STATA Training Camp Syllabus

Juncheng Jiang

2024 Nov-Dec

Course Objectives

This course is designed to lay the foundation for students new to STATA, helping them acquire basic data processing and statistical analysis skills to establish a solid groundwork for subsequent economic research. Through four sessions, students will learn to import data, clean data, perform basic statistical analysis and graphical visualization, and output empirical analysis charts and tables.

Course Schedule

Lecture 1: STATA Basics and Data Operations

Objective:

- Install STATA
- Understand the basic operation interface and command syntax of STATA software
- Learn how to import data in different formats and perform data cleaning

Course Content:

1. **Basic Software Operations**
    - Software interface (menu bar, command window, variables window, results window)
    - How to set the working directory (`cd` command)
    - Creating log files and recording results (`log` command)
    - Basic help documentation and commands (`help` and `search`)
2. **Data Import and Viewing**
    - Importing Excel, CSV, TXT files (`import excel`, `import delimited`)
    - Importing example datasets (`sysuse` command)
    - Viewing and browsing data (`browse`, `list` commands)
3. **Data Cleaning**
    - Generating, replacing, and deleting variables (`gen`, `replace`, `drop`)
    - Conditional statements (usage of `if`, `==`, `>`, `<`)
    - Variable labels and value labels (`label variable` and `label define`)
    - Saving data (`save`, `use`)

Lecture 2: Descriptive Statistics and Visualization

Objective:

- Master basic descriptive statistical analysis and plotting methods for data

Course Content:

1. **Descriptive Statistics** (`summarize`)
2. **Group Descriptive Statistics** (`tabstat`)
3. **Plotting Charts**
    - Scatter plot (`scatter`): Describing correlation between two variables
    - Line plot (`line`): Describing the time trend of a single variable
    - Histogram (`histogram`): Describing the distribution of a continuous single variable
    - Bar chart (`graph bar`): Expressing the distribution of discrete variables, group statistics
4. **Chart Customization**: Adjusting titles, labels, colors, etc.

Lecture 3: Regression Analysis and Output

Objective:

- Learn how to perform basic regression analysis and export results

Course Content:

- **Econometrics Basics:**
    - Meaning of estimated coefficients in linear regression models
    - The role of standard errors
- **STATA Application Methods:**
    - Linear regression (`reg`, `reghdfe`)
    - Interpreting regression results (coefficients, standard errors, etc.)
    - Exporting regression results to table formats (such as Word or Excel, using commands like `esttab`, `sum2docx`, `outreg2`)

Lecture 4: Paper Replication

Objective:

- Replicate the main results of a paper (DiD model)

Course Content:

- Generating variables, cleaning data, and descriptive statistics
- Replicating econometric results
- Replicating graphs (including dynamic effect graphs for DiD)
- Replicating output/results export

Expected Outcomes

Upon completion of this course, students will be able to:

1. Proficiently use STATA for basic data import, cleaning, and analysis
2. Master fundamental statistical analysis skills and perform simple regression analysis
3. Effectively present analysis results as charts or export them to documents