

STATA: A Brief Introduction to using Stata with MS Windows

A. Colin Cameron, Dept. of Economics, Univ. of Calif. - Davis

This January 2009 help sheet gives information on

- Data Sets in Stata
- Interactive Use
- Reading in a Stata dataset
- Reading in a Non-Stata dataset (a csv file)
- Summary statistics
- Linear Regression
- Two-way scatterplot with fitted regression line
- Stata do-file (A Script or program or Batch File)
- Help in Stata

STATA ACCESS AT U.C.-DAVIS

Some but not all UCD computer labs have Stata.

Schedules are available at <http://clm.ucdavis.edu/rooms/>

You need a campus computing account: <https://computingaccounts.ucdavis.edu/cgi-bin/services/index.cgi>

DATA SETS IN STATA

Stata stores data in a special format that cannot be read by other programs.

Stata data files have extension **.dta**

Stata can read data in several other formats.

A standard format is a comma-separated values file with extension **.csv** (which can be created by Excel for example).

INTERACTIVE USE

In interactive use we use a graphical-user interface and select commands from appropriate menus and dialog boxes.

This is similar to using Excel.

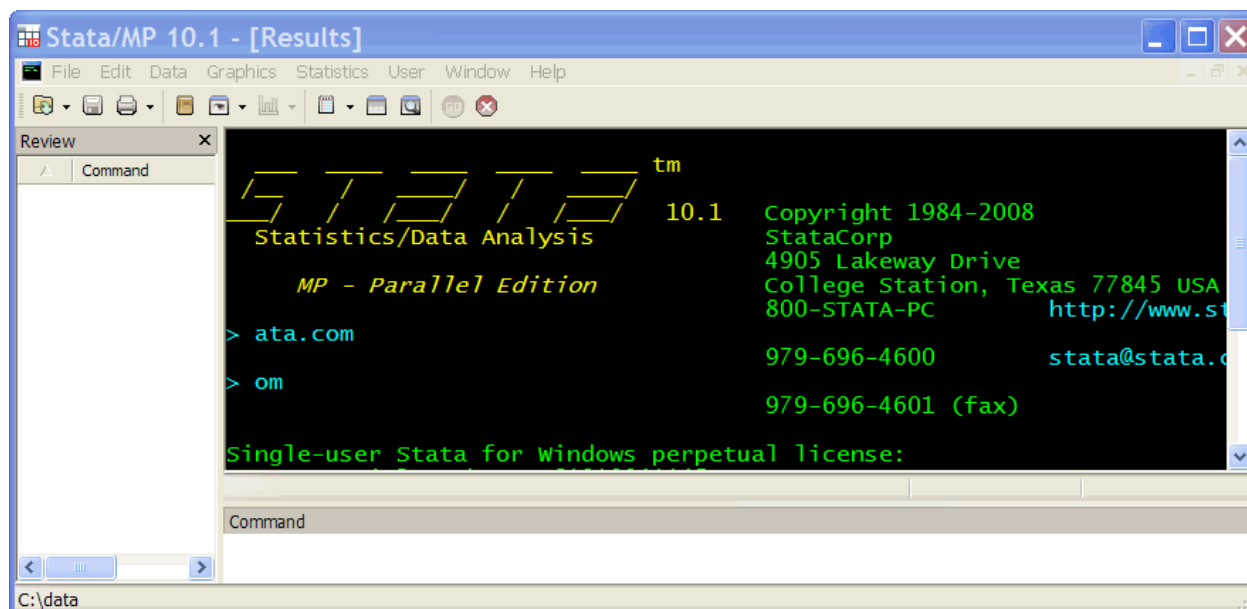
[Additionally one can combine commands in a file and execute the file.

This faster method for more experienced users is presented at the end of this file].

Interactive use can be initiated in several ways

- Click on the Stata icon in MS windows
- Click on the name of a file containing a Stata dataset (with extension **.dta**)
- Click on the name of a file containing a Stata do-file (with extension **.do**)

We do the first of these here. It yields:



Commands can be entered using the menus and consequent dialog boxes at the top.

Or commands can be typed in the Command line at the bottom.

READING IN A STATA DATA SET

Consider data in the Stata data file [carsdata.dta](#)

Here we suppose the file is in directory C:\stata (so the file is C:\stata\carsdata.dta)

1. The simplest method is in Windows go to the directory with file carsdata.dta and double-click on carsdata.dta
This initiates Stata and opens the data file.
2. Alternatively start STATA in Windows.

In the command line give the commands

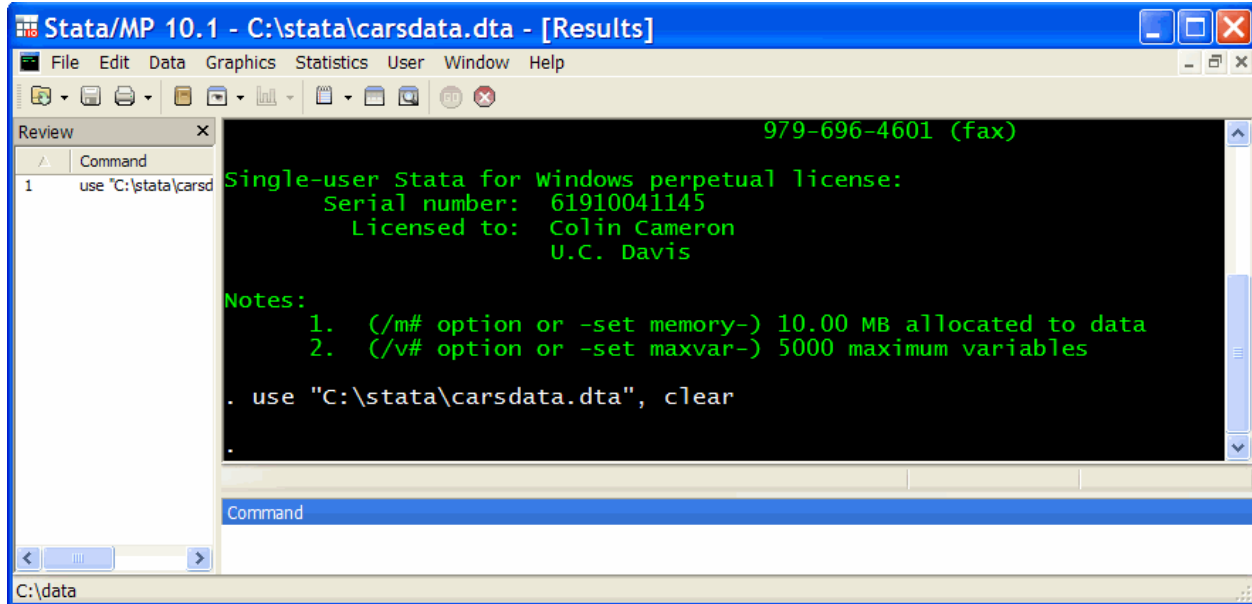
```
cd C:\stata
use carsdata.dta
or even more simply give the command
use "C:\stata\carsdata.dta"
```

3. Alternatively start STATA in Windows.

Use the File Menu and the Open submenu and browse to find the file and click on the file.

For more details see statareadinstatadataset.html

In all cases we obtain



READING IN A NON-STATA DATA SET: A CSV FILE

Stata can read in some other types of data file than a Stata dataset.

It cannot read in an Excel spreadsheet (with extension **.xls** or **.xlsx**).

A standard alternative format is a comma-separated file or comma-delimited file (with extension **.csv**).

For example in Excel an Excel worksheet can be saved as a **.csv** file.

An example is file [carsdata.csv](#)

Start STATA in Windows.

In the command line give the commands

```
cd C:\stata
insheet using carsdata.csv
or even more simply give the command
insheet using "C:\stata\carsdata.dta"
```

Alternatively Use the File Menu and the Import submenu.

Choose ASCII data created by a data sheet

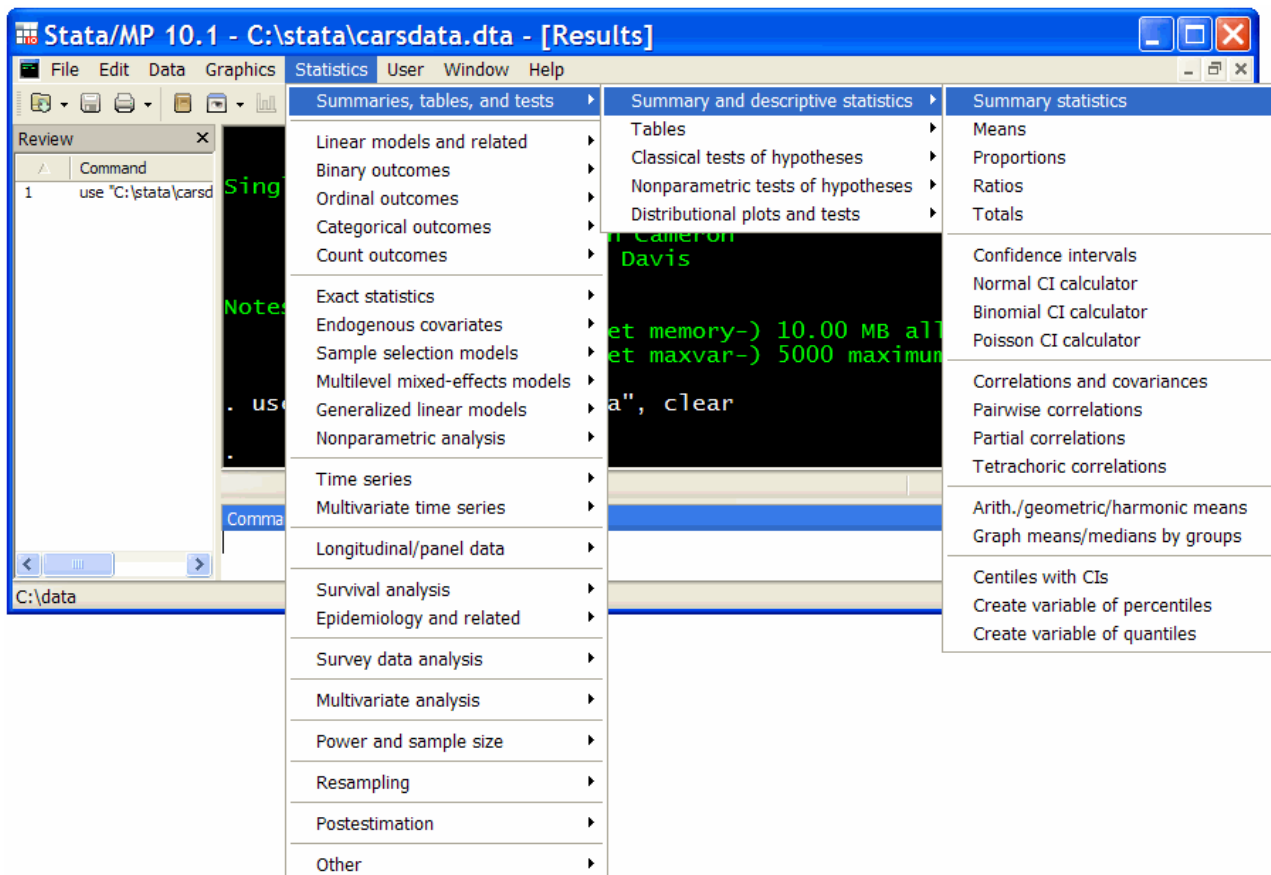
And browse to find the **.csv** file and click on the file.

SUMMARY STATISTICS

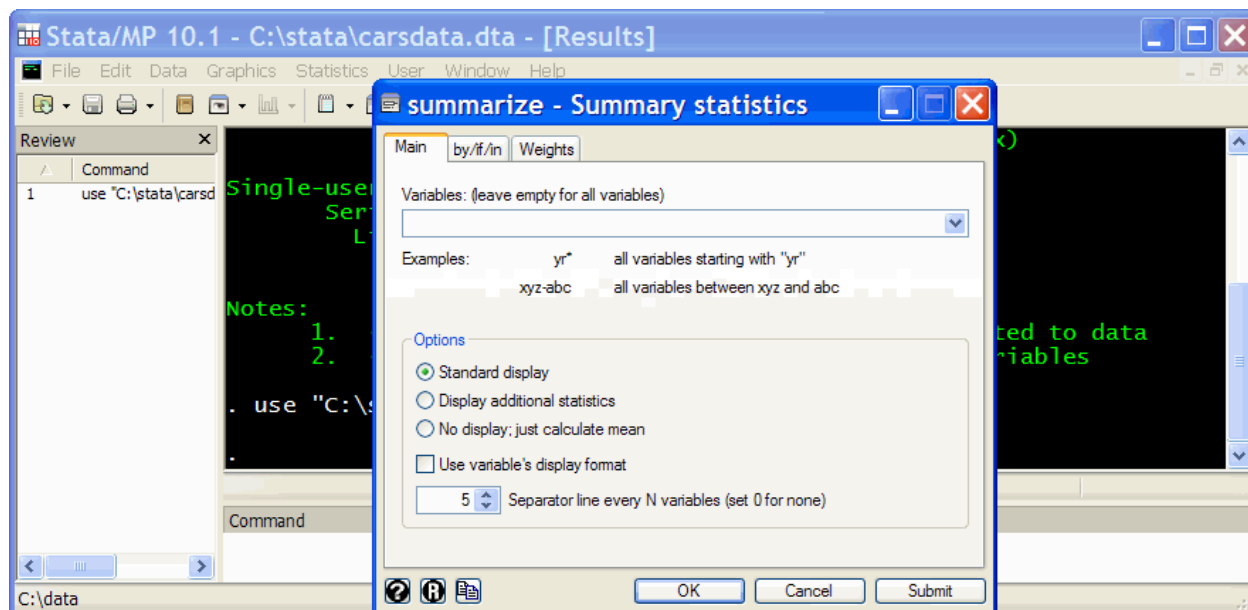
To obtain summary statistics we can simply type in the command line

```
summarize
and hit <enter>.
```

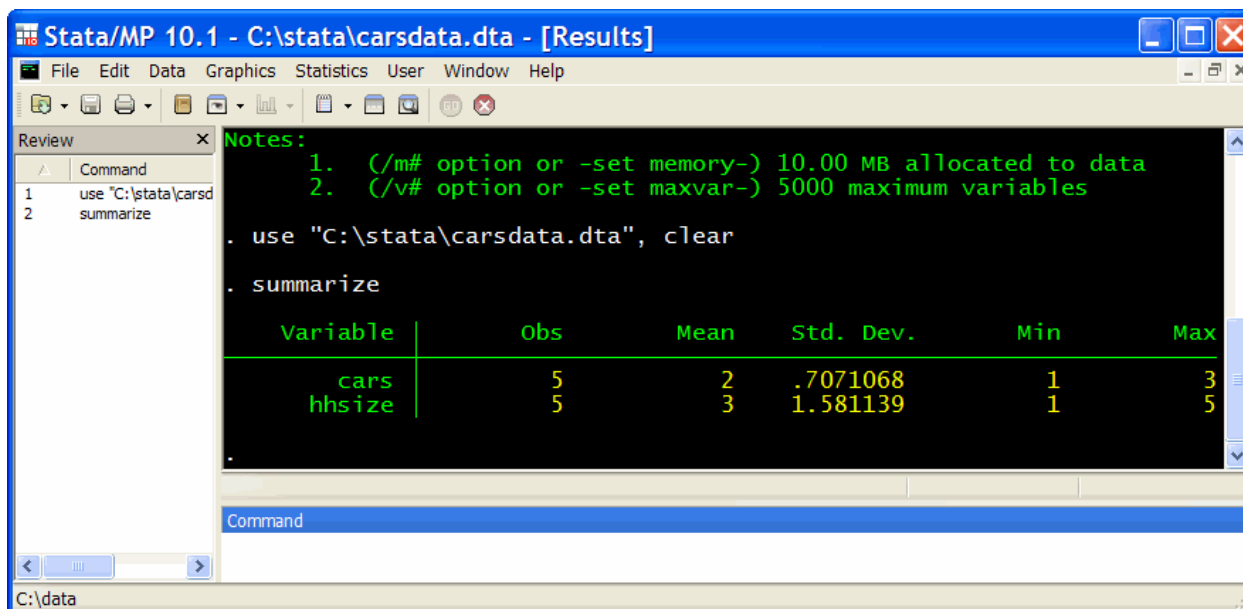
Alternatively we can use the Stata Statistics menu and subsequent submenus:



Then hit on summary statistics to get:



To obtain summary statistics for all variables simply hit the OK button.
This yields

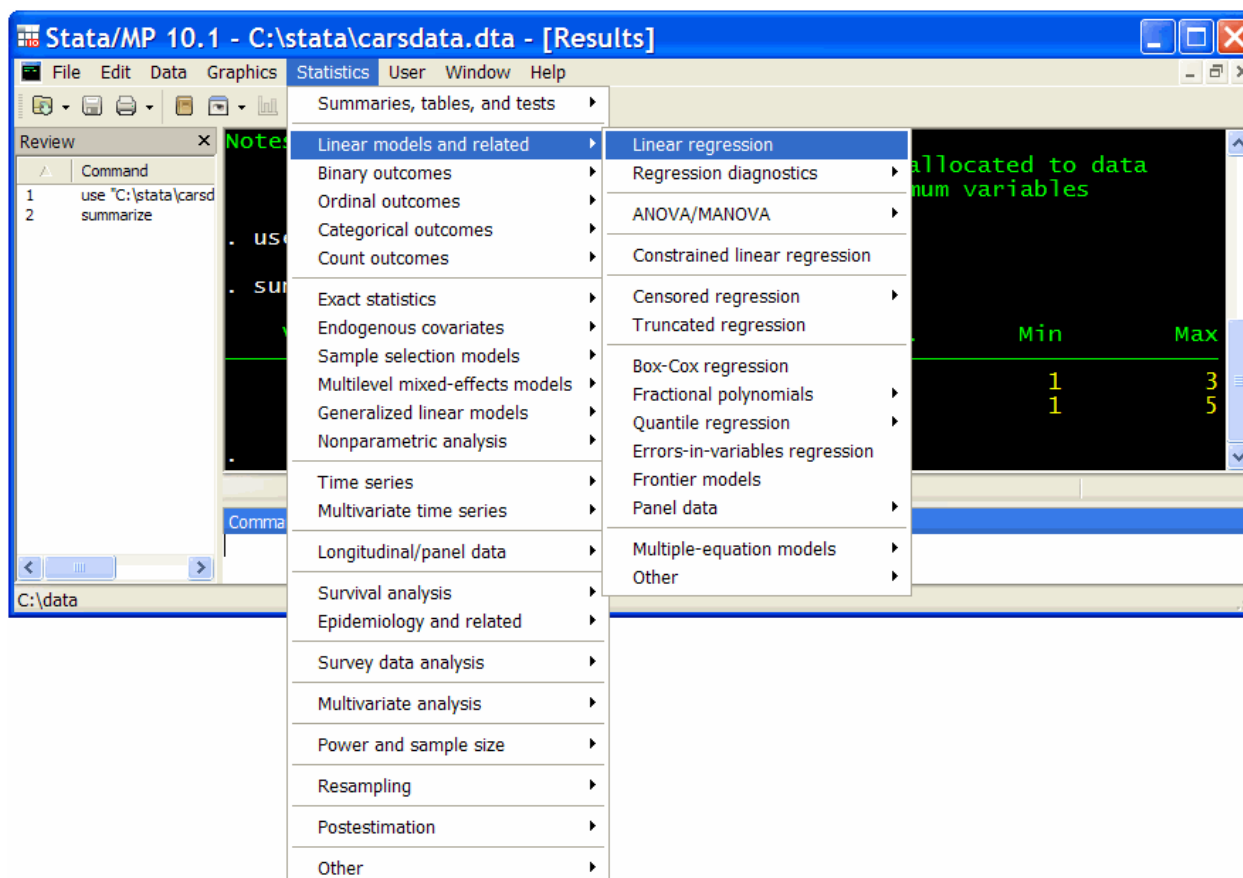


There are five observations on two variables: cars and hhsize. Summary statistics provided are the mean, standard deviation, minimum and maximum. Additional statistics would have been displayed if we had checked Display additional statistics.

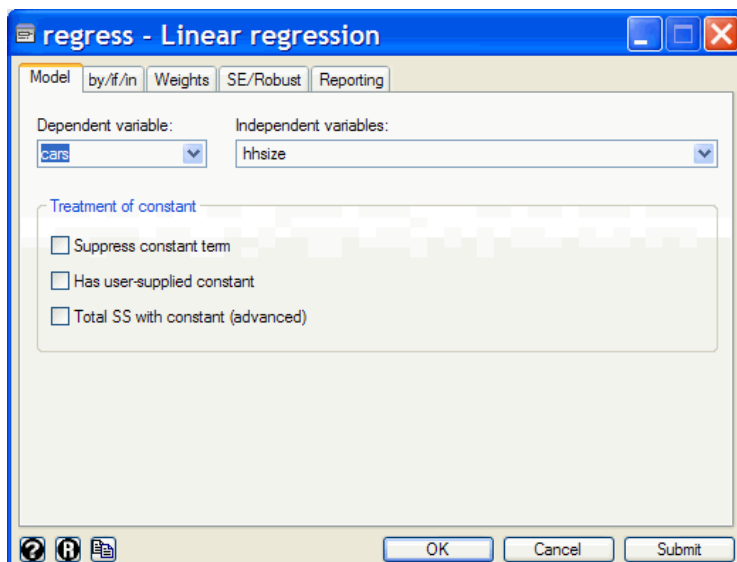
LINEAR REGRESSION

To regress variable cars on variable hhsize simply type in the command line
`regress cars hhsize`
 and hit <enter>.

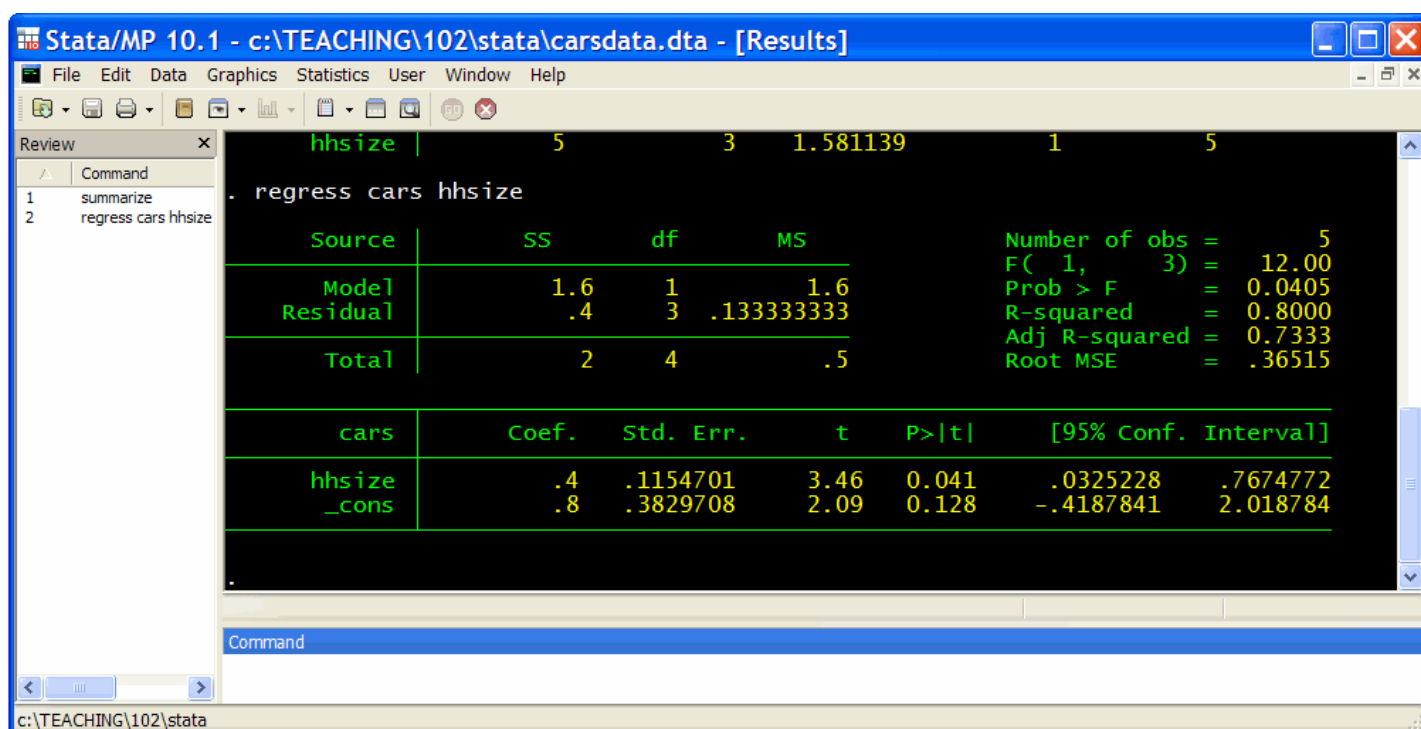
Alternatively we can use the Stata Statistics menu and subsequent submenus:



Then choosing Linear Regression yields a dialog box that we fill out as follows:



Hit OK (or directly give command `regress cars hhsz`) yields output

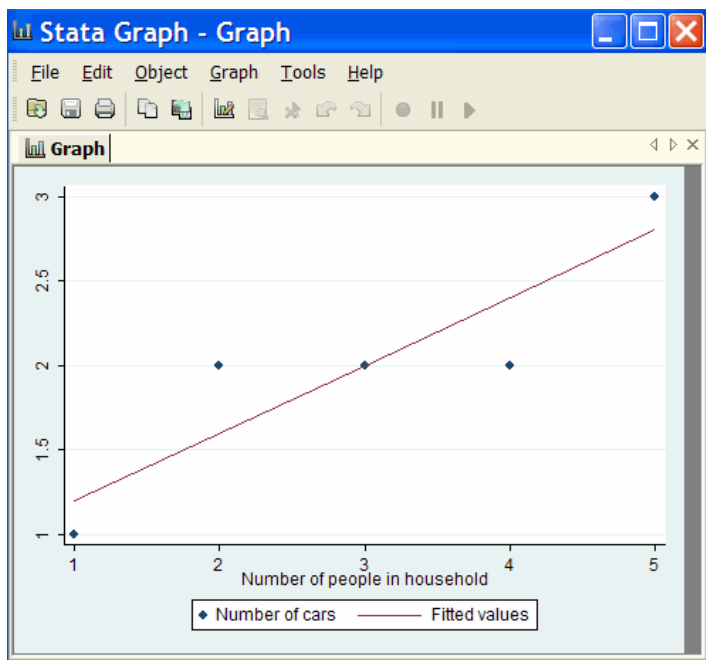


The estimated regression line is

$$\text{cars} = 0.8 + 0.4 \cdot \text{hhsz}$$

TWOWAY SCATTERPLOT WITH FITTED REGRESSION LINE

This can be obtained using the command
`twoway (scatter cars hhsz) (lfit cars hhsz)`



Alternatively use the Graphics menu and the Twoway Graph (scatter, line, etc.) submenu.

STATA DO-FILE (A Script or program or Batch File)

Stata commands can be combined in a text file with extension **.do** called a do-file.

The file [carsdata.do](#) has the following text

```
* Stata do-file carsdata.do written January 2009
* Create a text log file that stores the results
log using carsdata.txt, text replace
* Read in the Stata data set carsdata.dta
use carsdata.dta
* Describe the variables in the data set
describe
* List the dataset
list
* Provide summary statistics of the variables in the data set
summarize
* Provide an X,Y scatterplot with a regression line
twoway (scatter cars hhsz) (lfit cars hhsz)
* Save the preceding graph in a file in PNG (portable networks graphic) format
graph export carsdata.png
* Regress cars on hhsz
regress cars hhsz
```

The lines beginning with * are explanatory comments that are ignored by Stata.

To run this do-file simply click in Windows on filename carsdata.do
This file needs to be in the same directory as file carsdata.dta

Alternatively, start Stata, give command `cd C:\stata` (if file carsdata.do and carsdata.dta are in directory C:\Stata)
and then give command
`do carsdata.do`

The program does the preceding analysis.
Results are put in the text file [carsdata.txt](#)

HELP IN STATA

Stata provides extensive documentation on-line.

For example, to obtain help on the command `summarize`, in the command line type
`help summarize`

Alternatively use the Help menu

For further information on how to use STATA go to
<http://cameron.econ.ucdavis.edu/stata/stata.html>