESSION 3 VARIABLES

Readings

- **Feinstein & Thomas** ch. 2.1–2.4
- **Stata Guide** s. 7–8
- **Urdan** ch. 2–3

SESSION 4 DISTRIBUTIONS

Readings

- **Feinstein & Thomas** ch. 2.5–2.6
- **Stata Guide** s. 9
- **Urdan** ch. 4–5

SESSION 5 ESTIMATION **FIRST DRAFT**

Readings

- **Feinstein & Thomas** ch. 5
- **Stata Guide** s. 13–14
- **Urdan** ch. 6–7

SESSION 6 COMPARISON

Readings

- **Feinstein & Thomas** ch. 6–7
- **Stata Guide** s. 10
- **Urdan** ch. 9 and 14

SESSION 7 CORRELATION

Readings

- **Feinstein & Thomas** ch. 3
- **Stata Guide** s. 10
- **Urdan** ch. 8

SESSION 8 REGRESSION

Readings

- **Feinstein & Thomas** ch. 4
- **Stata Guide** s. 11
- **Urdan** ch. 13

SESSION 9 INTERACTIONS **REVISED DRAFT**

Readings

- **Feinstein & Thomas** ch. 8 (and optionally ch. 9)
- **Stata Guide** s. 13–15

SESSION 10 LINEAR REGRESSION

Readings

- **Feinstein & Thomas** ch. 10.1 (and optionally ch. 11)
- **Stata Guide** s. 11

SESSION 11 LOGISTIC REGRESSION

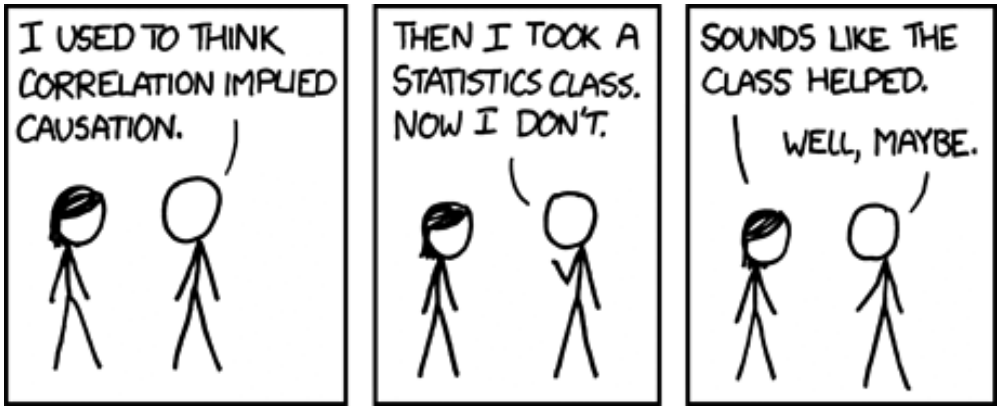
Readings

- **Feinstein & Thomas** ch. 13.1–13.3 (and optionally ch. 12.1–12.3)
- **Stata Guide** s. 11

SESSION 12 SPECIFICATION **FINAL PAPER**

Readings

- **Stata Guide** s. 16



COURTESY OF XKCD.COM