Tutorial: Getting Started with Stata

Getting Started

- Download the Framingham dataset to your computer. Next, open Stata.

Stata Menus

Most of the commands that you will use are available through the use of drop-down menus. The menus can be found on the Menu bar at the top of the Stata window. The following is a description of the most important Stata menus:

- *File* – open and save data; open and close log files; exit Stata.
- *Data* – data management commands (open data editor or browser; summarize dataset; label dataset and variables; generate new variables).
- *Graphics* –all of Stata's graphing tools.
- *Statistics* – all statistical commands.
- *Help* – resource for how to use Stata.

Stata Windows

The Stata interface is divided into five windows, four of which appear when you open Stata.

- *Command* Window - where you type in commands into Stata in order to perform statistical analyses, plot graphs, etc. Most of the commands that you will use in this course are available in the *Data*, *Statistics* and *Graphics* menus. You can choose between the "drop-down menu" or "command-line" approaches to using Stata.
- *Results* Window - where the results of any command performed by Stata appear.
- *Variables* Window – contains a list of variable names and labels for your data.
- *Review* Window - the *Review* window is useful because, as you send commands to Stata, they appear there (even if you use the drop-down menus to perform your analyses). If you are interested in learning the command-line approach to Stata, then noticing the commands in the *Review* window is a good place to start. You can call back commands in the *Review* window by clicking on them, and then editing them in the *Command* window. Or, you can re-run a command by double-clicking on the command in the *Review* window.
- *Graph* Window - appears when you create a graph.

Opening Existing Stata Data Files

Stata data files have the extension `.dta`. You can open `.dta` files in Stata. Other file types are *imported* into Stata. For instance, to import a tab or comma delimited text file, go to *File/Import/Text data created by a spreadsheet.*

The data set named `fhs.dta` contains data the Framingham dataset that we will use throughout this course.

- Open the data set `fhs.dta` using *File/Open* and selecting the correct file in the directory where you have saved the data. You will see the list of variable names in the data set appear in the variable window.

Also note that the command `use "…/fhs.dta", clear` appears in the *Results* and *Review* windows. Typing this command into the *Command* window instead of using *File/Open* would have also opened the data set.

"Opening" the data set means that Stata makes a copy of the data set into memory. The original is untouched unless you overwrite it by using the `save` command.

Data Editor

The Data Editor is where you can enter new data, make changes to the current dataset, or create new variables. Typically, when analyzing data, you will not make changes to your dataset using the Data Editor.

- Open the Data Editor by clicking the Data Editor button on the toolbar at the top of the screen. The button looks like a little spreadsheet. Alternatively you could choose *Data/Data editor* from the menu bar.

Note: The Data Editor is different from the Data Browser. The Browser only allows you to look at the data and does not allow you to make changes to the dataset. The button for the Data Browser looks like a magnifying glass, which is also located on the toolbar, to the right of the Data Editor button. If you just want to look at your data and don't want to make any changes, it is better to use the Browser for two reasons (1) you can't accidentally change your data, and (2) it is much faster to open the data (more relevant for large datasets).

- Sort the data on the variable `age1`.
    - Highlight the column for `age1` by clicking on the variable name.
    - Click the *Sort* button. Watch how the order of the observations changes.
    - Typing `sort age1` in the command line accomplishes the same thing. (Note that you can click on the name of the variable in the *Variables* window to bring the variable name into the *Command* window.)

- Now try deleting data.
    - Press the `delete` button and look at the resulting *dialog box*.
    - Try deleting an observation.
    - Then delete a variable.

Note that Stata does not save any of your changes to disk until you explicitly tell it to (*File/Save*), generally when exiting the program. **Please do not save your changes.** It is important that everyone is working with the same version of the `fhs.dta` dataset, and saving any changes would lead to inconsistency in versions of this dataset. If you accidentally change the dataset, download the dataset from the course website again.