

Introduction to Stata

Kristy Krivickas

CFDR Workshop Series

Fall 2008

Stata: Why should we use it?

- User friendly
- Great help links
- Can start out with dialog boxes
- More sophisticated analysis than in SPSS
 - Differences between Stata and SAS are mostly technical, i.e. reading in data (see *Stata for the Struggling SAS Mind* for more info)
- Constantly being updated (mostly by other Stata users)
 - Updates can be installed using the Internet from within Stata

Getting Started

- Preferences
- Four window screens
- Push pins
- Command structure
 - No punctuation at end of command
 - Case sensitive

The Basics

- Dialog boxes and command area
- Opening data
 - use “C:\Documents and Settings\kmmkrivi\Desktop\sampladata.dta”, clear
 - The cd command will list the path name of your data
- Data editor window
- Basic command examples
 - Describe
 - Summarize

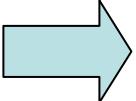
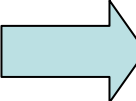
The Basics

- Operators
 - Symbols only, no mnemonic equivalents
 - Symbols need to be in correct order
 - \neq NOT \Rightarrow
 - Examples
 - $\&$ And
 - $=$ Is or is equal to (equality symbol)
 - $>$ Greater than or equal to
 - $|$ Or
 - \sim Not
 - \wedge Power

The “log” and “do” files

- The log file
 - Copy of everything sent to the Results window
 - Keeps all your output
 - Creates a text file which can be transferred to excel or word
 - Can use a new log file, add on to an existing file, and temporarily suspend logging
 - Automatically has extension .smcl can change to .log

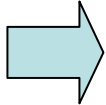
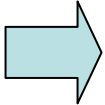
The “log” and “do” files

- How to create a log file
- File  Log  Begin
- Use the log command
 - log using *filename*
- Example of log file is attached to this handout

The “log” and “do” files

- The do file
- What the Stat program is written in
- Includes definition statements, commands, recodes
- Similar to syntax file in SPSS
- The do file is a separate file
 - .do file

The “log” and “do” files

- How to create a .do file
- Window  Do-file Editor  New Do-file
 - OR Click on the icon with the note pad and pencil
- Example .do file is attached to this handout
- You can also open your log file in your do file

Basic Commands

- *des* – List of variables
- *codebook* – Lists info about each variable
- *tab var1* – Creates frequency distributions
- *tab var1 var2* – Creates crosstab
- *su* – Displays means for all variables
- *su var1* – Mean for specific variable
- *su var1, detail* – Detailed descriptives
- *su gpa if sex == 2* – Mean gpa for girls

Creating a new variable

- *generate* – Creates new variable
 - Examples
 - *gen rate* = time*cost
 - *gen loginc* =log(inc)
 - *gen male* = (sex == 1)
 - *gen ymorelate* = mopraise + mocritr + mohelp

Sorting Data

- Sorting data
 - *sort var1 var2* – Sorts data by var1 and var2
 - *bysort* – Sorts variables for a command
 - *bysort gender: su gpa* – Mean gpa by gender

Online Help

- On-line STATA tutorials:
- <http://www.ats.ucla.edu/stat/stata/>
- <http://www.princeton.edu/~erp/stata/main.html>
- <http://library.syr.edu/information/mgi/nds/software/stata/intro/index.html>

Stata FAQ's and Resources

- <http://www.ats.ucla.edu/stat/stata/faq/default.htm>
- Acock, Alan C. 2006. *A Gentle Introduction to Stata*. Stata Press.

Additional Help

- CFDR Research Tools
 - Stata user guide
 - Annotated output
- Programming Hours