

# Stata Tutorial

First, go to Professor Berman's website:

<http://www.econ.ucsd.edu/~elib/120b/>

and download the coffee.dta file (right click and save it somewhere)

Second, open Stata (Go to the Start Menu and find Stata). You should see:

- The output window
- The command window
- The review of commands
- The variable list

\*Clickability: using your mouse to click on any variable in the variable list will automatically enter it into the command window. Similarly, clicking on any past command in the review of commands will also automatically enter it into the command window.

In the examples below, the basic Stata commands are in **bold** and you type them into the command window. Remember that some of the examples (such as loading the data) require you to enter your own parts of commands (e.g. a file name when loading the data).

## Loading Data

If in Stata Format (like our coffee data):

```
use "C:\Documents and Settings\Leah\My Documents\UCSD\Econ 120B -  
W 08\coffee08.dta", clear
```

\*\* Remember that you must include quotation marks ("s) around your file name

If in Excel or text format (i.e. notepad):

```
insheet using "C:\Documents and Settings\Leah\My  
Documents\UCSD\Econ 120B - W 08\coffee.txt"
```

```
insheet using "C:\Documents and Settings\Leah\My  
Documents\UCSD\Econ 120B - W 08\coffee.csv"
```

\*\* Remember that Excel data **MUST** be saved in the .csv format ("save as", if not already in that format).

You can also enter the data manually (or copy and paste from Excel or almost any format) by typing

```
edit
```

and then simply enter or paste your data into the columns and close the window when you finish. If you choose this option, Stata will automatically name your variables by column, e.g. the variable first column is “var1”. To change this, you can rename your variables by typing

```
rename var1 price
```

## Inspecting and Describing Data

To see the actual observations in the output window:

```
list price quantity
```

```
+-----+
| price  quantity |
+-----+
1. |      0         160 |
2. |     .4          35 |
3. |     .5          35 |
4. |    .75          23 |
5. |     .9          18 |
+-----+
6. |     1           16 |
7. |    1.25          7 |
8. |    1.4           6 |
9. |    1.5           5 |
10. |    1.6           4 |
+-----+
11. |    1.75          4 |
12. |     2           4 |
13. |    2.1           2 |
14. |    2.25          1 |
15. |    2.5           0 |
+-----+
```

or if there are a lot of observations:

```
browse price quantity
```

which opens a new window displaying the data you specify above

To see summary (descriptive statistics like the number of observations, the sample average and standard deviation, etc.) of the data.

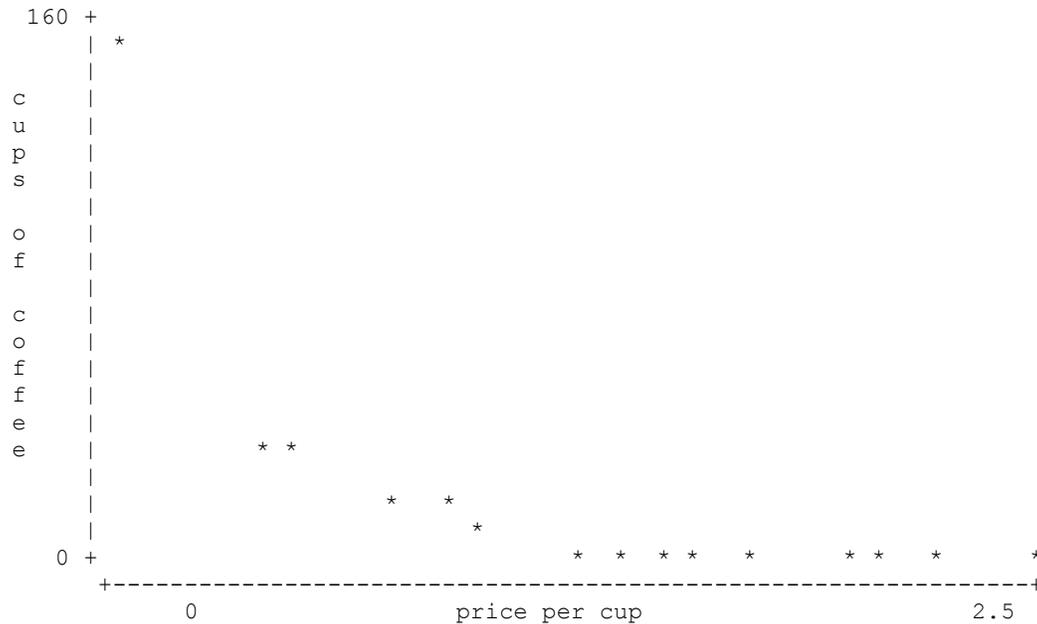
```
summarize price quantity
```

Variable	Obs	Mean	Std. Dev.	Min	Max
price	15	1.326667	.7318925	0	2.5
quantity	15	21.33333	40.08503	0	160

To get a basic plot of the data in the output window:

```
plot quantity price
```

\*\*Write the “y” variable first, then the “x” variable.



## Generating new variables

Example: generate new variables that are the log of price and log of quantity. First, write “generate” then give the new variable a name and then describe how to construct it.

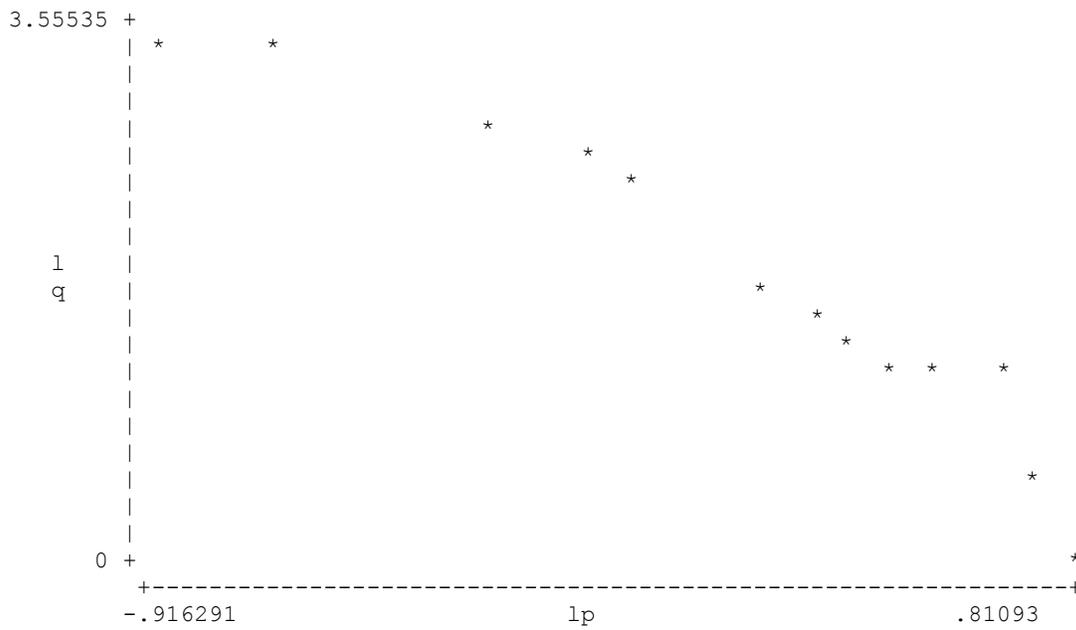
```
generate lp=log(price)  
(1 missing value generated)
```

\*\* Stata reports 1 missing value because  $\log(0)$  is undefined and we have an observation at  $p=0$ . This shows up as a “.” in Stata.

To see this, simply view the data using “browse” or “list”.

```
generate lq=log(quantity)  
(1 missing value generated)
```

```
plot lq lp
```



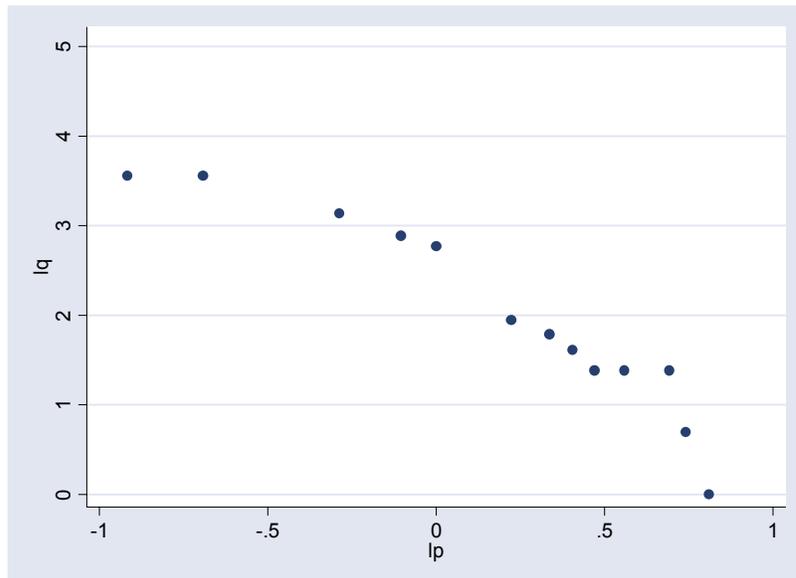
Notice how the points now seem to form a straight line!

## Graphing Data

We've already seen "plot" but if we want the graph in a separate window (e.g. to print easily for your homework assignment), we need to do one of the following:

```
gr7 lq lp
```

```
scatter lq lp
```



```
line quantity price
```

```
twoway connect quantity price
```

\*\*If you want to add an extra dependent “y” variable, then just add it in before the “x”.

To print any of these graphs, right click on the graph window and select “Print”.

\*\*Helpful resource for Stata graphs:

<http://www.stata.com/support/faqs/graphics/gph/statagraphs.html>

## Basic Regression

To run a basic OLS regression:

```
regress lq lp, robust
```

```
Regression with robust standard errors
```

Number of obs =	13
F( 1, 11) =	57.94
Prob > F	= 0.0000
R-squared	= 0.8962
Root MSE	= .37065

```
-----+-----
```

	lq	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
lp		-1.909033	.2508017	-7.61	0.000	-2.461044	-1.357023
_cons		2.337012	.1076816	21.70	0.000	2.100006	2.574017

```
-----+-----
```

\*\*Caution! The 95% confidence interval reported here assumes that we can use the Central Limit Theorem. But remember, the CLT only applies as the sample size “N” goes to infinity (or for practical matters, gets really “big”). Here, we have N=15, which is fairly small, so *these confidence intervals are invalid*.

To use your regression (the last one you’ve run) to predict values of your “y” variable (in this case, lq):

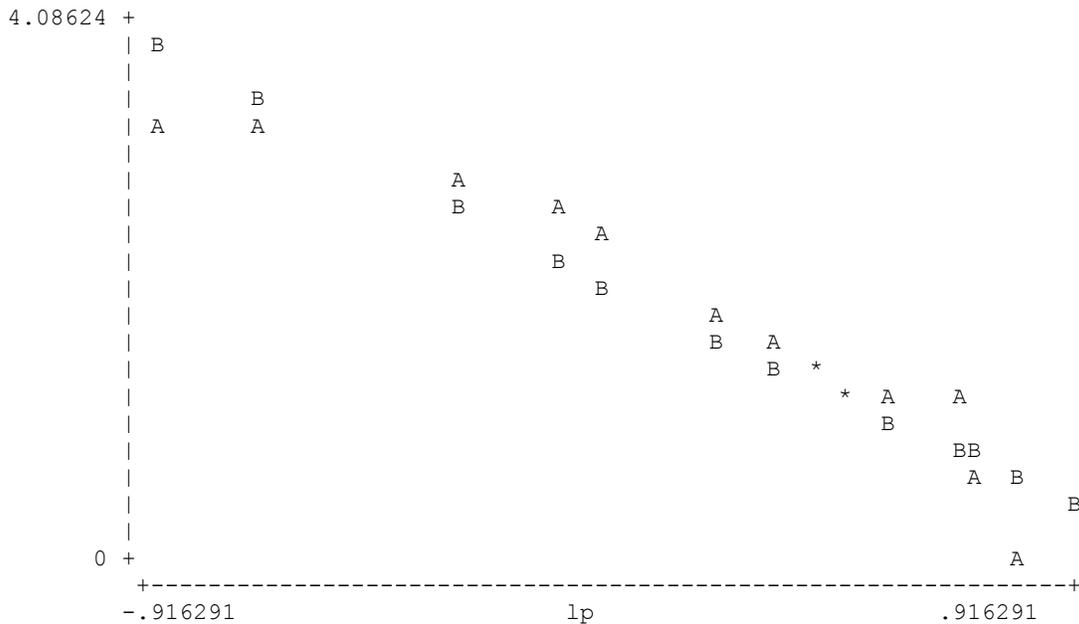
```
predict lqhat  
(option xb assumed; fitted values)  
(1 missing value generated)
```

\*\*Don’t forget to give the new predicted values a name (here, I’ve named it “lqhat”) – they’re a new variable!

```
browse lp lq lqhat
```

(or use **list**, your choice)

```
plot lq lqhat lp
```



```
scatter lq lqhat lp
```

(Or choose any graphing option)

**Remember, Stata doesn't save any of your output or the work you do (like running regressions, plotting graphs, etc.)** so if you want proof of your work, either print it or create a "do file" or a "log file." You can learn about these commands and how to perform them by checking out the links below.

**Other helpful (but sometimes advanced) Stata tutorials and guides:**

<http://www.princeton.edu/~erp/stata/main.html>

<http://www.ats.ucla.edu/stat/stata/>

\*particularly the "Stata Starter Kit"

(<http://www.ats.ucla.edu/stat/stata/sk/default.htm>) which includes a link to class notes with "movies" to help guide you through Stata basics.

The easiest way to find help for a specific command is to type "Stata help" and then whatever command you need help with into Google. The official Stata help pages will come up and are usually pretty helpful and include examples.