GESIS Summer School in Survey Methodology 2023

Syllabus for course:
“Introduction to Stata for Data Management and Analysis”

Lecturers: Irina Bauer, M.A.  Annika Stein, M.A.
E-mail: irina.bauer@gesis.org  annika.stein@gesis.org

Date: 02 – 04 August 2023
Time: 10:00-12:00 | 13:00-15:00 CEST
Venue: Online via Zoom

About the Lecturers:

Irina Bauer is a doctoral researcher working for the panel study FReDA (Family Research and Demographic Analysis) at GESIS Mannheim. Irina holds a bachelor’s and master’s degree in Social Sciences from the University of Koblenz-Landau. Her research interests include (visual) questionnaire design and data quality in web surveys.

Annika Stein is a doctoral researcher working for the panel study FReDA (Family Research and Demographic Analysis) at GESIS Mannheim. Annika holds a bachelor’s degree in Politics and Public Administration from the University of Konstanz and a master’s degree in Empirical Research in Democracy from the Johannes Gutenberg University Mainz. Her research interests locate in family sociology with a focus on the consequences of economic deprivation for family functioning.

Course Description:

This online short course will give a thorough introduction to the statistical software Stata. It is tailored to the needs of academics and other research practitioners who are new to Stata or who wish to refresh their skills. The course will not cover basic statistical methods and their underlying mathematics but will focus on their practical application using family demographic survey data in Stata.

In the first part of the course, we will cover the program's interface and introduce its syntax structure and basic rules for writing clean and reproducible Stata code. Subsequently, we aim at enhancing the participants' skills in hands-on data management, common data analyses, and the visualization of results. Depending on the participants’ prior knowledge, we might provide further insight into the automatization of data wrangling and analysis procedures and the export of publication-ready results. Additionally, we will review available help and support features (online and offline) to equip participants with the necessary knowledge to further develop their skills and solve occurring problems. This course will be interactive and case-based in nature.

Keywords:
data management, descriptive statistics, data analysis, data visualization, reproducibility

Course Prerequisites:

- Familiarity with quantitative (survey) data
- Basic knowledge of uni- and bivariate statistics (i.e., descriptive statistics, basics of regression analysis)
Target Group:
Participants will find the course useful if:

- they are new to statistical computing (with Stata).
- they are familiar with other statistical software but want to get to know Stata.

Course and Learning Objectives:
By the end of the course participants will:

- be familiar with Stata’s interface and facilities.
- understand how to integrate Stata into their research process to create reproducible and publication-ready results.
- know how to solve common data management problems in Stata and how to document all modifications of the data.
- be able to perform typical descriptive and inferential statistical procedures and use graphs to communicate their results effectively.
- know how to proceed from here and get additional support if needed.

Organizational Structure of the Course:
The course is offered as an expanded short course consisting of 12 hours of interactive group instruction, spread over three days. The course will alternate between short lectures and hands-on exercises using Stata. Additional assignments to get a handle on Stata will be discussed in the group the next day. The two lecturers will be available to the students and support work on assignments. All necessary materials will be provided.

Software and Hardware Requirements:
Participants need a laptop/desktop computer that enables them to access the internet and smoothly work with Stata. Participants will be provided with access to Stata licenses by GESIS but must install the software prior to the course on their own devices.

Day-to-day Schedule and Literature:

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Introduction to the Stata interface and syntax language</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Lunch Break</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Getting to know your data</strong></td>
</tr>
<tr>
<td>Suggested reading (optional):</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td><strong>Data management and data manipulation</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Lunch Break</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Bi- and multivariate statistics I</strong></td>
</tr>
</tbody>
</table>
Suggested reading (optional):


<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.00 - 12.00</td>
<td>Bi- and multivariate statistics II</td>
</tr>
<tr>
<td>12.00 - 13.00</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>13.00 - 15.00</td>
<td>Presentation and visualization of results</td>
</tr>
</tbody>
</table>

Suggested reading (optional):


Preparatory Reading:

- StataCorp YouTube channel provides some excellent preparatory material. Participants are especially encouraged to browse the “Tour of the Stata 17 interface”, “Quick help in Stata,” and any other of the available material according to personal interest: https://www.youtube.com/watch?v=FQ1MBQw_MT1
- Ludwig-Mayerhofer provides intuitive hands on material on using Stata for data analysis with a great search function in the “Internet Guide to Stata” https://wlm.userweb.mwn.de/Stata/

Additional Recommended Literature:

- Kohler, Ulrich, and Frauke Kreuter. 2012. Data Analysis Using Stata. 3rd Ed. College Station, TX: Stata Press.