

Introduction to Quantitative Methods

Instructor

[Alexandru Moise](#), Part-time Assistant Professor, Max Weber Fellow
Office hours: Mondays 14:00-16:00. Request by [email](#)

Course Assistant

[Inês Duarte](#), PhD researcher (Stata lab)

Contact

[Adele Battistini](#)

Lecture: Mondays, 11.00-13.00, Sala del Capitolo (on 2 November and 22 November: Seminar Room 3)

Lab: Tuesdays, 11.00-13.00, Seminar Room 3 (on 4 November: Sala del Capitolo)

Note: This course fulfills your requirement to take one quantitative course during your first two years. This course can be substituted with more advanced ones: Intermediate Quantitative Methods from the SPS Department or "Introduction to Econometrics and Regression", offered by the Economics Department.

This course is designed for those with little to no prior experience in using quantitative methods. By the end of the course, you will become informed consumers and producers of quantitative analysis. We will place a lot of emphasis on understanding published research, as well as on the fundamentals of probability theory and inference. By the end of the course, you will have a solid foundational understanding which will also allow you to take more advanced courses at the EUI or elsewhere. As you will see, all advanced methods build on the simple principles of probability and inference and are usually a form of extension of OLS (ordinary least squares) regression, which we will cover. Understanding these concepts and models will not only allow you to understand published quantitative research in social science, but in other fields as well. We will discuss in depth the limitations of these methods, including problems with p-values (and p-hacking!), publication bias, measurement bias, and assumptions in models. We will learn how to spot these pitfalls in other research, and how to avoid them in order to produce meaningful analysis.

The course starts with an overview of types of data, probability theory and statistical inference. We will explore hypothesis testing in the simple context of a t-test before moving on to linear regression. We will then explore the fundamental problem of inference and learn how multiple regression attempts to mimic a randomized control trial. We will then learn about how we can evaluate the quality and robustness of our models by looking at fit statistics. In the final part of the course, we will explore the use of categorical and binary predictor variables as well as data transformations.

The course consists of lectures and labs. Each lecture will have assigned readings, which you will need to do before the start of the class. You will also have weekly assignments that will test your theoretical understanding of the topics we cover, as well as help you practice your coding skills in Stata or your programming language of choice! The labs will be run in Stata. You can choose to work in another software if you prefer, but direct support will be offered for Stata. Your assignments need to be uploaded on the course site before the lab.

This course is an introduction to quantitative methods in the social sciences. It is mandatory and can be taken either during the first or the second year of the SPS doctoral programme. The course will be taught in person on campus, with the possibility of switching to Zoom if needed.

Coursebook

Agresti, Alan (2018). Statistical Methods for the Social Sciences, Fifth edition. Pearson.

This is the main book used in the course. The library has copies of it, but it is not a bad idea to purchase it. **Note:** Beware of websites that promise and fully deliver free PDFs of all kinds of books and articles such as:

<https://libgen.rs/>
<https://sci-hub.st/>

These websites might seem perfectly helpful in that they deliver exactly what they promise, but they are also free! In any case, you've been warned!

Other books that cover the issues and can be helpful include:

General statistics

- Gravetter, F. J., & Wallnau, L. B. (2013). Essentials of Statistics for the Behavioral Sciences (8 edition). Wadsworth Publishing.
- Huff, D. (1954). How to lie with statistics. Norton.

Popular books for data science

- Wheelan, C. (2013). Naked statistics: Stripping the dread from the data. WW Norton & Company.
- Taleb, N. N. (2007). The black swan: The impact of the highly improbable (Vol. 2). Random house.
- Silver, N. (2012). The signal and the noise: Why so many predictions fail—but some don't. Penguin.

Books for the R software:

- Adler, J. (2010). R in a nutshell (1st ed). O'Reilly.
- Field, A., Miles, J., & Field, Z. (2012). Discovering Statistics Using R (1 edition). SAGE Publications Ltd.

For Bayesian statistics: Lambert, B. (2018). A student's guide to Bayesian statistics. SAGE.

Soft introduction to causal inference: Pearl, J., & Mackenzie, D. (2018). The book of why:

The new science of cause and effect (First edition). Basic Books.

Schedule:

Week 1. Overview of the purpose of quantitative research and of social science data and measurement.

Reading: **Agresti, Ch. 1**

Popular reading: **Engber 2012**. Why do people love to say that correlation does not imply causation? Slate Magazine. <https://slate.com/technology/2012/10/correlation-does-not-imply-causation-how-the-internet-fell-in-love-with-a-stats-class-cliche.html>

Lecture: Monday, 4 October 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 5 October 2021, 11:00-13:00, Seminar Room 3

Week 2. Sampling and Descriptive Statistics

Textbook reading: **Agresti, Chs. 2-3**

Applied reading: **Geddes, B. (1990)**. How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics. *Political Analysis*, 2, 131–150.

Lecture: Monday, 11 October 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 12 October 2021, 11:00-13:00, Seminar Room 3

Week 3. Probability Theory and Inference

Textbook reading: **Agresti, Chs. 4-5**

Applied reading: **Mann, T. E., & Wolfinger, R. E. (1980)**. Candidates and Parties in Congressional Elections. *American Political Science Review*, 74(3), 617–632.
<https://doi.org/10.2307/1958145>

Lecture: Monday, 18 October 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 19 October 2021, 11:00-13:00, 11:00-13:00, Seminar Room 3

Week 4. Hypothesis testing

Textbook reading: **Agresti, Chs. 6-7**

Applied reading: **Matland, R. E. (1994)**. Putting Scandinavian Equality to the Test: An Experimental Evaluation of Applied reading: Gender Stereotyping of Political Candidates in a Sample of Norwegian Voters. *British Journal of Political Science*, 24(2), 273–292.

Lecture: Monday, 25 October 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 26 October 2021, 11:00-13:00, 11:00-13:00, Seminar Room 3

Week 5. Univariate regression and ordinary least squares estimation

Textbook reading: **Agresti, Ch. 9.**

Applied reading: **Segal, J. A., & Cover, A. D. (1989)**. Ideological Values and the Votes of U.S. Supreme Court Justices. *American Political Science Review*, 83(2), 557–565.

Lecture: **Tuesday, 2 November 2021, 11:00-13:00, Seminar Room 3**

Lab STATA: **Thursday, 4 November 2021, 13:00-15:00, Sala del Capitolo**

Week 6. Multiple regression: the fundamental problem of causal inference

Textbook reading: **Agresti, Ch. 10, 11.1-11.3**

Applied reading: **Cameron, D. R. (1978)**. The Expansion of the Public Economy: A Comparative Analysis. *American Political Science Review*, 72(4), 1243–1261.

<https://doi.org/10.2307/1954537>

Lecture: Monday, 8 November 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 9 November 2021, 11:00-13:00, Seminar Room 3

Week 7. Multiple regression continued: spurious effects, suppressor effects, and control variables

Textbook reading: **Agresti Chs. 11.5, 11.7**

Applied reading: **Fearon, J. D., & Laitin, D. D. (2003)**. Ethnicity, Insurgency, and Civil War. *American Political Science Review*, 97(01), 75–90.

<https://doi.org/10.1017/S0003055403000534>

Lecture: Monday, 15 November 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 16 November 2021, 11:00-13:00, Seminar Room 3

Week 8. Categorical predictors and model building

Textbook reading: **Agresti, Ch. 11.4, 12.1, 13.1-13.2, 14.1**

BBC podcast on the replication crisis: <https://www.bbc.co.uk/sounds/play/m00013p9>

Lecture: **Monday, 22 November 2021, 11:00-13:00, Seminar Room 3**

Lab STATA: Tuesday, 23 November 2021, 11:00-13:00, Seminar Room 3

Week 9. Lying to yourself and others with statistics: p-hacking, replication crisis

Popular reading 1: **Colquhoun 2016**. The problem with p-values. <https://aeon.co/essays/its-time-for-science-to-abandon-the-term-statistically-significant>

Popular reading 2: **Belluz 2016**. An unhealthy obsession with p-values is ruining science.

Vox. Available at <https://www.vox.com/2016/3/15/11225162/p-value-simple-definition-hacking>

Lecture: Monday, 29 November 2021, 11:00-13:00, Sala del Capitolo

Lab STATA: Tuesday, 30 November 2021, 11:00-13:00, Seminar Room 3

Week 10. Presenting regression results: interpretations, tables, and graphs

Theoretical reading: **Elkink 2013**. Some notes on presenting and interpreting regressions

http://www.joselkink.net/wp-content/uploads/2013/01/POL50050_Spring_2013_note_regression_presentation_and_interpretation.pdf

Applied reading: **Miller 2020**. What Do We Know About British Attitudes Toward Immigration? A Pedagogical Exercise of Sample Inference and Regression

<http://svmiller.com/blog/2020/03/what-explains-british-attitudes-toward-immigration-a-pedagogical-example/>

Lecture: Monday, 6 December 2021, 9:00-11:00, Sala del Capitolo
Lab (Stata): Tuesday, 7 December 2021, 11:00-13:00, Seminar Room 3

Week 11. Recap

Lecture: Monday, 13 December 2021, 11:00-13:00, Sala del Capitolo
Lab STATA: Tuesday, 14 December 2021, 11:00-13:00, Seminar Room 3