Mass production of Reports in Stata

Karl Keesman

2009

3rd Australian and New Zealand Stata Users Group Meeting

Karl Keesman Mass production of Reports in Stata

Subsection no.1.1

The Problem

Publication quality reports can take lot of time to produce especially where large numbers of tables and other information needs to be added

Subsection no.1.1

The Problem

The Australian Early Development Index has about 600 report each with 8 tables and a lot of information between text

An example of the changing information:

8.8 % of the children are considered developmentally vulnerable. The highest proportion of developmentally vulnerable children are in the Local TOWN X (27.3 %) and the lowest proportion are in TOWN V (0.0 %)

There are 29 % of children performing well. The highest proportion of children preforming well are in TOWN F (42.2 %) and lowest proportion are in TOWN Z (12.1 %)

Subsection no.1.1

The Problem

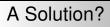
The Australian Early Development Index has about 600 report each with 8 tables and a lot of information between text

An example of the changing information:

8.8 % of the children are considered developmentally vulnerable. The highest proportion of developmentally vulnerable children are in the Local TOWN X (27.3 %) and the lowest proportion are in TOWN V (0.0 %)

There are 29 % of children performing well. The highest proportion of children preforming well are in TOWN F (42.2 %) and lowest proportion are in TOWN Z (12.1 %)

Section no.1	
Section no.2	Subsection no.2.2
Section no.3	
Section no.4	
Section no.5	Subsection no.2.5



There are some excellent user written programs: For example:

tabout

tabout is a table building program for oneway and twoway tables of frequencies and percentages, and for summary tables.

The output format can be csv, txt, html or LaTeX.

http://www.ianwatson.com.au/stata/tabout8examples.pdf

Section no.1	Subsection no.2.1
Section no.2	Subsection no.2.2
Section no.3	
Section no.4	
Section no.5	Subsection no.2.5

A Solution?

An example of tabout table:

	Ma	nufacturing origin	
	Domestic origin	Foreign origin	Tota
	(mean mpg)	(mean mpg)	(mean mpg)
Category of gear ratio			
Low	16.0		16.0
Middle	20.6	17.0	20.4
Upper	21.1	23.5	22.1
Highest	27.8	24.7	25.5
Total	20.0	23.6	21.0
Weight category			
Low	28.7	26.7	27.4
Medium	20.5	17.0	19.8
High	15.8		15.8
Total	20.0	23.6	21.0
Roominess of the vehicle			
Minimal headroom	23.4	23.0	23.4
Adequate headroom	20.2	23.6	21.4
Excellent headroom	18.1		18.1
Total	20.0	23.6	21.0

Table 11: Average fuel consumption by origin

Source: Stata Dataset. Population: All cars.

Section no.1	
Section no.2	Subsection no.2.2
Section no.3	
Section no.4	
Section no.5	Subsection no.2.5

A Solution?

estout

estout, estab etc. has a huge range of options for regression table, one way table, summarise out put etc.

The output format can be csv, rtf, txt, html or LaTeX.

http://repec.org/bocode/e/estout/

ection no.1	
ection no.2	Subsection no.2.2
ection no.3	Subsection no.2.3
ection no.4	
ection no.5	Subsection no.2.5

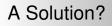


Table: Regression table

	(1)	(2)
	Price	Price
Weight (lbs.)	1.747	3.465
	(2.72)	(5.49)
Mileage (mpg)	-49.51	21.85
0 (10,	(-0.57)	(0.29)
Car type		3673.1
		(5.37)
Constant	1946.1	-5853.7
	(0.54)	(-1.73)
Observations	74	74
t statistics in nor		

t statistics in parentheses

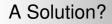
Section no.1	
Section no.2	
Section no.3	
Section no.4	Subsection no.2.
Section no.5	



Where the text and the analysis code are in the same document:

weave

ection no.1	
ection no.2	
ection no.3	
ection no.4	
ection no.5	Subsection no.2



Mass producing appendices using Stata and word processor mail merge

NASUG July 11, 2005

by: Michael Blasnik M. Blasnik & Associates, Boston, MA

Section	no.1
Section	no.2
Section	no 3
Section	

The data format for a mail merge

Identifier
data

Section no.1	
Section no.2	
Section no.3	Subsec
Section no.4	
Section no.5	

The data format for a mail merge

Identifier
data

table1,1,1	table1 $_r$ 1 $_c$ 2	table1 $_r$ 1 $_c$ 3	table1,1c41	table5 $_r2_c1$	table1,2
1	stuff	stuff	stuff	stuff	stuff

Section no.1	
Section no.2	
Section no.3	
Section no.4	
Section no.5	

The data format for a mail merge



table1,1,1	table1,1c2	table1 $_r$ 1 $_c$ 3	table1,1c41	table5 $_r2_c1$	table1,2
1	stuff	stuff	stuff	stuff	stuff

stvartable~1	stvartable~2	tablelevel~1	tablelevel~2	tablelevel~3	tablelevel~4	tableleve
Domestic	Foreign	1	2	3	4	

Clip from Stata data editor



To do this I will use Stata and Mata:

We could have used postfile but I think Mata is the logical choice, because:

Easy to manipulate matrices



To do this I will use Stata and Mata:

We could have used postfile but I think Mata is the logical choice, because:

- Easy to manipulate matrices
- Mata has lots of functions that allow you to do almost anything to the data eg. finding the max value



To do this I will use Stata and Mata:

We could have used postfile but I think Mata is the logical choice, because:

- Easy to manipulate matrices
- Mata has lots of functions that allow you to do almost anything to the data eg. finding the max value

Mata is fast



To do this I will use Stata and Mata:

We could have used postfile but I think Mata is the logical choice, because:

- Easy to manipulate matrices
- Mata has lots of functions that allow you to do almost anything to the data eg. finding the max value

Mata is fast

An example

Today we will look at how these are produced.



Assume that we wish to produce the following report in Word

Car Repair Report

It was found that highest category had cases XXXX and this was for XXXX cars with XXXX repairs per year.

		No of Repairs in 1978				
	1	2	3	4	5	
Foreign	2	8	27	9	2	
Domestic	0	0	3	9	9	



Assume that we wish to produce the following report in Word

Car Repair Report

It was found that highest category had cases XXXX and this was for XXXX cars with XXXX repairs per year.

	No of Repairs in 1978				
	1	2	3	4	5
Foreign	2	8	27	9	2
Domestic	0	0	3	9	9

The example in Stata