**STATA Tutorial**

- The Stata Interface:

  - When Stata starts up you see four docked windows.

    1. Results window: All outputs appear in this window. Graphics will appear in a separate window.
    2. Command window: This is where you enter your commands
    3. Variables window: Lists all variables in the open dataset. When you click on a variable the name can be transferred to the command window.
    4. Review window: Previous commands are listed here. This allows you to keep track of the commands you have used.

- Toolbar: Most important buttons

  - Open: opens a new data set. You can also just type use nameof dataset
  - Save: Save the current data file. You can also just type save, replace
  - New Do-file Editor: opens a new do file. You can also just type Crlt+d
  - Data Editor: Opens the data editor and allows changes
  - Data Browser: Opens the data editor and does not allow changes

- Typing Commands

  - Stata can work as a calculator using the display command. Try typing the following: display 2+2

- Commands can also be abbreviated: di 2+2
- The Page Up key recall command history using the Page Up and Page Down keys
- You can edit it using the arrow, insert and delete keys, which work exactly as you would expect.

- STATA Help

  - Stata has excellent online help. To obtain help on a command (or function) type **help** command name
  - Stata has a search command. Type help search TOPIC.
    * Try search Student's t. This will list all Stata commands and functions related to the t distribution.

- Do-Files:

  - Saves commands as a script in a simple text file with the ending .do.
  - Run the do-file
  - If all steps of a project have been logged in one or more do-files, all analyses and results can be reproduced

- General Commands

  - update: updates installed STATA
  - cd: changes directory. Stata uses a working directory if no path is entered.
  - set memory: you can increase the memory to allow for larger data sets: set memory 600m
  - insheet: If using non-Stata dataset you will need to read it into STATA.

- use: opens a STATA dataset.
- edit: automatically opens editor
- save: saves your data set. if you have made changes will need to use save, replace
- by: tells Stata to execute command across categories of a variable
- if: use at the end of a command to one use certain cases (males or females)
- e(sample): subcommand on only data in previous model

- Data Manipulation

  - generate: create new variable
  - egen: more extensive set of generation options (see also egenmore)
  - replace: change values of an existing variable
  - recode: useful to change categories of variable.
  - drop: delete variables or obs of variable
  - keep: opposite of drop; keeps only subset of variables or obs.
  - destring: changes string variable (with only numbers) to numeric variable
  - rename: allows you to rename a variable
  - label: two ways to label a variable
    1. label variable: label variable varname "newname"
    2. label define: label define labelname 1 "label" 2 "label"

- Data description

  - describe: displays information about the data

- codebook: gives information on one or more variables
- sort: sort the data by specific variables
- order: reorder variables
- list: value of variables can be listed in te results window
- summarize (sum): descriptive statistics on numerical variables.
- tabulate (tab): frequency table

- Data Merging

  - append: append data to the end of current data set.
  - merge: merge two data sets based on common variable

- Log Files

  - log using filename
  - use text or log...default scml can only STATA viewer.

- Loading a Sample Data File

  - Stata comes with a few sample data files. I will use lifeexp.dta, which has data on life expectancy and gross national product (GNP) per capita in 1998 for 68 countries.
  - To see a list of the files shipped with Stata type sysuse dir.
    * sysuse lifeexp