

Social-insurance Undergraduate Research Fellowship (SURF)

Stata Lecture 1

05/15/2023



Retirement & Disability
Research Center

Outline

- Stata Learning Resources
- Stata Interface
- Run Commands
- Stata File Types
- Make Notes in a Do File
- Operators in Stata
- Remarks

Stata Learning Resources

- **Free Webinar: Ready. Set. Go Stata.**

<https://www.stata.com/training/webinar/ready-set-go-stata/>

This free, one-hour webinar will introduce you to the basics of using Stata. It is not required, but I highly encourage you to register and attend. The first session is on Tuesday, May 23rd at 2:00 PM. The registration deadline is Sunday, May 21st

- **Stata Learning Modules**

<https://stats.idre.ucla.edu/stata/modules/>

<http://data.princeton.edu/stata/>

- **Stata Video Tutorials**

<https://www.stata.com/links/video-tutorials/>

- **Statalist: Official Stata Help Forum**

<https://www.statalist.org/forums/help>

- **Stata Visual Overview for Creating Graphs**

<https://www.stata.com/support/faqs/graphics/gph/stata-graphs/>

- **Google**

Simply google your questions

- **Stata Help Files**

Stata help files can be hard to understand, but examples towards bottom are often very useful

Learning by doing (repeatedly) is really the primary way to learn well.

Stata Interface

The screenshot shows the Stata MP 13.0 interface with the following components and annotations:

- Past commands appear here:** Points to the Review window on the left, which lists a sequence of commands from 1 to 8.
- Results are displayed here:** Points to the main command window, which shows the output of the `regress` command, including a table of sums of squares and a table of coefficients.
- Variable list appears here:** Points to the Variables window on the right, which lists all variables in the dataset with their labels.
- Data properties appear here:** Points to the Properties window on the right, which shows details for the selected variable 'make'.
- Current working directory appears here:** Points to the status bar at the bottom left, which shows the path `C:\Users\stata\Documents`.
- Commands are typed here:** Points to the Command window at the bottom, which shows the command `predict yhat`.
- Current log status appears here:** Points to the status bar at the bottom right, which shows `log on (rmc)`.
- Command log status appears here:** Points to the status bar at the bottom right, which shows `cmdlog on`.

Command Window Output:

```
. log using demonstration
. cmdlog using demonstrati...
. sysuse auto
. summarize
. generate gp100m = 100/m... 199
. generate gp100m = 100/m...
. regress gp100m weight
. predict yhat
```

Regression Results Table:

Source	SS	df	MS	Number of obs = 74		
Model	87.2964969	1	87.2964969	F(1, 72) =	194.71	
Residual	32.2797639	72	.448330054	Prob > F =	0.0000	
Total	119.576261	73	1.63003097	R-squared =	0.7300	
				Adj R-squared =	0.7263	
				Root MSE =	.66957	

Coefficient Table:

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
gp100m						
weight	.001407	.0001008	13.95	0.000	.001206	.0016081
_cons	.7707669	.3142571	2.43	0.017	.1443069	1.397227

Variables Window:

Variable	Label
make	Make and Model
price	Price
mpg	Mileage (mpg)
rep78	Repair Record 1978
headroom	Headroom (in.)
trunk	Trunk space (cu. ft.)
weight	Weight (lbs.)
length	Length (in.)
turn	Turn Circle (ft.)
displacement	Displacement (cu...
gear_ratio	Gear Ratio
foreign	Car type
gp100m	Fitted values
yhat	Fitted values

Properties Window (make):

Name	make
Label	Make and Model
Type	str18
Format	%c-18s
Value Label	
Notes	

Data Window:

Filename	auto.dta
Label	1978 Automobile
Notes	
Variables	14
Observations	74

Run Commands

Three ways to run commands:

- Uses **dropdown menus** to implement commands (almost never recommended except for first time implementing some tasks)
- **Command window**: Type code directly to enter simple data exploration commands while working (not usually recommended)
- **Do-file editor**: Type code and save it as you go. Can run code line-by-line from too (almost always recommended, allows for complete replicability)

Stata File Types

- Data file

Stata saves data in its own format, which ends with a “.dta” suffix

Data in excel, CSV or other formats must be imported into Stata and saved as a “.dta” file

- Do file

Text files ending in “.do” where code is written to execute commands

- Log file

A file Stata generates while you are working that captures everything happening in the command window (like a transcript of your session). You must open and close this with commands, we usually put these in a do file at the beginning and end

Make Notes in a Do File

“Commenting out” text in a do file is a way to make notes to yourself or others who may need to make sense of your code later

- “Comment out” a single line of text by starting it with *
- “Comment out” a section by starting with “/* ” and ending with “*/ ” (this is also how you can put a comment “in line” with a command, for example

```
sysuse auto.dta /* this line opens up a data file stored in  
the software */
```
- “Comment out” notes and put them “in line” with a command with “// ”, for example

```
sysuse auto.dta // this line opens up a data file stored in  
the software
```

Commenting at the top of your do file saying what it accomplishes, what data it uses, etc. is good practice

Operators in Stata

Arithmetic	Relational	Logical
+ plus	== equal	& and
- minus	!= not equal	or
* multiply	~= not equal	! not
/ divide	> greater than	~ not
^ power/root	< less than	
	>= greater than or equal to	
	<= less than or equal to	

A double equal sign (==) is used for equality testing

Remarks

- Stata does not capitalize its *commands*
- However, *variables* are case sensitive, i.e. “wage” and “Wage” are two different variables