



## Cross-Sectional & Panel Data Using STATA

### 3-Day Professional Development Workshop

**East Asia Training & Consultancy Pte Ltd** invites you to attend a three-day professional development workshop reviewing econometrics methods and using Stata to analyse the course databases. Stata is the well-known statistics and econometrics software package developed by StataCorp (USA).

#### About STATA

STATA is a statistical software package that offers a broad range of statistics to professional researchers in many disciplines. Stata is particular useful to professionals working in areas of biostatistics, medical research and economic research.

#### Course Programme

The participants will learn STATA tools used for data management, statistical tests, regression analysis using cross-sectional and panel data, and some advanced econometric techniques such as limited dependent variable, sample-selection model, and instrumental variable regressions. The workshop also covers programming with STATA. The topics and tools covered in the workshop are general, useful to anyone who is interested in statistical analysis of real data. Statistical and econometric theories will be explained for each topic, but the emphasis is on learning and understanding the tools provided by STATA and interpreting the results.

During the six practice sessions, the participant will have opportunities to analyze real data using STATA for various interesting issues. The data are selected mostly from those used for actual research projects. For example, we analyze data on Singapore's fertility trend. Using data from the US Census 2000 and National Longitudinal Study, we estimate how an individual's economic status and labor market status are determined. We also use STATA to analyze some household decision issues, with data from Bangladesh and Indonesia.

#### Who should Attend

The target group of the workshop is those who work with cross-sectional or panel data from household or individual surveys---for example, those who conduct market surveys and analyse them, researchers, government workers who work with social survey data and analyse them or researchers who are interested in statistical analysis of real data.

#### Fee & Registration

The fee covers extensive course materials and databases, lectures, luncheons, morning and afternoon tea/coffee breaks, refreshments, receptions and the opportunity to meet and network with researchers & STATA users across the various industries in Asia.

**The number of participants is restricted.** Please register early to guarantee your place. Please complete the official registration form and fax to (65)-62506369 or email it to us at [stata@eastasiatc.com.sg](mailto:stata@eastasiatc.com.sg) to reserve your place. Confirmation will only be made upon receiving full payment. Further instructions will be sent to confirmed participants.

The workshop is hands-on. Delegates are required to bring their own laptops with Stata software installed.

### MAS Financial Sector Development Fund (FSDF)

Participants may be eligible for Financial Sector Development Fund (FSDF) support on a case by case basis. Interested applicants should submit their applications to the FSDF Secretariat directly. For enquiries, please contact the FSDF secretariat at 65- 6229 9396 or via email at fsdf@mas.gov.sg.

You may use the printable MAS FSDF application form in Word format posted on our website. Please submit your applications to the FSDF Secretariat directly at least 6 weeks prior to the commencement of the course.

### Course Outline

Each day will comprise three formal sessions:

- A talk on various aspects and procedures in Stata
- Worked examples
- Assignment

There will also be the opportunity each day to work on your own data, and discuss procedures not listed below.

## Day 1

### Morning Session 1

- Introduction to STATA
  - o Settings of STATA
  - o Getting help and updates
  - o Command syntax
  - o Running STATA in a batch mode and logging STATA
- Getting data into STATA
  - o Reading ASCII data
  - o Conversion of data from other programs
  - o Variable types and names
- Data management I
  - o Getting basic information on data
  - o Generating/replacing/labeling/deleting variables
  - o Renaming/reordering/sorting variables
  - o Creating categorical variables

Tea-Break

### Morning Session 2

- Practice session
  - o We create and use a simple data set for basic introduction to STATA. Then we use data on Singapore's fertility trend to practice the commands we learned in the session.

Lunch

### Afternoon Session 1

- Data management II
  - o Random sampling
  - o Combining data sets
  - o Reshaping data
  
- Descriptive analysis of data
  - o Getting summary statistics
  - o Tabulations
  - o Graphs

Tea-Break

### Afternoon Session 2

- Practice session
  - o We practice data management commands using STATA sample data sets on automobiles. For the descriptive analysis, we use data on Singapore's fertility trend to analyze secular changes in fertility patterns and other relevant variables.

## Day 2

### Morning Session 1

- Linear regression
  - o Running regressions using STATA
  - o Categorical explanatory (dummy) variables
  - o Diagnostic tests
  - o Hypothesis tests
  - o Prediction
  
- Regression with a binary dependent variable
  - o Linear probability model
  - o Probit and Logit regression

Tea-Break

### Morning Session 2

- Practice session
  - o For this session, we use data from the US Census in 2000. Using linear regressions, we estimate how an individual's work status and earnings are determined. We analyze how education level, age, sex, race, marital status, and language skills affect an individual's economic status.

Lunch

Afternoon Session 1

- Instrumental variable (IV) and two-stage regression (2SLS)
  - o Introduction to Instrumental Variables
  - o IV regression using STATA
  - o Two-stage least squares
  - Tests on validity of IV
  
- Limited dependent variable
  - o Truncated data
  - o Multiple categorical dependent variable

Tea-Break

Afternoon Session 2

- Practice session
  - o For 2SLS exercise, we use data from Bangladesh to examine the relationship between household income and expenditures. For analysis of limited dependent variable, we estimate how remittances received and sent by a family are determined using data from Indonesia.

**Day 3**

Morning Session 1

- Regression with panel data
  - o Random effect model
  - o Fixed effect model
  - o Panel IV model
  - o Tests
  - o Panel limited dependent variable model
  
- Analysis of survey data
  - o Statistical issues with survey data
  - o Analyzing complex survey data with STATA
  - o Regression with survey data using STATA

Tea-Break

Morning Session 2

- Practice session
  - o For exercise with panel data, we use well-known National Longitudinal Study of the US that has been used for numerous studies in the US. Using the data, we estimate a wage equation and work status equation with different setups.

- o For analysis of survey data, we use the US Census 2000 data. We revisit estimation of wage and work status equations and compare the results to the previous results.

Lunch

Afternoon Session 1

- Programming STATA
  - o Macros
  - o Program arguments
  - o Accessing/Saving results
  - o Ado-files
  - o Writing help file

Tea-Break

Afternoon Session 2

- Practice session
  - We practice some programming problems. For example, how to repeat the same task, how to instruct STATA to do different tasks by cases, etc. A few programming problems will be given to the participants.