ECON 4003 Econometrics I

Introduction to STATA

About Lab Sessions

- 9 hours tutor-led computing sessions (6 x 1.5 hour): Week 3, 6-10
- GOAL: Use statistical software (Stata) to perform econometric analysis on empirical data
- Download and have a look at materials (exercises and datasets) from Moodle page in advance

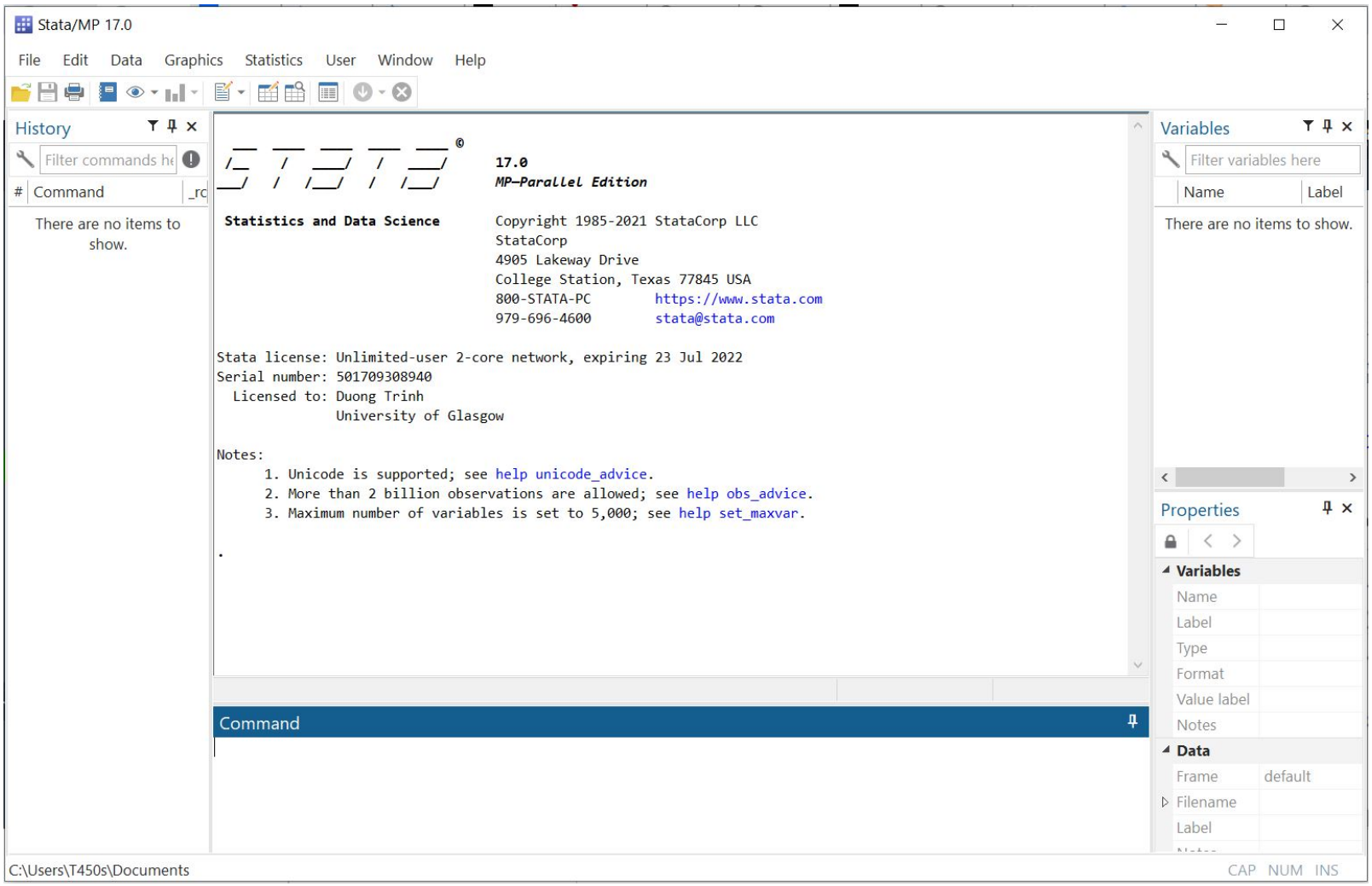
Overview - Introduction to Stata

- Little or no previous experience in Stata before
- Learning Objectives:
  - Familiarise yourself with the Stata interface
  - Set the working directory and get data in and out of Stata
  - Perform basic data analysis: Data Description, Graphics, Data Management

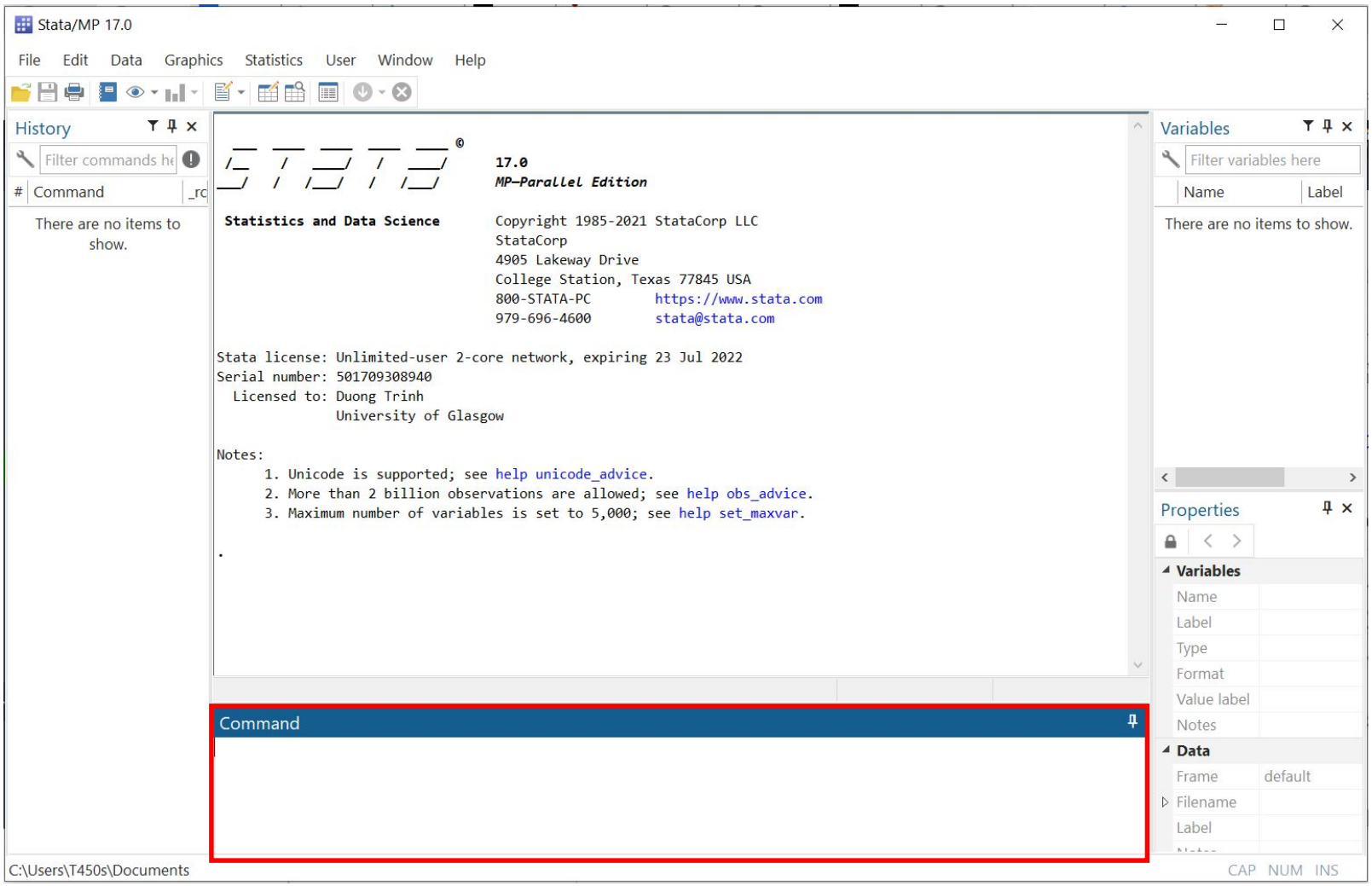
What is Stata? Statistics and Data Analysis

- A statistical package that includes a wide variety of capabilities
  - Data management
  - Statistical and econometric analysis
  - Graphics etc.
- Widely used in the fields of economics, finance, political science, sociology, biomedicine and epidemiology
- Three main versions
  - Stata/IC (Intercooled): mid-sized datasets
  - Stata/SE (Special Edition): large datasets
  - Stata/MP (Multi-processor): fastest version (for quad-core, dual-core, and multicore/multiprocessor computers)

Stata Interface

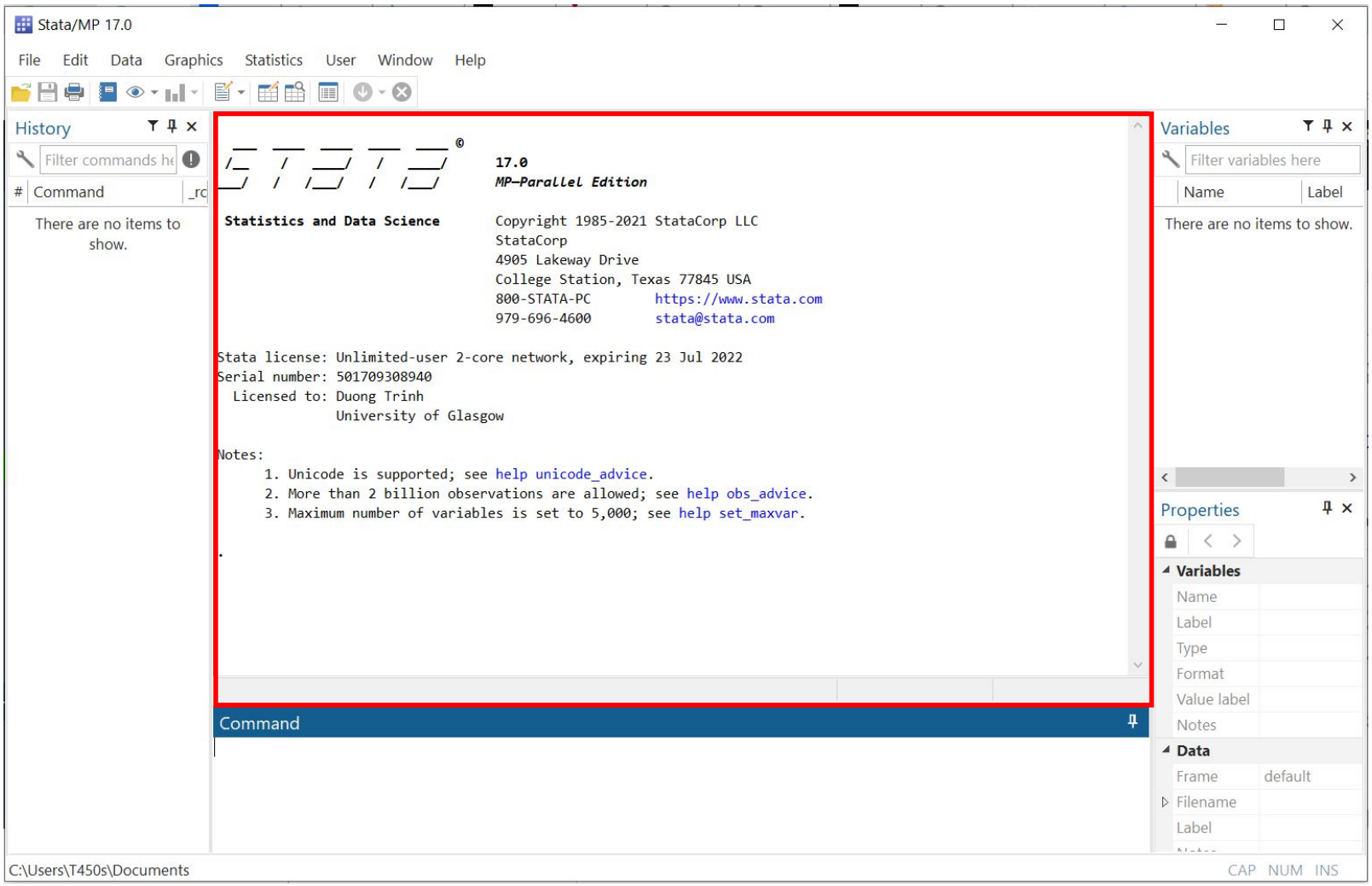


Stata Interface: Command Window



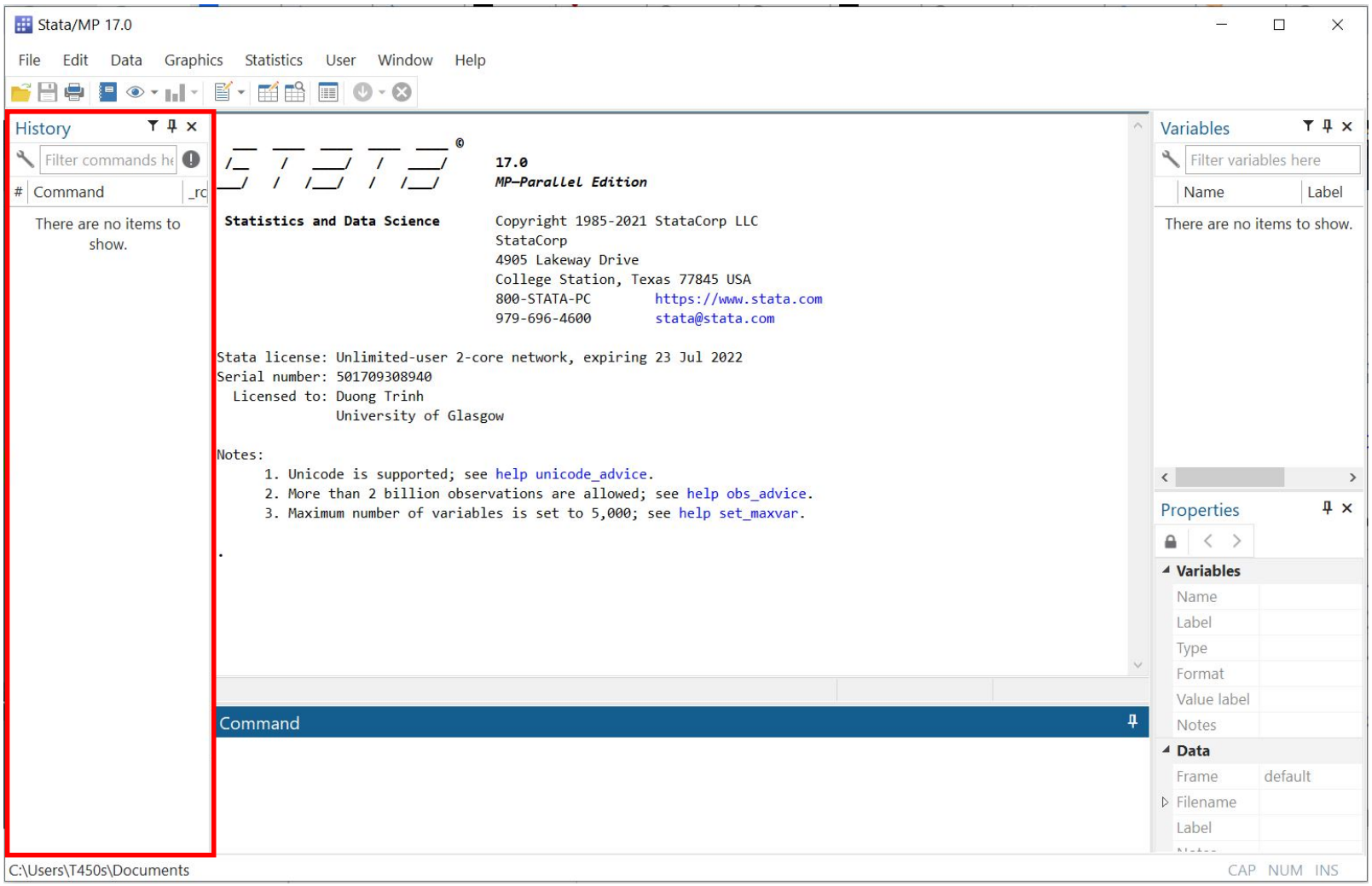
Type the instructions we want Stata to execute

Stata Interface: Results Window



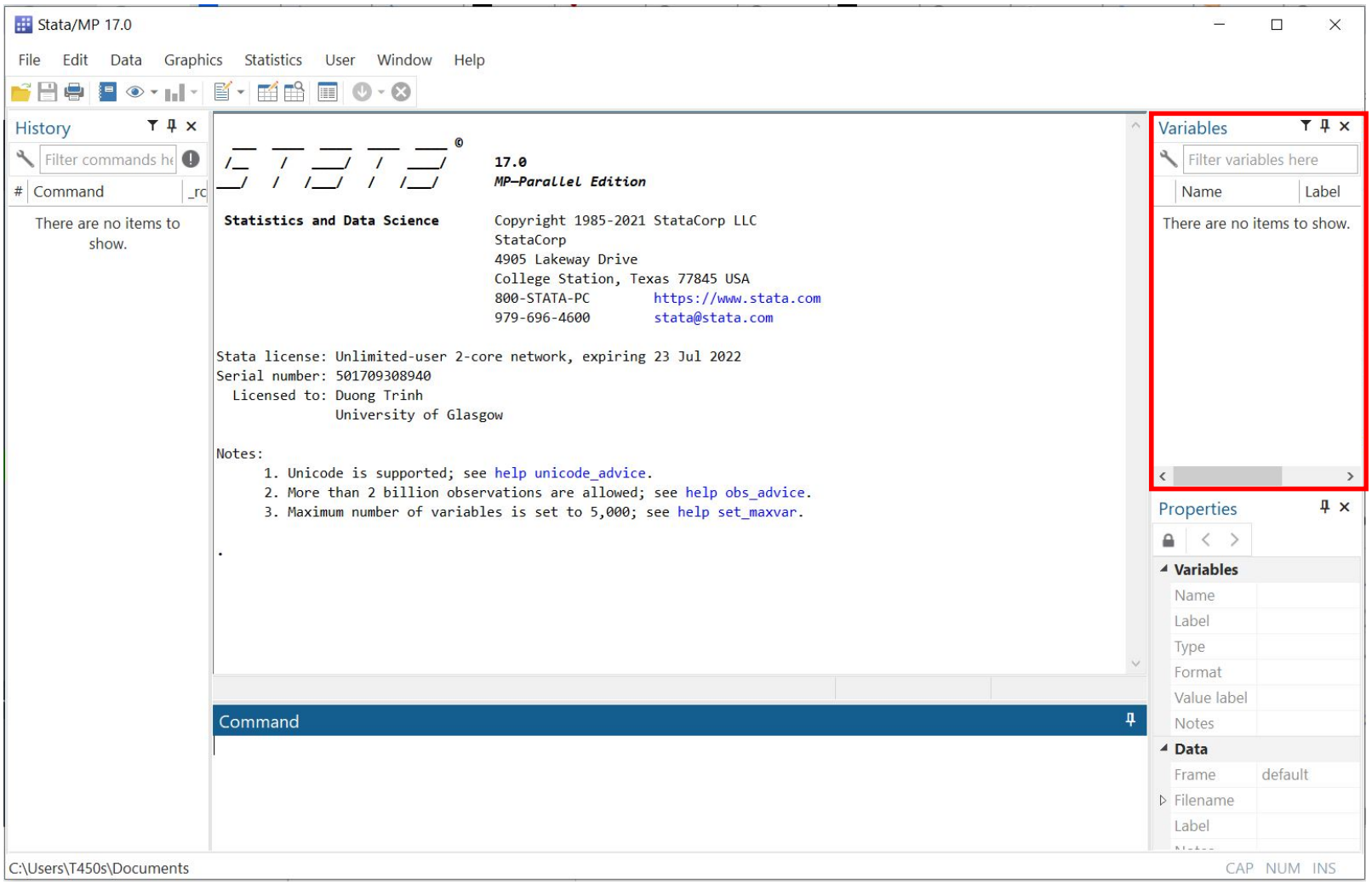
Displays the results and outputs after a command is executed

Stata Interface: Review Window



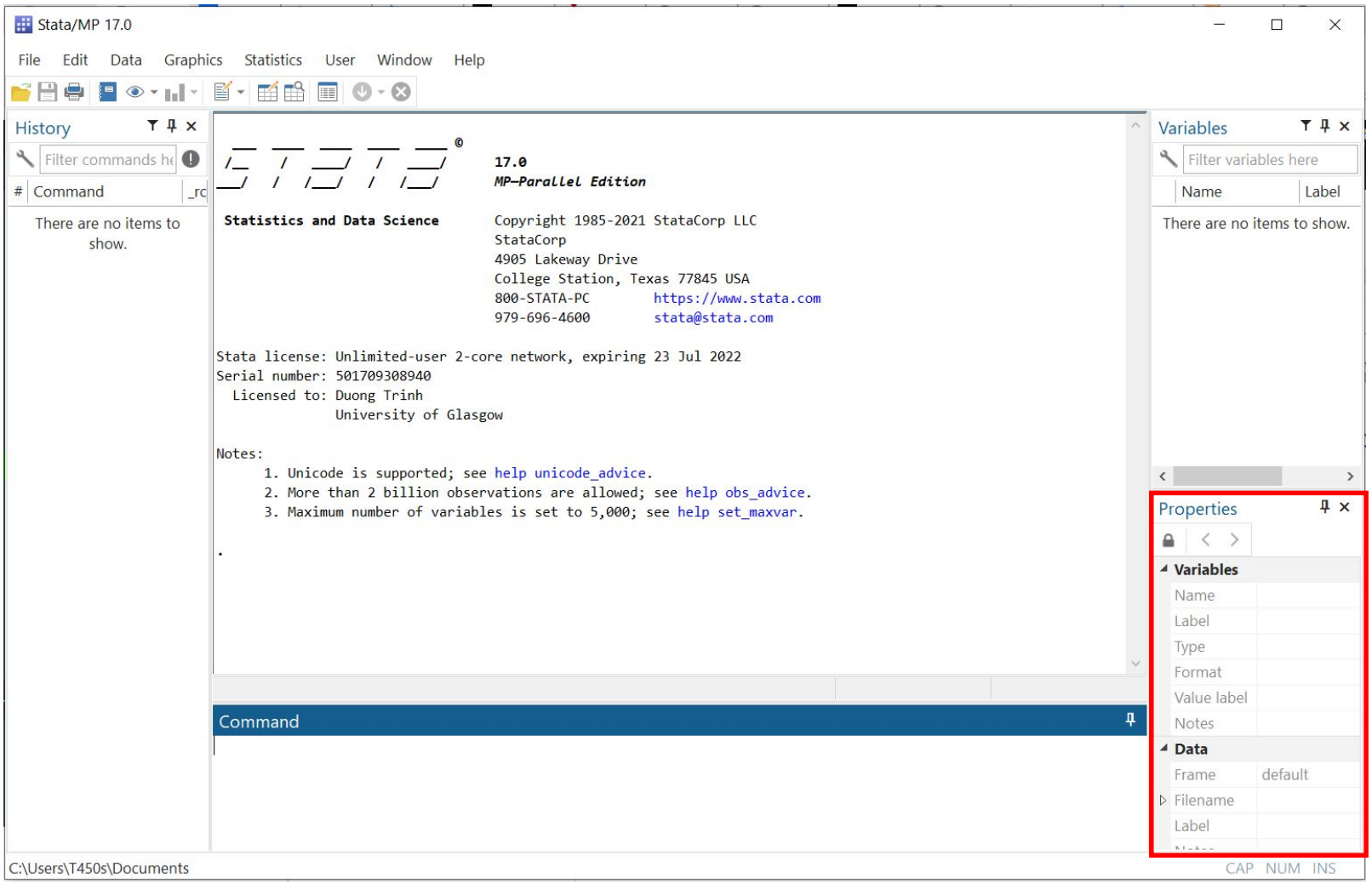
Keeps a record of all the commands used/the input history

Stata Interface: Variables Window



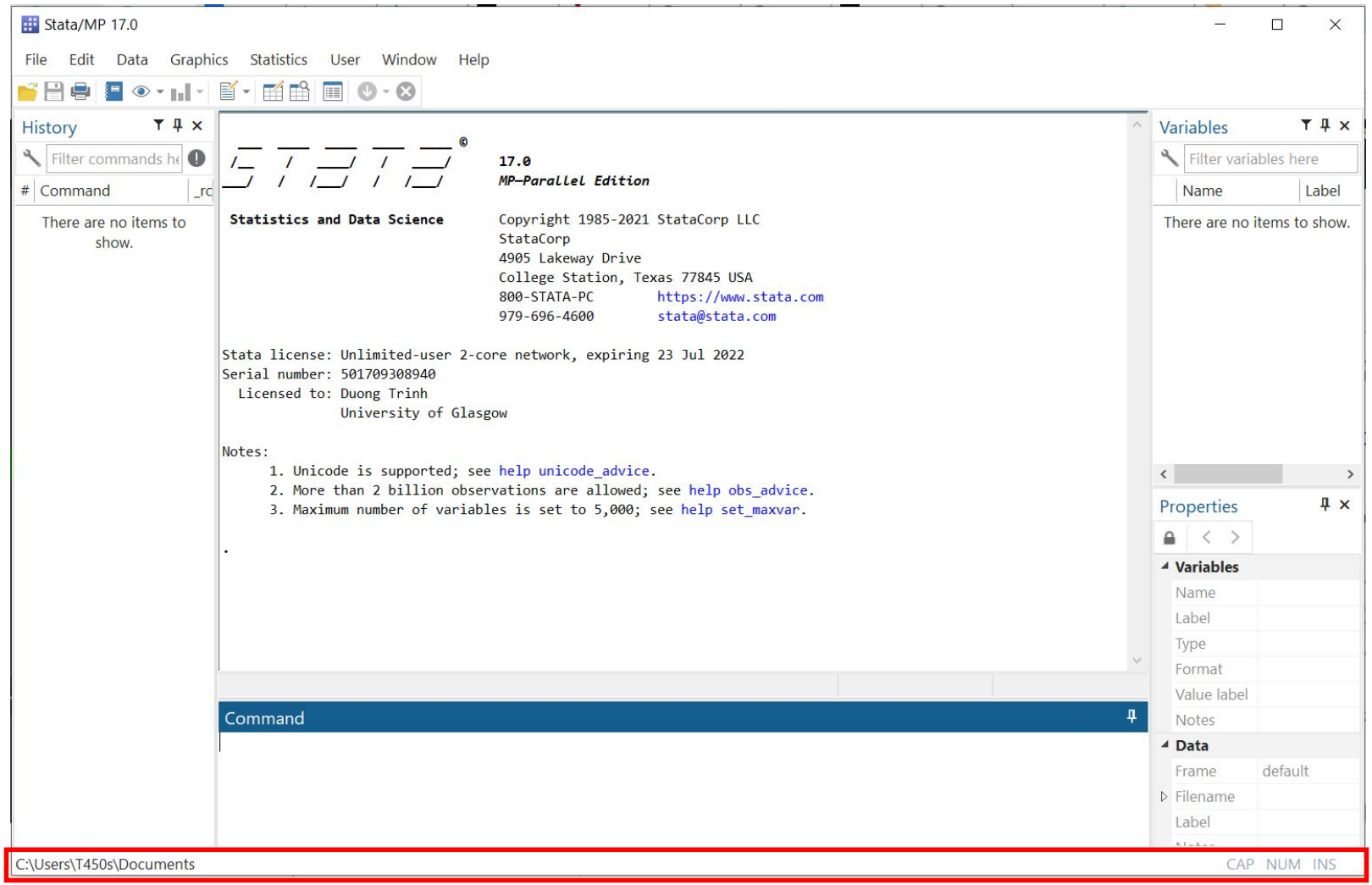
Shows all the variables in the dataset

Stata Interface: Properties Window



Indicates the properties of a highlighted variable

Stata Interface: Current Working Directory



Shows current directory in file system from where Stata will read or save any file

Stata Characteristics

- Basic syntax:

```
command varlist , options
```

- Relational operations
  - $>$ greater than
  - $<$ less than
  - $>=$ greater than or equal
  - $<=$ less than or equal
  - $==$ equal
  - $!=$ not equal

  (!)CAUTION $=$ is NOT equal to $==$
- Logical operators
  - & and
  - | or
  - ! not

Stata Characteristics

- Mathematical operations
  - $+$ addition
  - $-$ subtraction
  - $*$ multiplication
  - $/$ division
  - $\wedge$ exponentiation
- Use $*$ for comments
- Use "" for strings
- Capitalization consistency
  - Stata is sensitive to capitalized words for variables
  - Stata does not understand capitalized commands

PRACTICE: Data Description

*Ex1: which is the mean of the variable tenure for those women who live in the south?

*Ex2: which is the working hour's mean for married women who belong to a union?

*Ex3: is it higher or lower than the working hours average of single women who belong to a union?

PRACTICE: Graphics

*Ex1: draw pie chart for "occupation"

*Ex2: create a bar graph for the variable "married"

*Ex3: create a horizonal bar graph for the variable "industry"

*Ex4: repeat the hist exercise above for the variable "hours"

*Ex5: create a scatter plot of "wage" on "ttl_exp"

*Ex6: create a scatter plots of "wage" on "ttl_exp" by "married"

PRACTICE: Data Management

*Ex1: create a new variable called “totwage” which is “wage” multiplied by “hours”.

*Ex2: replace “totwage” with the natural log of “totwage”? (hint: natural log of x is “log(x)”).

*Ex3: what is the mean of ”totwage”?