

Stata Markdown Tutorial

Cyrus Samii

January 2019

Overview

Here are some notes and examples for using the Stata Markdown package. The Stata Markdown package was written by German Rodriguez. These notes offer some basic guidance on using the package. For instructions on installation and dependencies, refer to the Stata Markdown website:

<https://data.princeton.edu/stata/markdown/>

I give examples of some things we might want to do in social science related projects.

Markdown

Markdown is a simple markup language that, through Pandoc, can be rendered in a variety of formats, including pdf (via tex), html, or docx.

If you are used to writing latex or html, then markdown will be easy, since it admits a lot of the syntax used in those languages.

There are lots of cheatsheets out there, such as:

<https://github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet>

Lots of things are done very simply in Markdown. E.g., here is a numbered list:

1. Foo
2. Foo 2
3. Foo 3

The header of this document is a YAML header for Markdown, which contains meta instructions for the Markdown->Pandoc compilation. I am rendering this document in PDF through Tex, and so you will see that in my YAML header I have included some Tex instructions.

Workflow

The way I work is to type into this document and then compile by running the requisite commands that I have put into a separate .do file called “stata-markdown-example-do.do”. That way, I can load the various compilation options (that is, the options to the `markstat` function in a way that I can easily recall them later. Using the `do` button in the Stata .do file editor gives me one button compilation. I also have my commands to set the working directory and also load in dependencies (e.g., the `stata.sty` file needed to compile to PDF).

I may also have another Stata .do file that I use as a scratch pad for working out the kinks of the Stata code that I then insert as code chunks into this document.

“Simple Script” Example

Here we replicate the simple example from German Rodriguez’s “Simple Script” example, tweaking a few things to make some additional points.

Stata code appears below in “chunks” that are demarcated in the following manner:

For code chunks that you want to appear in the rendered document:

```
```{s}  
[code here]
```
```

For code chunks that you DO NOT want to appear in the rendered document:

```
```{s/  
[code here]
```
```

Now we can proceed with the simple example. First read in the fuel efficiency data that is shipped with Stata:

```
. sysuse auto, clear  
(1978 Automobile Data)
```

To study how fuel efficiency depends on weight it is useful to transform the dependent variable from “miles per gallon” to “gallons per 100 miles”:

```
. gen gphm = 100/mpg
```

We can then plot the relationship. We will run this code in a manner that is not echoed in the resulting output file (PDF, docx, etc.).

Regression table with `esttab`

Something that we frequently need to do is to report regression tables. We can use the `esttab` function in Stata and insert its output here:

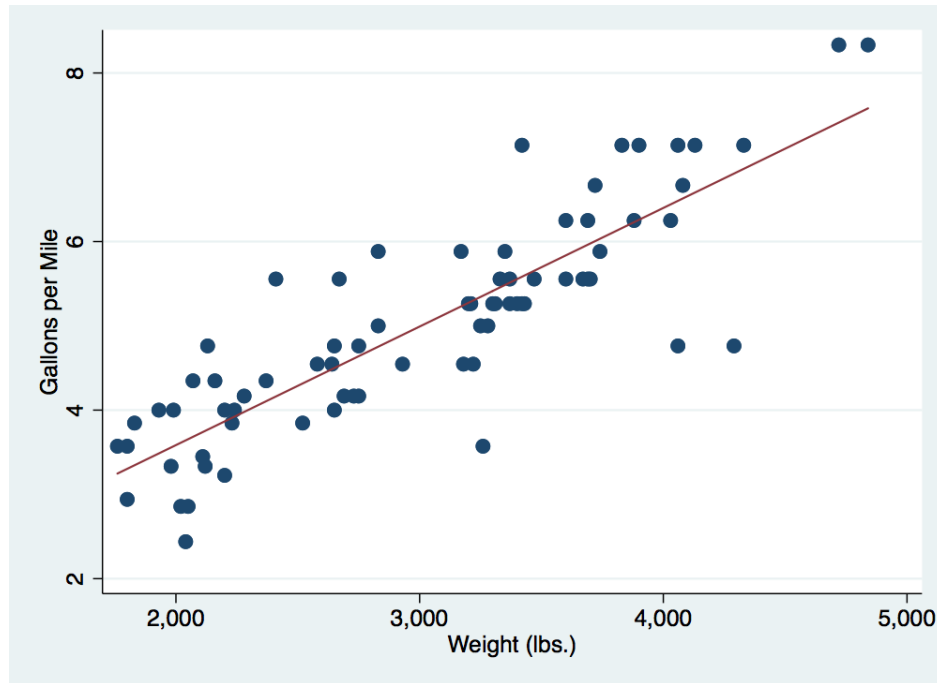


Figure 1: Fuel Efficiency

| | (1) |
|--|-------------------|
| | Gall/100 mi. |
| Weight (lbs.) | 0.00***
(0.00) |
| Constant | 0.77*
(0.33) |
| Observations | 74 |
| r2 | 0.73 |
| Standard errors in parentheses | |
| * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ | |

(If you look at the Stata Markdown .stmd file, you will see that I used tex commands to insert the regression table and center it.)

Summary stats with `esttab`

Sometimes we want nice summary stats tables. Here is an example:

Table 1: Summary Stats.

| | Min. | Mean | Med. | Max. | SD | Obs. |
|---------------|---------|---------|---------|----------|---------|-------|
| Price | 3291.00 | 6165.26 | 5006.50 | 15906.00 | 2949.50 | 74.00 |
| Weight (lbs.) | 1760.00 | 3019.46 | 3190.00 | 4840.00 | 777.19 | 74.00 |

Loop with display

```
2. sum `varUp`, detail
3. hist `varUp`
4. }
```

| Price | | | | |
|-------------------------------------|----------|---------|-------------|----------|
| Percentiles | Smallest | | | |
| 1% | 3291 | 3291 | | |
| 5% | 3748 | 3299 | | |
| 10% | 3895 | 3667 | Obs | 74 |
| 25% | 4195 | 3748 | Sum of Wgt. | 74 |
| 50% | 5006.5 | | Mean | 6165.257 |
| | | Largest | Std. Dev. | 2949.496 |
| 75% | 6342 | 13466 | | |
| 90% | 11385 | 13594 | Variance | 8699526 |
| 95% | 13466 | 14500 | Skewness | 1.653434 |
| 99% | 15906 | 15906 | Kurtosis | 4.819188 |
| (bin=8, start=3291, width=1576.875) | | | | |
| Weight (lbs.) | | | | |
| Percentiles | Smallest | | | |
| 1% | 1760 | 1760 | | |
| 5% | 1830 | 1800 | | |
| 10% | 2020 | 1800 | Obs | 74 |
| 25% | 2240 | 1830 | Sum of Wgt. | 74 |
| 50% | 3190 | | Mean | 3019.459 |
| | | Largest | Std. Dev. | 777.1936 |
| 75% | 3600 | 4290 | | |
| 90% | 4060 | 4330 | Variance | 604029.8 |
| 95% | 4290 | 4720 | Skewness | .1481164 |
| 99% | 4840 | 4840 | Kurtosis | 2.118403 |
| (bin=8, start=1760, width=385) | | | | |