

Getting to know Stata

Stata is a general-purpose statistical software package developed by StataCorp for data manipulation, visualization, statistics, and automated reporting. It is used by researchers in many fields, including economics, sociology, political science, biomedicine, and epidemiology. When you open your Stata, the interface looks like the following one in Figure 3 more or less (depends on your version). The official website of Stata has a series of video tutorials for you to become familiar with its interface. <https://www.stata.com/links/video-tutorials/>. You could watch it to learn how to use Stata if you prefer watching a video rather than reading instructions in this section.

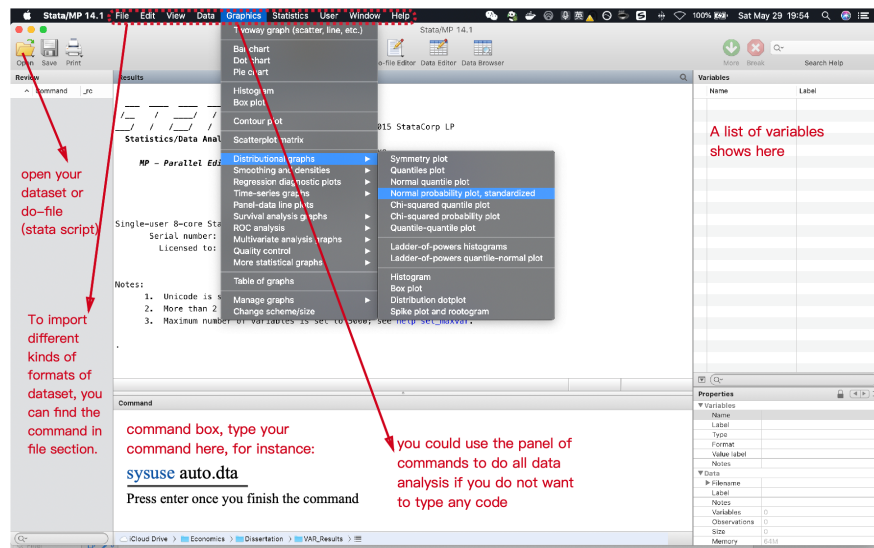


Figure 3: Stata Interface

Once you install Stata and open it, then you could try to type the following code.

```
1 sysuse auto.dta
2 summarize
```

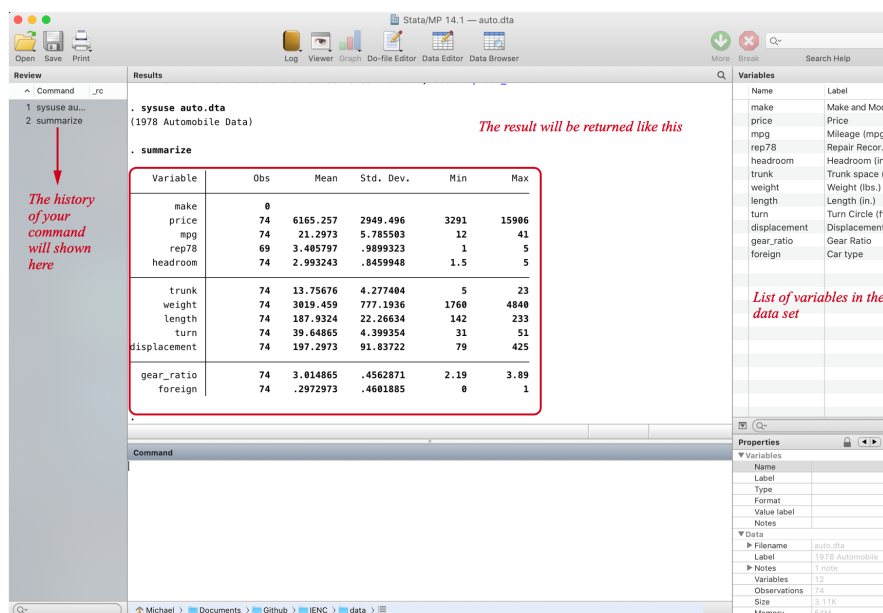


Figure 4: Stata results

The command or syntax of Stata shares the same grammar, knowing this common syntax will help you to get used to it. You *do not need to memorize the command of Stata*, but understanding the grammar of syntax is helpful. If you do not use the correct syntax you will get an error message. The basic command syntax is almost always of the following form:

```
1 command [varlist] [if] [in] [, options]
```

Now, let's learn to install a program in Stata. Type the following code in your command area and follow the instructions in Figure 5.

```
1 search heads
```

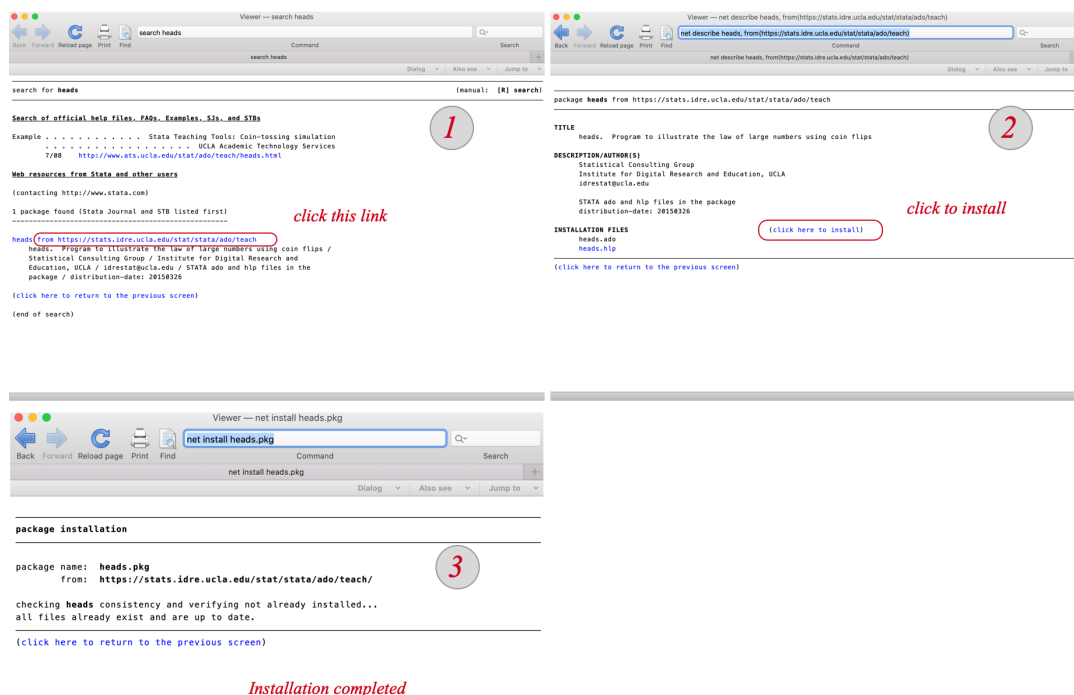


Figure 5: Stata program installation

The heads program we installed is for running simulation of flipping coins. After installing this program, then you could try to simulate the flipping of coins in Stata².

```
1 heads, flips(10)
2 heads, flips(50)
3 heads, flips(50) prob(0.3)
```

The command `heads, flips(50) prob(0.3)` runs 50 times simulation of flipping coins with the probability of head is 0.3, $P(\text{head}) = 0.3$, and it returns the average number of heads and plot it in a graph. Since the **model** of the world in this experiment assumes $P(\text{head}) = 0.3$, which means the chance of getting a head is *three times out of ten times*. Since we run 50 times simulation, then the average should be around:

$$\mathbb{P}(\text{head}) = \frac{15}{50} = \frac{3 \times 5}{10 \times 5} = 0.3$$

Every time when you run the simulation via command `heads, flips(50)`, the result will vary as each experiment is random in Stata.

I hope now you would feel comfortable to run a command in Stata. If not, please do not worry as we will run all essential code together step by step in our tutorial and I will answer to your questions.

²I would like to remind you, you need to type `enter` after each line of command. If you want to copy and past code from the box, you have to copy it line by line. Later I would give you the `do-file`, which includes all the code and you could copy, past and run code easily.