Handout 1
PO8006 - Andrea Salvi
Hilary Term 2018/2019

Reference: Stata Companion 1, 11

1 Admin Stuff

What's the purpose of this?

- Furthering what we did in class and explore its applications;

- Deepen the concepts of quants and statistical modeling;

- Hands-on STATA;

- Real-world applications of theories, concepts and quantitative methods;

- Revision of Homeworks;

- Q&A;

- **NB: tutorials do not replace the lectures and viceversa!.**

2 Hi STATA!

2.1 Get to know the interface:

The interface is built upon 4 main panels:

- Command: type commands (almost NEVER use it)

- Results: output

- Variables: names of variables in current dataset

- Review: keep record of commands

    - (you can click on it and repeat the same command).
    - You can choose the order of appearance of the commands by clicking on Command (decreasing, increasing).

        You also have the databrowser to look at your data.

        If you accidentally close one window Window click and add it again.

        -use "filename.dta"-

        -use13 "filename.dta"-

2.2 Ways to interact with STATA

2. How to work with STATA

Using pull-down menus, 'working interactively', and writing do files.

Pull-down menus You will see in the pull-down menus under Statistics, for example, a series of options for various types of statistical tests. It should be pointed out early on to the new user that there are more statistical possibilities in Stata than there are options in the pull-down menus which open dialogue boxes.

Working interactively' thorugh Command Window Simply type in the Command window. Ex try to type -summarize-

Writing do files THAT'S THE BEST WAY TO WORK WITH STATA. Instead of just typing commands into the Command window or using the pull-down menus, it is likely that you will want to keep a record of your commands so that you can refer to them (and run them again) later. Write code and execute is using Do. N.B. Do exectute the whole file, if you want to execute only one line you have to select it before running it.

Data Storage Data Editor and the one on the right is the Data Browser (with the magnifying glass). You can physically change the data in your data file in the Data Editor, but not in the Data Browser. Only the Data Editor or the Data Browser may be open at one time (not both). Also, you must close the Data Editor and Data Browser before Stata will run any commands.

3. Directory, Do file, Log file, Memory

To check directory type: -pwd- To set wd: -cd- and then put the directory in quotes Using the pull-down menu: File Change working directory Then browse to the location of the directory/folder you wish to use. Now if you save any files without explicitly defining another directory, your files will be saved in this directory.

Always use a do file for your analyses! Open it from the icon on the menu bar. And save it. -delimit ;- for long lines of code (you can also simply always use it if u want, but then remember to use it everytime!)

-*- beginngin and end of comment, only the asterisk for short comment (1 line!)

File –¿ Log –¿ View here u can open and save a log file to save all otuput (Not necessary I usually do no use it, you don't need it if you ALWAYS work on a nice do.file)

-clear- -clear all- Clear memory

4. Help Installing packages

-search- use it when you know what u want to do, but have no idea of the command.

-help- use it when you know the command but do not remember how it works.

-scc install package name- to install package

To find all packages at SSC that start with word, type -ssc describe word-

5. Data in and out File –¿ Import (here u can choose the format of dataset)

Open icon if u are using stata files

Stata format: easiest format, simply download and is ready to use. p. 191 some datasets ...

-use filename.dta, clear- if you have already defined the directory you do not need to insert file directory here.

Excel format: make sure that the cells type in excel is numeric (change it on excel). When u copy and paste do not select the excel label, stata creates label when u insert the data (var 1, var 2 ...). You can change it using Tools (Data on stata11) –¿ Variable manager. This is the way your manual explains how to do it. You can if you want, but copying and pasting is error prone... -import excel filename- (or use File–¿import) Remember to click firstrow if you want the first row to be the variables labels.

CSV format: -import delimited using "filename.csv", varn(1) clear-

SPSS format: usespss using "pathandfilename", clear

Html format: here u could need a preliminary Excel phase to fix data (see example p. 194) and then simply copy in stata editor as before.

Other formats see p. 6 PDF Handbook.

You can also insert data into the File Editor and then save as .dta

6. Describe Datasets

Never modify the original dataset! NEVER!!! Just use the .do file to reproduce your modified dataset/s.

Open gss1012. NB: you can always open files using File–¿Open. It is probably better to simply set your wd and then type -use- and file name "use gss2012".

-describe- describe dataset that you have just opened.

-codebook varname- describe variable

nb "." indicates missing observations (STATA will not include it)

the -more- option (keep it, and just use the green arrow if u want to arrive at the end without looking at the whole result output). To print: select what u want to print, right-click, select selection and print.

If you want to keep or drop variables: -keep varnames- -drop varnames-

JOINING DATASETS (P. 11 Handbook) Data –¿ combine datasets –¿ merge or append

Append: TO ADD NEW OBSERVATIONS (stack datasets on top of each other) ex. Join two rounds of a survey.

-append using newdataset-

Merge: TO ADD NEW VARIABLES (a bit more complex) 1. find a unique identifier that u will use to merge data.

2. Sort data by your identifier: -sort identifier-

3. Merge the new dataset: -merge 1:1 identifier using newdataset-

Note: u will find a new variable $_merge(givesuinfoonmergeddata).Examplewithimport GDPpercapitaper$

NB: You can have a look at File –¿ Example Dataset on STATA.

https://www.youtube.com/watch?v=2Fqz4ARUaFc

6 Create Variables

-generate newvarname = expression-

The egen command is useful for working across groups of variables or within groups of observations.

-egen newvarname = expression-

Generally speaking, if you do not remember how to do something in STATA: 1. DON'T PANIC 2. HAVE A LOOK AT THE MANUAL 3. IF YOU DO NOT FIND THE ANSWER IN THE MANUAL 4. DON'T PANIC 5. GOOGLE IT! (www.stata.com; www.stackoverflow.com; http://www.timberlake.co.uk; http://statcomp.ats.ucla.edu/stata/ are reliable sources)