Recitation 3: Get help

McCourt School of Public Policy, Georgetown University

Key Ideas:

- Help files
- Minimal reproducible example

Often the fastest answer is the one you find yourself

- It's important to try to answer your own questions first using resources such as:
    - Stata help files
    - Stata pdf manuals
    - www.statalist.org
    - Stata youtube channel: www.youtube.com/user/statacorp
    - Google
- If the answer to your question is in the help file or the top hit on Google, the answer to your question will be, "Read documentation" or "Google it"

Help Files:

- We have used several complicated commands with options and if statements.
- You do not have to memorize these commands.
- Every command has a help page to tell you how to use it.

Example 1: count

`help count`

- Overview of help page, including `Also See`
- Explain syntax elements relevant to count command:
    - **bold** - type as is
    - *italics* - replace with your own variables or expression
    - [brackets] - optional
    - underline - shortest abbreviation
    - blue - hyperlinks

```
clear
sysuse auto
count if foreign==1
cou    if price > 5000
coun   in 1/5
```

Example 2: list

- Look at another help page: list

  ```
  help list
  ```

- Two new elements: [varlist]' and[, options]'

- Both are optional, so the command works fine without either one.

- list the observations that we counted previously:

  ```
  list if foreign==1
  l     if price > 5000
  li    in 1/5
  ```

- in 1/l will list all of the elements

- l refers to the last element in the list

- you can also list from the end of the list using the - sign

```
list in -2/l
```

varlist'* Follow link tovarlist' help page

- Don't worry about factor variables and time-series operators for now

- Examples of varlist

  ```
  list make in 1/5
  list make price mpg in 1/5
  list make-mpg in 1/5
  list t* in 1/5
  ```

`options`

- a list of one or more options can follow the comma

- you only need one comma, even if there are multiple options

- remember, there are no commas between a list of variables

- the comma is always placed after any if or in statements

- options can be in any order

- options can be abbreviated, just like regular commands

- Examples of options

```
list make in 1/5 , noobs
list in 1/5 , abbrev(5)
list make price in 1/5 , noobs divider mean(price)
list make if price > 10000 , sepby(foreign)
list make price foreign if price > 10000 , sepby(foreign)
```

Example 3: tabstat

- notice varlist (at least one variable) is required for this command `tabstat` if foreign==1 results in an error `tabstat price mpg weight`

- The `by(varname)` option requires a single variable
`tabstat price mpg weight , by(foreign)`

- The second option `statistics(statname [...])` is more complicated

- Scroll down to the Options section of the help page for more info

```
tabstat price mpg weight , by(foreign) stats(mean sd)
tabstat price mpg weight , by(rep78) stats(mean sd max min)
```

For more help on reading help pages, see:

- Pdf documentation:
  - GSW 10: Listing data and basic command syntax
  - U 11: Language syntax
- Help pages
  - `help language`

Leaning New Commands

- Most help pages are easy to find by guessing the command name and `See Also`
  - look at the `See Also` menu from the `summarize` help page
- There are a few help pages that are worth remembering:
  1. help contents
  2. help language
  3. help operators
  4. help resources - includes link to Stata youtube channel

In Class Activity 1

Using the auto.dta data set, attempt to answer the questions below with `tabstat` command

Create a do-file with comment before and after each question. - 1. What is the median value of the mpg variable? - 2. What is the 95th percentile value of the mpg of domestic cars? - 3. Challege Question: Try reproduce the following result:

```
foreign       variable |     mean       sd       min      max      range
-----------------------+-----------------------------------------------------
Domestic         price |   6,072.4   3,097.1     3,291   15,906    12,615
                weight |   3,317.1   695.364     1,800    4,840     3,040
                   mpg |   19.8269    4.7433        12       34        22
                 rep78 |   3.02083   .837666         1        5         4
-----------------------+-----------------------------------------------------
Foreign          price |   6,384.7   2,621.9     3,748   12,990     9,242
                weight |   2,315.9   433.003     1,760    3,420     1,660
                   mpg |   24.7727   6.61119        14       41        27
                 rep78 |   4.28571   .717137         3        5         2
-----------------------------------------------------------------------------
```

```
sysuse auto.dta
* 1. What is the median value of the mpg variable?
tabstat mpg, statistics(median)
* 2. What is the 95th percentile value of the mpg of domestic cars?
tabstat mpg, by(foreign) s(p95)
* 3. Challege Question: Try reproduce the following result:
tabstat price weight mpg rep78, by(foreign) stat(mean sd min max range) ///
nototal long col(stat) format
```

Minimal reproducible example

Why?

> My stata prompted error when I try to count the missing values.

When you ask other people for help, you'll get the most useful advice if you know how to make a *minimal reproducible example*. A reproducible example is a sample of code and data that any other user can run and get the same results as you do. A minimal reproducible example is the smallest possible example that illustrates the problem.

Four parts of minimal reproducible example

A minimal reproducible example consists of the following items : - A minimal dataset, necessary to reproduce the error - The minimal (runnable) code necessary to reproduce the error, which can be run on the given dataset. - The necessary information on the Stata version and system it is run on.

Example

Provide a minimal dataset

- Use example data set in STATA

  ```
  sysuse auto
  ```

- Use a subset of data you are using. For example, draw a 10% sample from your data

  ```
  sysuse auto, clear
  sample 10
  ```

Provide minimal (runnable) code

```
* (WRONG)
sysuse auto, clear
count if rep78=.
```

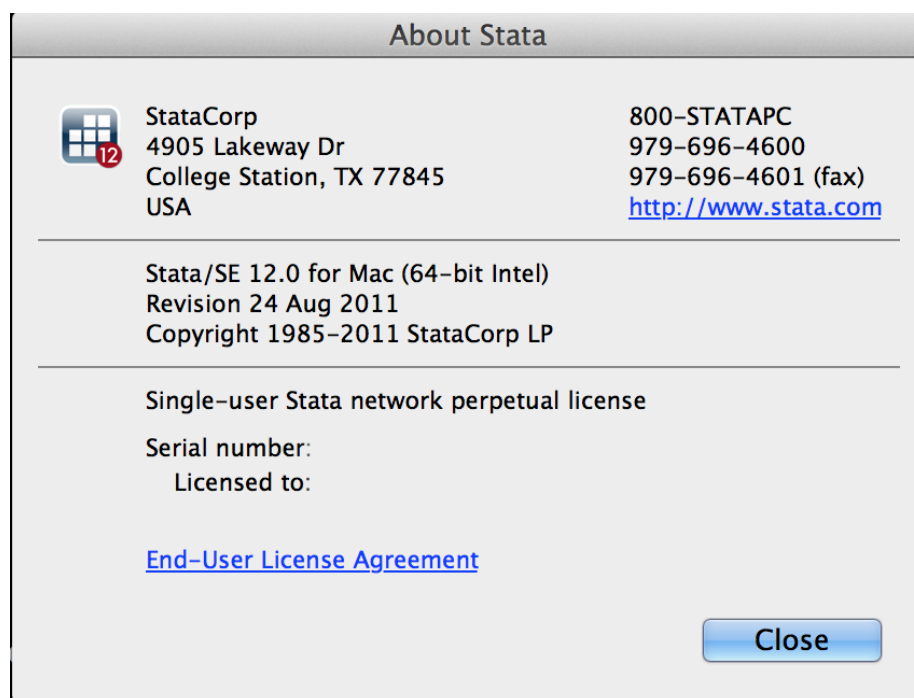### STATA version and system

- Click About Stata ...

Figure 1: Stata Version

Challenge Questions (Optional)

- When invalid commands are put into Stata, it generates an error

- The error message gives you an idea of what went wrong

- Generate the following errors

```
varlist required
variable _____ not found
option by not allowed
invalid syntax
type mismatch
```