

Reproducible Research Tools for Stata

January 2017

Boriana Pratt
Princeton University



Tools for Stata

Weaver (uses Markdoc) – created by E.F. Haghish

http://www.stata.com/meeting/italy14/abstracts/materials/it14_haghish.pdf

Install weaver:

```
ssc install weaver
```

Or (better) use:

```
net install weaver, replace from("https://raw.githubusercontent.com/haghish/weaver/master/")
```

Install MarkDoc:

```
ssc install markdoc
```

And statax

```
ssc install statax
```

To make pdf or use mathematical notations also need to install:

wkhtmltopdf (<http://wkhtmltopdf.org/>) and MathJax (<https://www.mathjax.org/>)

Do this from within Stata (select “Install automatically”)

Tools for Stata

Weaver has two sets of commands:

weave commands:

<code>weave using <i>filename</i></code>	<i>create html or LaTeX log file</i>
<code>weave close</code>	close the Weaver log
<code>Weave query</code>	report the status of the Weaver log
<code>weave on off</code>	temporarily deactivate or reactivate the Weaver log
<code>weave pdf</code>	render pdf from the Weaver log anytime

annotation commands:

<code>div [code] [result] <i>command</i></code>	display Stata commands and outputs in the Weaver log
<code>txt [code] <i>display_directive</i></code>	display dynamic text in the Weaver log
<code>img [<i>filename</i>]</code>	import a figure
<code>tbl (*[,*...] [\ *[,*...] [\ [...]])</code>	creating dynamic table

Documentation:

<http://www.haghish.com/statistics/stata-blog/reproducible-research/weaver.php>

Tools for Stata

Example file - example.do

```
1  weave using weave_ex1, replace date author("B. Pratt") ///
2  title("A weaver example") ///
3  summary("My first time using weaver.")
4
5  sysuse auto, clear
6
7  summarize mpg
8  scalar mpgN      = r(N)
9  scalar mpgMean   = r(mean)
10 scalar mpgSD     = r(sd)
11
12 txt "-// Basic linear regression -//"
13
14 txt "The auto.dta includes [mono]mpg[#] and [mono]weight[#] variables "
15 txt "which include data about mileage and weight of" r(N)
16 txt " observation. The variables are summarized in Table 1. "
17
18 div result summarize mpg weight
19 correlate mpg weight
20
21 txt "There is a high and negative correlation of " r(rho) ///
22 " between the [mono]mpg[#] and [mono]weight[#] variables. Figure 1 shows the fitter regres:
23 " line in the scatter plot."
24 div code graph twoway (scatter mpg weight) (lfit mpg weight)
25 graph export scatter.png, replace
26 img using scatter.png, center w(250) h(180) ///
27 title("Figure 1. Regression line fitter on mpg and weight variables")
28
29
30 weave close
31
```

Tools for Stata

example.do – after running it in Stata – output window

```
. graph export scatter.png, replace
(file scatter.png written in PNG format)

. img using scatter.png, center w(250) h(180) ///
> title("Figure 1. Regression line fitter on mpg and weight variables")

.
.
. weave close
```

```
pdf : weave\_ex1.pdf
name : weave\_ex1.html
log : C:\bpratt\projects\ReproducibleR\weave_ex1.html
log type : html
closed on : 7 Nov 2016, 14:41:30
```

Tools for Stata

The result - weave_ex1.pdf

1

Basic linear regression

The `auto.dta` includes `mpg` and `weight` variables

which include data about mileage and weight of 74

observation. The variables are summarized in Table 1.

Variable	Obs	Mean	Std. Dev.	Min	Max
<code>mpg</code>	74	21.2973	5.785503	12	41
<code>weight</code>	74	3019.459	777.1936	1760	4840

There is a high and negative correlation of -0.81 between the `mpg` and `weight` variables. Figure 1 shows the fitted regression line in the scatter plot.

```
graph twoway (scatter mpg weight) (lfit mpg weight)
```

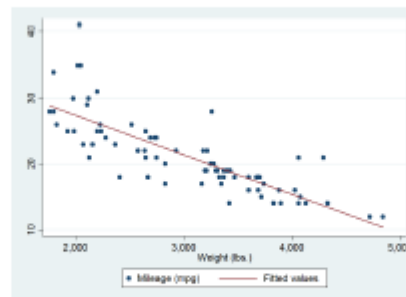


Figure 1. Regression line fitter on mpg and weight variables

Tools for Stata

Markdown – created by Germán Rodríguez

To install:

```
net from http://data.princeton.edu/stata
```

Also install `whereis` command (from the same place)

First run:

```
whereis pandoc "path_to_pandoc.exe"
```

```
whereis pdflatex "path_to_pdflatex.exe"
```

Documentation and examples here: <http://data.princeton.edu/stata/markdown>

Tools for Stata

Markdown – how it works

1. You create a .stmd file (through Stata do-file editor)
using either the simple syntax or the strict syntax

2. Then run in the Stata command window:

```
markdown using filename [, pdf mathjax strict ]
```


Tools for Stata

exMkdown.stmd – using simple syntax

```
Stata_exMkdown.stmd × Stata_exMkdown1.stmd × Untitled1.do ×
1  Stata Markdown example
2  -----
3
4  Basic linear regression
5
6      sysuse auto, clear
7
8      quietly summarize mpg
9      scalar mpgN      = r(N)
10     scalar mpgMean   = r(mean)
11     scalar mpgSD     = r(sd)
12
13
14  The auto.dta includes mpg and weight variables
15  which contain data about mileage and weight. These two variables are summarized in the table below:
16
17     summarize mpg weight
18
19  The next Stata command is not executed because it's not indented:
20
21  correlate mpg weight
22
23
24
25  Figure 1 shows the fitted regression line in the scatter plot.
26
27     graph twoway (scatter mpg weight) (lfit mpg weight)
28     graph export scatter.png, replace width(400) height(400)
29
30  ![Figure 1. Regression line fitted on mpg and weight variables](scatter.png)
31
32
```

In Stata command prompt run: `markdown using exMarkdown`

Stata Markdown example

Basic linear regression

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

```
. quietly summarize mpg  
. scalar mpgN      = r(N)  
. scalar mpgMean   = r(mean)  
. scalar mpgSD     = r(sd)
```

The auto.dta includes mpg and weight variables which contain data about mileage and weight. These two variables are summarized in the table below:

```
. summarize mpg weight
```

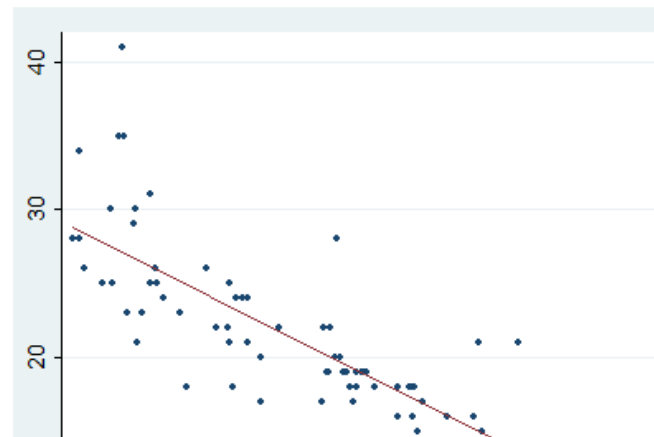
variable	Obs	Mean	Std. Dev.	Min	Max
mpg	74	21.2973	5.785503	12	41
weight	74	3019.459	777.1936	1760	4840

The next Stata command is not executed because it's not indented:

```
correlate mpg weight
```

Figure 1 shows the fitted regression line in the scatter plot.

```
. graph twoway (scatter mpg weight) (lfit mpg weight)  
. graph export scatter.png, replace width(400) height(400)  
(file scatter.png written in PNG format)
```



Tools for Stata

Let's create a `.stmd` file using the strict syntax, that:

- includes a title
- reads the auto dataset
- computes the correlation between mpg and weight and includes the value in the text
- includes a plot of mpg by weight

Then run the `stmd` file through `markdown` and produce a `pdf` file.

Let's try it...

Any questions?

Thank you