Introduction to Stata

Universitat de Barcelona - MSc in Economics

Academic Year 2025/2026

1 Information

- Instructor: Salvatore Viola
- Email: sviola@ub.edu
- Course materials: https://github.com/salvatoreviola
- Office: 4320 (Diagonal 690, Tower 4)
- Meetings by appointment

2 Course Objectives

The objective of this course is to give students an overview of the Stata statistical software. This course will introduce students to a comprehensive set of basic commands and actions in Stata which are necessary for descriptive and econometric analysis. Students will learn the fundamentals of manipulating and visually representing data in addition to some best practices for saving and presenting tables and figures. Particular attention will be given to understanding the Stata interface as well as proper file and project management. By the end of the course, students should be equipped with the tools to conduct the basics of an economic research project. All of the course materials will be made available with each topic on my GitHub.

3 Evaluation

- **Problem Sets** (50 %): After the first and second classes, students will be given a problem set to solve on their own at home during the week in between. The problem sets cover topics taught in class and aim to encourage students "dig deeper" into the functionality of relevant commands and get use to problem solving in Stata.
- **Final Exam** (50 %): The final exam will be held in class on the **November 7th at 12:00h**. The questions will reflect the core topics covered in class, including performing the fundamental tasks of any research project in Stata, understanding data structure and interpreting basic econometric results. 1 hour of this session will be allotted for students to complete the exam. The use of any LLM (ChatGPT, Claude, etc.) is prohibited during the exam.

4 Organization

4.1 Class 1 (October 17th)

In the first class we will familiarize ourselves with Stata, importing and preparing data, and exporting results.

- Understand the interface of Stata
- Import files in different formats

- Rename, replace, drop variables, change labels, using conditionalities
- Descriptives and basic estimations
- Saving and exporting graphs and tables
- Graphs and tables for descriptive analysis
- Post-editing (for aesthetic purposes)
- In-class exercises

4.2 Class 2 (October 24th)

The second class will be devoted to some advanced data manipulation techniques in Stata

- Predictions, residuals and plotting them
- Merging datasets
- Different types of data (cross-section, time series, panel data)
- Collapsing and reshaping data
- In-class exercises

4.3 Class 3 (October 31st)

Class 3 will cover some more advanced functions of Stata.

- Advanced Stata topics
- Project structuring and management
- Real code examples
- In-class exercises

5 Additional Resources

The first three are websites that can often provide solutions to problems you may be having in Stata. The 4th link is to the online version of a book by Scott Cunningham which covers a very broad range of economic methods and includes code sample in Stata, R and Python. The CRAI UB online library may have access to the newest versions of the listed textbooks, while older versions may be available online for free. Stata is a proprietary statistical software, a characteristic which has proven to limit the helpfulness of LLMs in providing solutions to Stata-related problems. Therefore, it is generally best practice to look elsewhere for help when working in Stata, especially when using more advanced techniques.

- (Statalist)
- (Stack Exchange)
- (Stack Overflow)
- (Causal Inference: The Mixtape)
- *An Introduction to Modern Econometrics Using Stata* by Christopher F. Baum
- *Introductory Econometrics: A Modern Approach* by Jeffrey M. Wooldridge
- *Mostly Harmless Econometrics: An Empiricist's Companion* by Joshua D. Angrist and Jörn-Steffen Pischke