Training course on Advanced Statistical Software for Social Protection Research (Stata, R)

Professional Development Training Course Outline

Category	Social Protection	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

Mastering advanced statistical software is no longer a luxury but a necessity for robust and credible social protection research. In an increasingly data-driven world, researchers and practitioners need to efficiently manage, analyze, and visualize complex datasets to inform policy and program design. This intensive 10-day training course focuses on equipping participants with the practical skills to leverage two of the most powerful and widely used statistical software packages in social sciences and development: Stata and R. Participants will gain hands-on experience in data manipulation, advanced econometric modeling, causal inference techniques, and high-quality data visualization, ensuring they can independently conduct sophisticated analyses and contribute effectively to evidence-based decision-making in social protection. Training Course on Advanced Statistical Software for Social Protection Research (Stata, R) is meticulously designed to equip with the advanced theoretical insights and intensive practical tools necessary to excel in Advanced Statistical Software for Social Protection Research. We will delve into the foundational concepts of data management and programming within both Stata and R, master the intricacies of various statistical models, and explore cutting-edge approaches to causal inference, panel data analysis, and reproducible research. A significant focus will be placed on hands-on application, analyzing real-world complex social protection datasets, and developing tailored scripts and workflows. By integrating industry best practices, analyzing complex case studies, and engaging in intensive practical exercises, attendees will develop the strategic acumen to confidently lead and implement high-quality data analysis, fostering unparalleled analytical rigor, efficiency, and evidence generation for social protection outcomes.

Programme Curriculum

Training Course on Advanced Statistical Software for Social Protection Research (Stata, R)

Introduction

Mastering advanced statistical software is no longer a luxury but a necessity for robust and credible social protection research. In an increasingly data-driven world, researchers and practitioners need to efficiently manage, analyze, and visualize complex datasets to inform policy and program design. This intensive 10-day training course focuses on equipping participants with the practical skills to leverage two of the most powerful and widely used statistical software packages in social sciences and development: **Stata** and **R**. Participants will gain hands-on experience in data manipulation, advanced econometric modeling, causal inference techniques, and high-quality data visualization, ensuring they can independently conduct sophisticated analyses and contribute effectively to evidence-based decision-making in social protection.

Training Course on Advanced Statistical Software for Social Protection Research (Stata, R) is meticulously designed to equip with the advanced theoretical insights and intensive practical tools necessary to excel in **Advanced Statistical Software for Social Protection Research**. We will delve into the foundational concepts of data management and programming within both Stata and R, master the intricacies of various statistical models, and explore cutting-edge approaches to causal inference, panel data analysis, and reproducible research. A significant focus will be placed on hands-on application, analyzing real-world complex social protection datasets, and developing tailored scripts and workflows. By integrating industry best practices, analyzing complex case studies, and engaging in intensive practical exercises, attendees will develop the strategic acumen to confidently lead and implement high-quality data analysis, fostering unparalleled **analytical rigor, efficiency, and evidence generation** for social protection outcomes.

Course Objectives

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

1. Analyze the **strengths and weaknesses of Stata and R** for social protection research.
2. Comprehend **advanced data management and cleaning techniques** in both software.
3. Master **linear and non-linear regression models** using Stata and R.
4. Develop expertise in implementing **panel data models** (Fixed Effects, Random Effects).
5. Formulate strategies for conducting **causal inference analyses** (DiD, PSM, IV) in both environments.
6. Understand and apply methods for **handling complex survey data** and sampling weights.
7. Implement robust approaches to **data visualization and customized graphing**.
8. Explore key strategies for **automating tasks and writing reproducible code** in both Stata and R.
9. Apply methodologies for **integrating data from various sources** and formats.
10. Understand the principles of **efficient coding practices** for large datasets.
11. Develop preliminary skills in **troubleshooting common statistical software issues**.
12. Conduct a **complete data analysis workflow** for a social protection research question using both Stata and R.
13. Examine **global best practices and collaborative tools** for statistical research.

Target Audience

This course is essential for professionals seeking to advance their statistical software skills for social protection:

1. **Researchers & Academics:** Specializing in quantitative social science and development.
2. **M&E Specialists & Data Analysts:** Requiring advanced tools for impact evaluation.
3. **Economists & Statisticians:** Working with large-scale social protection datasets.
4. **Social Protection Program Managers:** Overseeing data-driven decision-making.
5. **Government Officials:** From statistical offices and policy analysis units.
6. **Development Practitioners:** From NGOs and international organizations.
7. **Consultants:** Providing data analysis and research services.

8. **Graduate Students:** Pursuing research in social sciences.

Course Duration: 10 Days

Course Modules

Module 1: Introduction to Stata and R Ecosystems for Social Protection

- Overview of Stata: interface, command syntax, do-files, help system.
- Overview of R: RStudio interface, packages, scripts, CRAN.
- Compare and contrast Stata and R strengths for social protection research.
- Install and set up both software environments.
- Understand the philosophy behind each software for data analysis.

Module 2: Advanced Data Management and Cleaning in Stata

- Master efficient data import/export from various formats (Excel, CSV, databases).
- Learn advanced data manipulation: reshape, merge, append, collapse.
- Discuss strategies for identifying and handling missing data.
- Understand string manipulation, date variables, and egen commands.
- Automate data cleaning workflows using do-files.

Module 3: Advanced Data Management and Cleaning in R

- Master efficient data import/export using readr, haven, rio packages.
- Learn advanced data manipulation using dplyr, tidyr, data.table.
- Discuss strategies for identifying and handling missing data with dplyr and na.omit.
- Understand string manipulation (stringr), dates (lubridate), and purrr for iteration.
- Automate data cleaning workflows using R scripts and RMarkdown.

Module 4: Linear and Non-Linear Regression in Stata

- Master Ordinary Least Squares (OLS) regression: robust standard errors, clustering.
- Implement logistic and probit regression for binary outcomes.
- Understand multinomial and ordinal logit/probit models.
- Explore count models (Poisson, Negative Binomial) for event data.
- Interpret marginal effects and interaction terms in non-linear models.

Module 5: Linear and Non-Linear Regression in R

- Master lm() for OLS regression: robust standard errors (sandwich), clustering (lfe).
- Implement glm() for logistic and probit regression.
- Understand multinomial (nnet::multinom) and ordinal (MASS::polr) models.
- Explore count models (glm with family=poisson/quasipoisson, MASS::glm.nb).
- Interpret marginal effects (margins package) and interaction terms.

Module 6: Panel Data Analysis in Stata

- Understand panel data structures and xtset command.
- Implement Fixed Effects (FE) models (xtreg, fe).
- Explore Random Effects (RE) models (xtreg, re) and the Hausman test.
- Discuss dynamic panel models (e.g., Arellano-Bond GMM basics using xtabond).
- Analyze unobserved heterogeneity and within-cluster correlation.

Module 7: Panel Data Analysis in R

- Understand panel data structures using plm package.
- Implement Fixed Effects (FE) models (plm::plm with model="within").
- Explore Random Effects (RE) models (plm::plm with model="random") and phtest.
- Discuss advanced panel methods using packages like fixest.
- Compare and contrast FE/RE estimation approaches in R.

Module 8: Causal Inference Techniques in Stata

- Apply Difference-in-Differences (DiD) using regression (reg, xtreg).
- Implement Propensity Score Matching (PSM) using psmatch2.
- Understand Instrumental Variables (IV) and Two-Stage Least Squares (2SLS) using ivregress.
- Explore Regression Discontinuity Design (RDD) using rdrobust.
- Discuss sensitivity analysis and robustness checks for causal estimates.

Module 9: Causal Inference Techniques in R

- Apply Difference-in-Differences (DiD) using fixest or did package.
- Implement Propensity Score Matching (PSM) using MatchIt or cem packages.
- Understand Instrumental Variables (IV) and 2SLS using ivreg or AER package.
- Explore Regression Discontinuity Design (RDD) using rdrobust package.
- Practice implementing different causal inference strategies with real data.

Module 10: Complex Survey Data Analysis and Power Analysis

- Understand survey design features: stratification, clustering, sampling weights.
- Apply svy commands in Stata for correct standard errors and estimates.
- Utilize survey package in R for complex survey analysis.
- Learn to calculate appropriate sample sizes and conduct power analysis in both software.
- Discuss the implications of complex survey design on statistical inference.

Module 11: Advanced Data Visualization and Reporting

- Master creating publication-quality graphs in Stata (graph editor, schemes).
- Utilize ggplot2 in R for highly customized and aesthetic plots.
- Discuss principles of effective data visualization for social protection.
- Learn to integrate results and graphs into dynamic reports (Stata's putexcel/putdocx, RMarkdown).
- Explore interactive visualizations with R packages (e.g., plotly, leaflet).

Module 12: Reproducible Research and Collaborative Workflows

- Understand the importance of reproducible research practices.
- Learn version control basics with Git/GitHub for Stata and R projects.
- Discuss best practices for organizing research projects and scripts.
- Explore tools for collaboration and sharing code (e.g., Jupyter Notebooks with R/Stata kernels, cloud environments).
- Debugging and troubleshooting strategies in both software.

Training Methodology

- **Interactive Workshops:** Facilitated discussions, group exercises, and problem-solving activities.
- **Case Studies:** Real-world examples to illustrate successful community-based surveillance practices.
- **Role-Playing and Simulations:** Practice engaging communities in surveillance activities.
- **Expert Presentations:** Insights from experienced public health professionals and community leaders.
- **Group Projects:** Collaborative development of community surveillance plans.
- **Action Planning:** Development of personalized action plans for implementing community-based surveillance.
- **Digital Tools and Resources:** Utilization of online platforms for collaboration and learning.
- **Peer-to-Peer Learning:** Sharing experiences and insights on community engagement.
- **Post-Training Support:** Access to online forums, mentorship, and continued learning resources.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call +254769199797

Certification

Upon successful completion of this training, participants will be issued with a globally recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

- Participants must be conversant in English.
- Upon completion of training, participants will receive an Authorized Training Certificate.
- The course duration is flexible and can be modified to fit any number of days.
- Course fee includes facilitation, training materials, 2 coffee breaks, buffet lunch, and a Certificate upon successful completion.
- One-year post-training support, consultation, and coaching provided after the course.
- Payment should be made at least a week before the training commencement to Fineskill Training Center account, as indicated in the invoice, to enable better preparation.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com