Basic Data Analysis Course on STATA

Introduction to Stata

Md Monowar Hossain

Institute of Applied Statistics and Data Science (IASDS)
University of Dhaka
mmhossain1@isrt.ac.bd
mdmonowarhossain.netlify.app

July 8, 2026

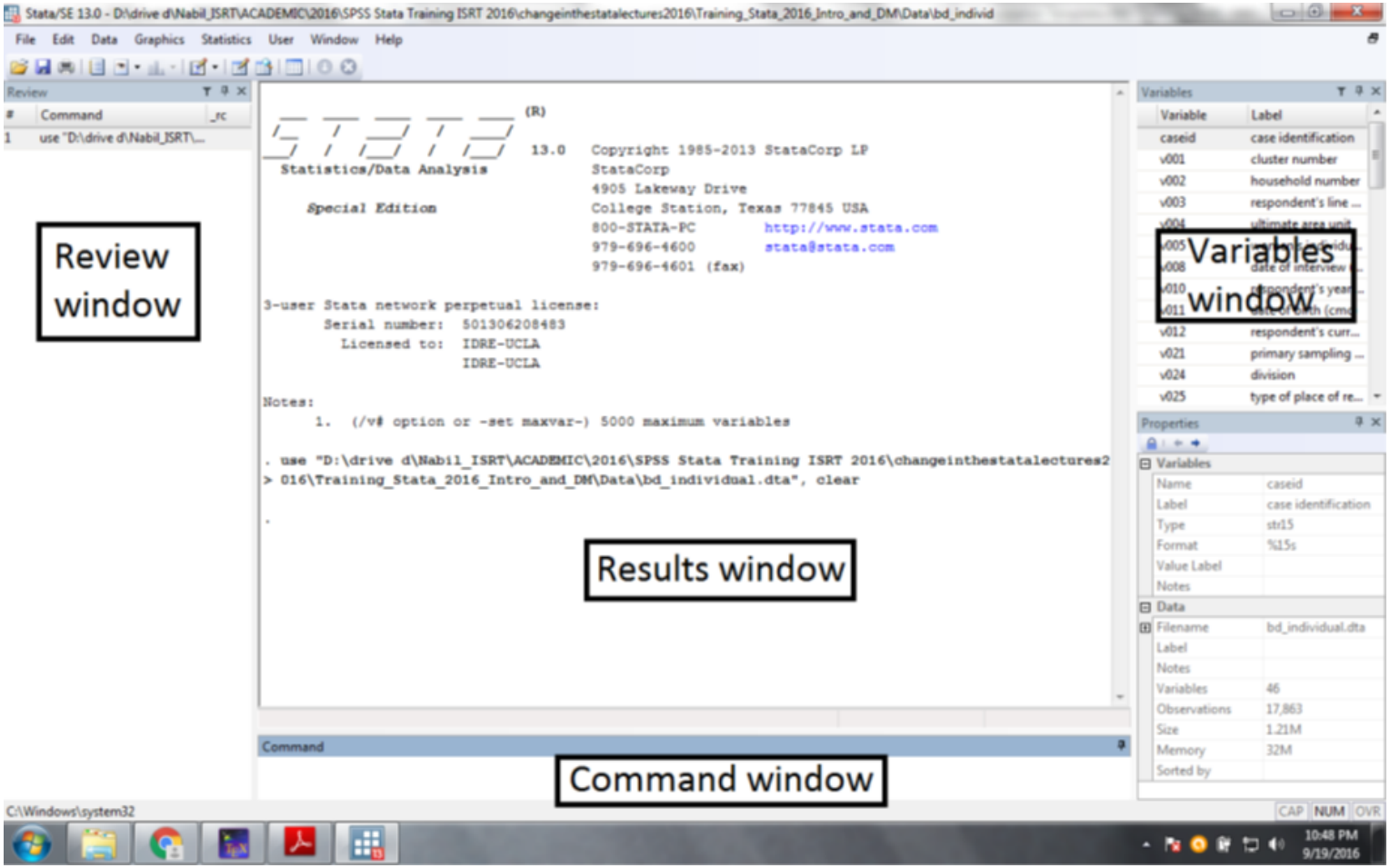
About Stata software

- Stata is a software package used for statistical analysis. The name Stata is a syllabic abbreviation of the words statistics and data.
- Stata was developed by StataCorp in 1985.
- Stata has always emphasized a command-line interface, which facilitates replicable analyses. Starting with version 8.0, however, Stata has included a graphical user interface which uses menus and dialog boxes.
- The latest release is Stata version 18, released in April, 2023.
- Purchase link: https://www.stata.com/order/dev/

Overview of Stata file extensions

- Stata data file: .dta, example: mydata.dta
- Stata syntax file: .do, example: mysyntax.do
- Stata output file: .smcl, example: myoutput.smcl, smcl stands for
- Stata Markup and Control Language

Stata windows



Stata windows

There is an additional window for graph that appears when a graph is produced in Stata.

- Command window: where commands are entered. Press the Enter key on your keyboard to execute any command.
- Results window: where results appear (i.e, are printed).
- Variables window: lists the variables in the current dataset.
- Review window: lists past (i.e, already executed) commands.
- Graph window: where graphs are displayed (appears when graphs are generated).

GUI and CLI

- You can either use the menu-based **Graphical User Interface** (GUI) or the command-driven **Command Line Interface** (CLI) to interact with Stata. But there are some advantages of using CLI:
  - **Command syntax** has the advantage of reproducibility.
  - Many of the commands are short and easy to execute, and as proficiency grows, it is often faster to type commands rather than use the menu.

We will mostly cover command syntax in this course. We will be providing commands so that you will become familiar with them. The commands will be useful for constructing longer and more complicated programs.

- The command executes if it is entered properly
- Issues an error message (in red) if it cannot figure out what to do, or if there is something which could inadvertently change the dataset.
- If Stata gives you an error, it will provide you with a return code, r(#). You can find out more about the error by typing: lookup rc #.
- To get help regarding any command, e.g. the regression command, type: help regression

THANK YOU