



# STATA WORKSHOP 2

ERL Workshop for Sociology Fall 2014

NITTY GRITTY DATA SAVVY





# STEPS TO SUCCESS

- Write your codebook
  - By doing this first you take care of all the decision making, so you don't have to go back to make changes later
- Entering your data into Excel
- Importing your data into Stata
- Setting up to do your analysis
- Running your statistics

# WRITING YOUR CODEBOOK

- Your codebook will include the exact wording of the question
  - “Who is your favorite Ninja Turtle?”
- The variable name
  - NinjaTurtle
- The numbers used to represent each category (for variables that have categories)
  - 1 = Michaelangelo
  - 2 = Donatello
  - 3 = Raphael
  - 4 = Leonardo
- You can use the GSS [Codebook](#) as a model (or not)
- I usually list the questions in the order I asked them (if I asked them always in the same order) in a Word document.

# ASSIGNING VARIABLE NAMES

- Call your variables something that will be easy for you to type in Stata, but still be meaningful for you
- Stata will allow letters, numbers, and \_ in variable names
- Variable names cannot begin with numbers
- Stata distinguishes upper & lower case (Gender != gender) \*not equal
- Don't call your variables a common command in Stata

# SETTING UP YOUR DATA IN EXCEL

- List variable names along the first row
- Use unique numbers to identify participants
- Each row below the first represents a unique participant
- Enter responses for each participant
- Remember to enter the NUMBERS for the categories in your nominal & ordinal variables!
  - Without the numbers, Stata can't do the statistics.





**DOUBLE CHECK YOUR  
DATA!!!!**



**JUST DO IT!!!!**



# TRANSFERRING YOUR DATA FROM EXCEL TO STATA

## THE EASIEST WAY:

- In Excel
  - Select the entire data (CNTL A, or use your mouse)
  - Copy it (CNTL C, or right click, or use the ribbon)
- In Stata
  - Click **EDIT Data** button
  - Paste the data (CNTL V, or right click, or use Edit in menu)
  - Select 'Treat first row as variable names'.
  - Now **EXIT** edit mode. Any typing you do here will edit the data and you may not be able to undo it!!
    - Either look at your data in Browse mode or return to the Command window

TA-DA!!

# USEFUL COMMANDS FOR ORGANIZING YOUR DATA

Since everything is in numbers, it is worthwhile to label your variables (it will make interpreting your analysis easier).

- Labeling categorical values
  - *label define labelvariable 1 "label 1" 2 "label 2"*
  - *label values variablename labelvariable*

You may or may not need this depending on how your questions are set up.

- Recoding values
  - *recode variable (x = y) (a/c = z), generate newvariable*
    - where a, b, c, x, y, & z represent actual numbers.

# SAVING YOUR DATA

This is important to do especially after you've labeled and recoded your variables. You won't be able to view the data file without Stata, but if you need to run more analyses later, you don't have to redo the labeling.

- Data file extension is .dta
- Click "save file" icon in menu, go to File → Save, or CNTL S
- Notice that the extension is .dta (if it isn't, then you are NOT saving your data)

# LOG FILE

Remember to start a Log file so that you can view the output on another computer!

- \*\*\*\* Log files \*\*\*\*
  - File → Log → Begin
  - File Name → whatever makes sense to you
  - **Save as type: Log** \*so that you can review the output on a computer without Stata
  - \*\*\* Remember to close the log file when you are done
    - File → Log → Close

# DO FILE

Do files: allow you to save all of the commands you run. This can be very helpful if you find an analysis doesn't make sense and you need to rerun the same tests again (or maybe your log file got lost or you forgot to start it...)

- \*\*\*\* Do files \*\*\*\*
  - Click on Do File Editor button (to the right of Data Editor button)
  - New window
  - You can copy and paste any commands in your command Review box (to the left).
  - You can type commands directly here.
  - After typing commands here, you can also run the commands from here (instead of typing them in the command line)



# THANKS!

Questions about Stata can always be asked at the ERL during our walk-in hours.

- Mon 10-8
- Tues 10-8
- Wed 10-8
- Thurs 10-6
- Fri 10-6
- Sun 1-7
- [ERL.barnard.edu/calendar](http://ERL.barnard.edu/calendar)