Producing Automated Tables using Stata

Monday
18/12/2023

Roxanne Connelly
University of Edinburgh
roxanne.connelly@ed.ac.uk

Banco De Portugal
Workshop

Materials are available here:
https://osf.io/6h7gm/

Workshop Resources

Materials are available here: https://osf.io/6h7gm/

Producing Automated Tables

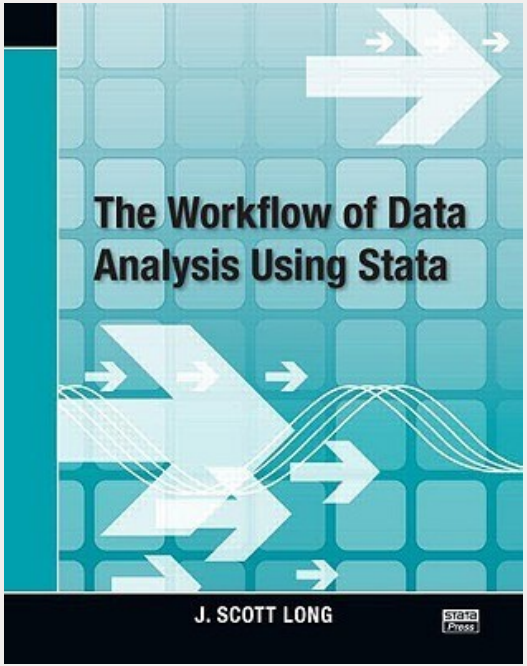
- Automation is the backbone of many procedures to promote efficiency, transparency and reproducibility.
- Automation saves time and energy and reduces the opportunity for errors.
- Copying and pasting numbers will lead to errors and waste time.

The Workflow of Data Analysis Using Stata

"The workflow involves the entire process of data analysis including planning and documenting your work, cleaning data and creating variables, producing and replicating statistical analyses, presenting findings, and archiving your work."

(Long 2009: 1)

A little planning goes a long way to ensuring that your research stays on track, and that you work effectively.



Long, J.S. (2009). The Workflow of Data Analysis Using Stata. Stata Press.

Troubleshooting

Some of these commands are new to Stata 18 (or Stata 17)

Many features have been updated since the initial release of Stata 18

```
update query
update all
```

Old(er) Commands

Community-contributed commands for creating and exporting tables:

- outreg (John Gallup)
- tabout (Ian Watson)
- estout (Ben Jann)
- outreg2 (Roy Wada)
- asdoc (Attaullah Shah)
- And many more...

Collect Commands

- The collect commands are a suite of commands new to Stata 17.
- The collect commands allow you to ‘collect’ results from any Stata command, create fully customizable table layouts, and export the table in your choice of format (or insert it automatically into a document).
- Collections contain the results from one or more commands.
- Collections also contain styles which determine the format of the table and it’s contents (e.g. headers, font styles, numeric formats).
- You can modify anything in a collection (i.e. any label or any style).
- you cannot modify values, these are produced by your commands.

StataCorp. 2023. Stata 18 Reporting Reference Manual. College Station, TX: Stata Press.

export()

- docx (Microsoft Word)
- html (HTML 5 with CSS)
- pdf (PDF)
- xlsx (Microsoft Excel 2007/2010 or newer)
- xls (Microsoft Excel 1997/2003)
- tex (LaTeX)
- smcl (SMCL)
- txt (Plain Text)
- Markdown (Markdown)

StataCorp. 2023. Stata 18 Reporting Reference Manual. College Station, TX: Stata Press.

dtable and etable

- The dtable and etable commands use the collect commands to create commonly used tables.
- The dtable command creates tables of descriptive statistics.
- The etable command creates tables of estimation results (e.g. modelling results).
- These commands can be used to create attractive and effective tables in a single step, without knowledge of the collect commands.

StataCorp. 2023. Stata 18 Reporting Reference Manual. College Station, TX: Stata Press.

StataCorp Stata Blog ‘Not Elsewhere Classified’: Reporting Blog Posts

StataCorp YouTube Playlist: Customizable Tables

Demo Outline

- Tables of Descriptive Statistics
  - The collect Commands
  - The dtable Command
- Tables of Modelling Results
  - The etable Command
- Outputting Saved Results

```
webuse nhanes2b, clear

numlabel, add

svyset psuid [pweight=finalwgt], strata(stratid)
```

Tables of Descriptive Statistics

```
webuse nhanes2b, clear

numlabel, add

svyset psuid [pweight=finalwgt], strata(stratid)
```

```
-> tabulation of sex

       Sex |      Freq.     Percent        Cum.
-----------+-----------------------------------
    1. Male |      4,915       47.48       47.48
  2. Female |      5,436       52.52      100.00
-----------+-----------------------------------
     Total |     10,351      100.00

-> tabulation of rural

     Rural |      Freq.     Percent        Cum.
-----------+-----------------------------------
  0. Urban |      6,548       63.26       63.26
  1. Rural |      3,803       36.74      100.00
-----------+-----------------------------------
     Total |     10,351      100.00

-> tabulation of agegrp

 Age group |      Freq.     Percent        Cum.
-----------+-----------------------------------
  1. 20–29 |      2,320       22.41       22.41
  2. 30–39 |      1,622       15.67       38.08
  3. 40–49 |      1,272       12.29       50.37
  4. 50–59 |      1,291       12.47       62.84
  5. 60–69 |      2,860       27.63       90.47
    6. 70+ |        986        9.53      100.00
-----------+-----------------------------------
     Total |     10,351      100.00

-> tabulation of diabetes

Diabetes status |      Freq.     Percent        Cum.
----------------+-----------------------------------
0. Not diabetic |      9,850       95.18       95.18
   1. Diabetic |        499        4.82      100.00
----------------+-----------------------------------
          Total |     10,349      100.00

.
.       summ height weight

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
      height |     10,351    167.6509    9.655916       135.5        200
      weight |     10,351    71.89752    15.35642       30.84     175.88
```

The Collect Commands

```
collect clear

table (var) (), statistic(fvfrequency sex agegrp rural) ///
                statistic(fvpercent sex agegrp rural) ///
                statistic(mean height weight) ///
                statistic(sd height weight)

collect dir

collect preview
```

	Factor-variable frequency	Factor-variable percent	Mean	Standard deviation
Sex=1. Male	4,915	47.49		
Sex=2. Female	5,434	52.51		
Age group=1. 20-29	2,320	22.42		
Age group=2. 30-39	1,621	15.66		
Age group=3. 40-49	1,271	12.28		
Age group=4. 50-59	1,291	12.47		
Age group=5. 60-69	2,860	27.64		
Age group=6. 70+	986	9.53		
Rural=0. Urban	6,547	63.26		
Rural=1. Rural	3,802	36.74		
Height (cm)			167.6526	9.655687
Weight (kg)			71.89977	15.35705

	Factor-variable frequency	Factor-variable percent	Mean	Standard deviation
Sex=1. Male	4,915	47.49		
Sex=2. Female	5,434	52.51		
Age group=1. 20-29	2,320	22.42		
Age group=2. 30-39	1,621	15.66		
Age group=3. 40-49	1,271	12.28		
Age group=4. 50-59	1,291	12.47		
Age group=5. 60-69	2,860	27.64		
Age group=6. 70+	986	9.53		
Rural=0. Urban	6,547	63.26		
Rural=1. Rural	3,802	36.74		
Height (cm)			167.6526	9.655687
Weight (kg)			71.89977	15.35705

Table 1: Descriptive Statistics

	n	%
Sex		
1. Male	4,915	47.49%
2. Female	5,434	52.51%
Age group		
1. 20-29	2,320	22.42%
2. 30-39	1,621	15.66%
3. 40-49	1,271	12.28%
4. 50-59	1,291	12.47%
5. 60-69	2,860	27.64%
6. 70+	986	9.53%
Rural		
0. Urban	6,547	63.26%
1. Rural	3,802	36.74%
	Mean	SD
Weight (kg)	71.90	15.36
Height (cm)	167.65	9.66
n		10349

Data Source: nhanes2b

```
collect remap result[fvfrequency mean] = Col[1 1]
collect remap result[fvpercent sd] = Col[2 2]

collect get resname = "Mean", tag(Col[1] var[mylabel])
collect get resname = "SD", tag(Col[2] var[mylabel])

collect get empty = "  ", tag(Col[1] var[empty])
collect get empty = "  ", tag(Col[2] var[empty])

count
collect get n = `r(N)', tag(Col[2] var[n])

collect layout (var[1.sex 2.sex ///
                    1.agegrp 2.agegrp 3.agegrp ///
                    4.agegrp 5.agegrp 6.agegrp ///
                    0.rural 1.rural ////
                    empty mylabel ///
                    weight height ///
                    empty n]) (Col[1 2])
```

Table 1: Descriptive Statistics

	n	%
Sex		
1. Male	4,915	47.49%
2. Female	5,434	52.51%
Age group		
1. 20-29	2,320	22.42%
2. 30-39	1,621	15.66%
3. 40-49	1,271	12.28%
4. 50-59	1,291	12.47%
5. 60-69	2,860	27.64%
6. 70+	986	9.53%
Rural		
0. Urban	6,547	63.26%
1. Rural	3,802	36.74%
	Mean	SD
Weight (kg)	71.90	15.36
Height (cm)	167.65	9.66
n		10349

Data Source: nhanes2b

```
collect label levels Col 1 "n" 2 "%"

collect style header Col, title(hide)

collect style header var[empty mylabel], level(hide)
collect style row stack, nobinder

collect style cell var[sex agegrp rural]#Col[1], nformat(%6.0fc)
collect style cell var[sex agegrp rural]#Col[2], nformat(%6.2f) sformat("%s%%")
collect style cell var[weight height], nformat(%6.2f)

collect style cell border_block[item row-header], border(top, pattern(nil))

collect title "Table 1: Descriptive Statistics"

collect note "Data Source: nhanes2b"
```

Table 1: Descriptive Statistics

	n	%
Sex		
1. Male	4,915	47.49%
2. Female	5,434	52.51%
Age group		
1. 20-29	2,320	22.42%
2. 30-39	1,621	15.66%
3. 40-49	1,271	12.28%
4. 50-59	1,291	12.47%
5. 60-69	2,860	27.64%
6. 70+	986	9.53%
Rural		
0. Urban	6,547	63.26%
1. Rural	3,802	36.74%
	Mean	SD
Weight (kg)	71.90	15.36
Height (cm)	167.65	9.66
n		10349

Data Source: nhanes2b

```
collect preview

collect export "table.docx", replace
```

Table 1: Descriptive Statistics

	n	%
Sex		
1. Male	4,915	47.49%
2. Female	5,434	52.51%
Age group		
1. 20-29	2,320	22.42%
2. 30-39	1,621	15.66%
3. 40-49	1,271	12.28%
4. 50-59	1,291	12.47%
5. 60-69	2,860	27.64%
6. 70+	986	9.53%
Rural		
0. Urban	6,547	63.26%
1. Rural	3,802	36.74%
	Mean	SD
Weight (kg)	71.90	15.36
Height (cm)	167.65	9.66
n		10349

Data Source: nhanes2b

Table 1: Descriptive Statistics

	n	%
Sex		
1. Male	4,915	47.49%
2. Female	5,434	52.51%
Age group		
1. 20–29	2,320	22.42%
2. 30–39	1,621	15.66%
3. 40–49	1,271	12.28%
4. 50–59	1,291	12.47%
5. 60–69	2,860	27.64%
6. 70+	986	9.53%
Rural		
0. Urban	6,547	63.26%
1. Rural	3,802	36.74%
	Mean	SD
Weight (kg)	71.90	15.36
Height (cm)	167.65	9.66
n		10349

Data Source: nhanes2b

Table 1: Descriptive Statistics

	Unadjusted	Unadjusted	Adjusted
	n	%	%
Sex			
1. Male	4915	47.49%	47.95%
2. Female	5434	52.51%	52.05%
Age group			
1. 20–29	2320	22.42%	28.05%
2. 30–39	1621	15.66%	20.42%
3. 40–49	1271	12.28%	16.83%
4. 50–59	1291	12.47%	16.72%
5. 60–69	2860	27.64%	13.35%
6. 70+	986	9.53%	4.62%
Rural			
0. Urban	6547	63.26%	68.26%
1. Rural	3802	36.74%	31.74%
		Mean (SD)	Mean (SD)
Weight (kg)		71.90 (0.15)	71.90 (0.17)
Height (cm)		167.65 (0.09)	168.46 (0.15)
n	10349		

Data Source: nhanes2b.

Percentages, mean and standard deviation are adjusted for sample design.

```
collect clear

table () (result), ///
    command(prop sex, percent) ///
    command(prop agegrp, percent) ///
    command(prop rural, percent) ///
    command(mean weight) ///
    command(mean height)


table () (result), ///
    command(svy: prop sex, percent) ///
    command(svy: prop agegrp, percent) ///
    command(svy: prop rural, percent) ///
    command(svy: mean weight) ///
    command(svy: mean height) name(Table) append
```

```
collect style row stack, nobinder

collect style cell result[_r_b]#colname[1.sex 2.sex c1 c2 ///
                                        1.agegrp 2.agegrp 3.agegrp ///
                                        4.agegrp 5.agegrp 6.agegrp ///
                                        0.rural 1.rural], sformat(%s%%)


collect get _r_b = "Mean (SD)", tags(cmdset[1] colname[myvar])

collect get _r_b = "Mean (SD)", tags(cmdset[2] colname[myvar])


collect get freq = "n", tags(cmdset[1] colname[myvar0])

collect get _r_b = "%", tags(cmdset[1] colname[myvar0])

collect get _r_b = "%", tags(cmdset[2] colname[myvar0])

collect get _r_b = "  ", tag(cmdset[1]  colname[empty])

collect get _r_b = "  ", tag(cmdset[2]  colname[empty])
```

```
count

collect get freq = `r(N)', tag(cmdset[1] colname[n])

collect remap result[_r_se] = result[se2], fortags(colname[weight height])

collect style cell result[se2], sformat((%s))

collect composite define meansd = _r_b se2

collect style cell result[meansd], nformat(%6.2f)

collect style header colname[myvar], level(hide)

collect style header colname[myvar0], level(hide)

collect style header colname[empty], level(hide)

collect style header result, level(hide)
```

```
collect label levels cmdset 1 "Unadjusted", modify

collect label levels cmdset 2 "Adjusted", modify

collect style header cmdset, title(hide)

collect title "Table 1: Descriptive Statistics"

collect note "Data Source: nhanes2b."

collect note "Percentages, mean and standard deviation are adjusted for sample design."

collect layout (colname[myvar0 sex agegrp rural empty myvar ///
                        weight height empty n]) ///
                        (result[freq]#cmdset[1] result[meansd]#cmdset) ()
```

```
collect preview

collect export "table.docx", replace
```

Table 1: Descriptive Statistics

	Unadjusted	Unadjusted	Adjusted
	n	%	%
Sex			
1. Male	4915	47.49%	47.95%
2. Female	5434	52.51%	52.05%
Age group			
1. 20–29	2320	22.42%	28.05%
2. 30–39	1621	15.66%	20.42%
3. 40–49	1271	12.28%	16.83%
4. 50–59	1291	12.47%	16.72%
5. 60–69	2860	27.64%	13.35%
6. 70+	986	9.53%	4.62%
Rural			
0. Urban	6547	63.26%	68.26%
1. Rural	3802	36.74%	31.74%
		Mean (SD)	Mean (SD)
Weight (kg)		71.90 (0.15)	71.90 (0.17)
Height (cm)		167.65 (0.09)	168.46 (0.15)
n	10349		

Data Source: nhanes2b.
Percentages, mean and standard deviation are adjusted for sample design.

Table 1: Descriptive Statistics

	Unadjusted	Unadjusted	Adjusted
	n	%	%
Sex			
1. Male	4915	47.49%	47.95%
2. Female	5434	52.51%	52.05%
Age group			
1. 20–29	2320	22.42%	28.05%
2. 30–39	1621	15.66%	20.42%
3. 40–49	1271	12.28%	16.83%
4. 50–59	1291	12.47%	16.72%
5. 60–69	2860	27.64%	13.35%
6. 70+	986	9.53%	4.62%
Rural			
0. Urban	6547	63.26%	68.26%
1. Rural	3802	36.74%	31.74%
		Mean (SD)	Mean (SD)
Weight (kg)		71.90 (0.15)	71.90 (0.17)
Height (cm)		167.65 (0.09)	168.46 (0.15)
n	10349		

Data Source: nhanes2b.

Percentages, mean and standard deviation are adjusted for sample design.

The dtable Command

```
dtable i.sex i.agegrp i.rural weight height, export(table.docx, replace)
```

```
                    Summary
-------------------------------------------
N                                    10,349
Sex
  1. Male                     4,915 (47.5%)
  2. Female                   5,434 (52.5%)
Age group
  1. 20-29                    2,320 (22.4%)
  2. 30-39                    1,621 (15.7%)
  3. 40-49                    1,271 (12.3%)
  4. 50-59                    1,291 (12.5%)
  5. 60-69                    2,860 (27.6%)
  6. 70+                        986 (9.5%)
Rural
  0. Urban                    6,547 (63.3%)
  1. Rural                    3,802 (36.7%)
Weight (kg)          71.900 (15.357)
Height (cm)          167.653 (9.656)
-------------------------------------------
(collection DTable exported to file table.docx)
```

	Summary
N	10,349
Sex	
1. Male	4,915 (47.5%)
2. Female	5,434 (52.5%)
Age group	
1. 20–29	2,320 (22.4%)
2. 30–39	1,621 (15.7%)
3. 40–49	1,271 (12.3%)
4. 50–59	1,291 (12.5%)
5. 60–69	2,860 (27.6%)
6. 70+	986 (9.5%)
Rural	
0. Urban	6,547 (63.3%)
1. Rural	3,802 (36.7%)
Weight (kg)	71.900 (15.357)
Height (cm)	167.653 (9.656)

```
dtable, ///
continuous(weight height, statistics( mean min max)) ///
factor(sex agegrp rural, statistics(fvfrequency fvpercent)) ///
column(summary( , hide)) ///
nformat(%9.0f) ///
title(Table of Descriptive Statistics) ///
note(Data Source: nhanes2b.) ///
export(table.docx, replace)
```

Table of Descriptive Statistics	
N	10349
Weight (kg)	72 31 176
Height (cm)	168 136 200
Sex	
1. Male	4915 (47%)
2. Female	5434 (53%)
Age group	
1. 20–29	2320 (22%)
2. 30–39	1621 (16%)
3. 40–49	1271 (12%)
4. 50–59	1291 (12%)
5. 60–69	2860 (28%)
6. 70+	986 (10%)
Rural	
0. Urban	6547 (63%)
1. Rural	3802 (37%)

Data Source: nhanes2b.

```
dtable, by(sex) ///
continuous(weight height, statistics( mean min max)) ///
factor(agegrp rural, statistics(fvfrequency fvpercent)) ///
column(summary( , hide)) ///
nformat(%9.0f) ///
title(Table of Descriptive Statistics) ///
note(Data Source: nhanes2b.) ///
export(table.docx, replace)
```

Table of Descriptive Statistics

	Sex		
	1. Male	2. Female	Total
N	4915 (47%)	5434 (53%)	10349 (100%)
Weight (kg)	78 31 176	66 35 159	72 31 176
Height (cm)	175 145 200	161 136 189	168 136 200
Age group			
1. 20–29	1116 (23%)	1204 (22%)	2320 (22%)
2. 30–39	770 (16%)	851 (16%)	1621 (16%)
3. 40–49	610 (12%)	661 (12%)	1271 (12%)
4. 50–59	602 (12%)	689 (13%)	1291 (12%)
5. 60–69	1369 (28%)	1491 (27%)	2860 (28%)
6. 70+	448 (9%)	538 (10%)	986 (10%)
Rural			
0. Urban	3023 (62%)	3524 (65%)	6547 (63%)
1. Rural	1892 (38%)	1910 (35%)	3802 (37%)

Data Source: nhanes2b.

```
dtable, ///
continuous(weight height, statistics(mean min max)) ///
factor(sex agegrp rural, statistics(fvrawfrequency fvpercent)) ///
column(summary( , hide)) ///
nformat(%9.0f) ///
title(Table of Descriptive Statistics) ///
note(Data Source: nhanes2b.) ///
svy ///
export(table.docx, replace)
```

Table of Descriptive Statistics

N	117131111
Weight (kg)	72 31 176
Height (cm)	168 136 200
Sex	
1. Male	4915 (48%)
2. Female	5434 (52%)
Age group	
1. 20–29	2320 (28%)
2. 30–39	1621 (20%)
3. 40–49	1271 (17%)
4. 50–59	1291 (17%)
5. 60–69	2860 (13%)
6. 70+	986 (5%)
Rural	
0. Urban	6547 (68%)
1. Rural	3802 (32%)

Data Source: nhanes2b.

Tables of Modelling Results

The etable Command

```
regress weight ib1.sex ib0.rural ib3.agegrp height, allbaselevels

etable

collect export "table.docx", replace
```

	weight
Sex	
2. Female	-1.783
	(0.371)
Rural	
1. Rural	0.464
	(0.271)
Age group	
1. 20–29	-5.982
	(0.463)
2. 30–39	-1.797
	(0.496)
4. 50–59	1.187
	(0.524)
5. 60–69	-0.038
	(0.450)
6. 70+	-0.825
	(0.569)
Height (cm)	0.727
	(0.020)
Intercept	-47.725
	(3.476)
Number of observations	10349

```
collect clear

regress weight ib1.sex ib0.rural ib3.agegrp height, allbaselevels

etable

collect style showbase all
collect label levels etable_depvar 1 "Coef. (SE)", modify

etable, replay cstat(_r_b, nformat(%4.2f))  ///
        cstat(_r_se, nformat(%6.2f))  ///
        showstars showstarsnote  ///
        stars(.05 "*" .01 "**" .001 "***", attach(_r_b)) ///
        mstat(N) mstat(aic) mstat(bic) mstat(r2_a)  ///
        title("Table 3: Linear Regression Model of Weight") ///
        titlestyles(font(Arial Narrow, size(14) bold)) ///
        note("Data Source: nhanes2b") ///
        notestyles(font(Arial Narrow, size(10) italic)) ///
        export("table.docx", replace)
```

Table 3: Linear Regression Model of Weight

	Coef. (SE)	
Sex		
1. Male	0.00	
	(0.00)	
2. Female	-1.78	***
	(0.37)	
Rural		
0. Urban	0.00	
	(0.00)	
1. Rural	0.46	
	(0.27)	
Age group		
1. 20–29	-5.98	***
	(0.46)	
2. 30–39	-1.80	***
	(0.50)	
3. 40–49	0.00	
	(0.00)	
4. 50–59	1.19	*
	(0.52)	
5. 60–69	-0.04	
	(0.45)	
6. 70+	-0.83	
	(0.57)	
Height (cm)	0.73	***
	(0.02)	
Intercept	-47.73	***
	(3.48)	
Number of observations	10349	
AIC	82833.57	
BIC	82898.78	
Adjusted R-squared	0.26	

**** p<.001, ** p<.01, * p<.05*

Data Source: nhanes2b

```
collect clear

svy: regress weight ib1.sex ib0.rural ib3.agegrp height, allbaselevels

etable

collect style showbase all
collect label levels etable_depvar 1 "Coef. (SE)", modify

etable, replay cstat(_r_b, nformat(%4.2f))  ///
        cstat(_r_se, nformat(%6.2f))  ///
        showstars showstarsnote  ///
        stars(.05 "*" .01 "**" .001 "***", attach(_r_b)) ///
        mstat(N) mstat(r2)  ///
        title("Table 3: Linear Regression Model of Weight") ///
        titlestyles(font(Arial Narrow, size(14) bold)) ///
        note("Data Source: nhanes2b, results adjusted for sample design.") ///
        notestyles(font(Arial Narrow, size(10) italic)) ///
        export("table.docx", replace)
```

Table 3: Linear Regression Model of Weight

	Coef. (SE)	
Sex		
1. Male	0.00	
	(0.00)	
2. Female	-3.13	***
	(0.58)	
Rural		
0. Urban	0.00	
	(0.00)	
1. Rural	0.93	**
	(0.30)	
Age group		
1. 20–29	-5.82	***
	(0.65)	
2. 30–39	-2.10	***
	(0.56)	
3. 40–49	0.00	
	(0.00)	
4. 50–59	1.06	
	(0.62)	
5. 60–69	-0.06	
	(0.44)	
6. 70+	-0.97	
	(0.61)	
Height (cm)	0.72	***
	(0.03)	
Intercept	-45.79	***
	(5.02)	
Number of observations	10349	
R-squared	0.29	

*** p<.001, ** p<.01, * p<.05

Data Source: nhanes2b, results adjusted for sample design.

```
collect clear

quietly regress weight ib1.sex, allbaselevels
etable

quietly regress weight ib1.sex ib0.rural, allbaselevels
etable, append

quietly regress weight ib1.sex ib0.rural ib3.agegrp height, allbaselevels
etable, append

collect style showbase all

etable, replay column(index) export("table.docx", replace)


collect style showbase all

collect label levels etable_depvar 1 "Model 1" ///
                                   2 "Model 2" ///
                                   3 "Model 3", modify

collect style cell, font(Times New Roman)

etable, replay column(depvar) ///
    cstat(_r_b, nformat(%4.2f))  ///
    cstat(_r_se, nformat(%6.2f))  ///
    showstars showstarsnote  ///
    stars(.05 "*" .01 "**" .001 "***", attach(_r_b)) ///
    mstat(N) mstat(r2_a) mstat(aic) mstat(bic)  ///
    title("Table 1: Linear Regression Models of Weight, Coef. (SE)") ///
    titlestyles(font(Times New Roman, size(12) bold)) ///
    notestyles(font(Times New Roman, size(11) italic)) ///
    export("table.docx", replace)
```

Table 1: Linear Regression Models of Weight, Coef. (SE)

	Model 1		Model 2		Model 3	
Sex						
1. Male	0.00		0.00		0.00	
	(0.00)		(0.00)		(0.00)	
2. Female	-11.59	***	-11.56	***	-1.78	***
	(0.28)		(0.28)		(0.37)	
Rural						
0. Urban			0.00		0.00	
			(0.00)		(0.00)	
1. Rural			0.87	**	0.46	
			(0.29)		(0.27)	
Age group						
1. 20–29					-5.98	***
					(0.46)	
2. 30–39					-1.80	***
					(0.50)	
3. 40–49					0.00	
					(0.00)	
4. 50–59					1.19	*
					(0.52)	
5. 60–69					-0.04	
					(0.45)	
6. 70+					-0.83	
					(0.57)	
Height (cm)					0.73	***
					(0.02)	
Intercept	77.98	***	77.65	***	-47.73	***
	(0.20)		(0.23)		(3.48)	
Number of observations	10349		10349		10349	
Adjusted R-squared	0.14		0.14		0.26	
AIC	84325.41		84318.47		82833.57	
BIC	84339.90		84340.20		82898.78	

*** $p<.001$, ** $p<.01$, * $p<.05$

```
collect clear

logit heartatk ib1.sex ib0.rural ib3.agegrp

etable

margins, dydx(sex rural agegrp) asobserved nose

etable, append margins  ///
    cstat(_r_b, nformat(%4.2f))  ///
    cstat(_r_se, nformat(%6.2f))showstars showstarsnote  ///
    stars(.05 "*" .01 "**" .001 "***", attach(_r_b)) ///
    mstat(N) mstat(r2_p)

collect style showbase all

collect layout

collect label levels etable_depvar 1 "Log Odds (SE)" ///
                                   2 "AME", modify

collect layout

collect export "table.docx", replace
```

	Log Odds (SE)		AME
Sex			
1. Male	0.00		0.00
	(0.00)		
2. Female	-0.89	***	-0.04
	(0.10)		
Rural			
0. Urban	0.00		0.00
	(0.00)		
1. Rural	0.12		0.01
	(0.10)		
Age group			
1. 20–29	-3.60	***	-0.02
	(1.03)		
2. 30–39	-1.86	***	-0.01
	(0.55)		
3. 40–49	0.00		0.00
	(0.00)		
4. 50–59	1.28	***	0.04
	(0.26)		
5. 60–69	1.87	***	0.08
	(0.23)		
6. 70+	2.17	***	0.10
	(0.25)		
Intercept	-3.82	***	
	(0.23)		
Number of observations	10349		10349
Pseudo R-squared	0.17		

*** p<.001, ** p<.01, * p<.05

Outputting Saved Results

```
    putdocx begin

    putdocx paragraph, style(Title)

    putdocx text ("My Research Report")

    putdocx textblock begin
This report presents analysis of heart attacks using the NHANES data.
    putdocx textblock end

    putdocx paragraph, style(Heading1)

    putdocx text ("Results")
```

My Research Report

This report presents analysis of heart attacks using the NHANES data.

Results

```
    logit heartatk ib1.sex ib0.rural ib3.agegrp, allbaselevels

    local r2 : display %6.3f e(r2_p)
    local n = e(N)
    local sex : display %6.3f _b[2.sex]

    putdocx textblock begin

There was a significant association between sex and suffering a heart attack.

The log odds for women is <<dd_docx_display: `sex'>>.

There are <<dd_docx_display: `n'>> sample members.

The Pseudo R2 of the logistic regression model is <<dd_docx_display: `r2'>>.

    putdocx textblock end
```

There was a significant association between sex and suffering a heart attack. The log odds for women is -.889. There are 10349 sample members. The Pseudo R2 of the logistic regression model is .167.

```
putdocx paragraph

etable, replay cstat(_r_b, nformat(%4.2f))  ///
        cstat(_r_se, nformat(%6.2f))  ///
        showstars showstarsnote  ///
        stars(.05 "*" .01 "**" .001 "***", attach(_r_b)) ///
        mstat(N) mstat(r2)  ///
        title("Table 1: Logistic Regression: Prior Heart Attack") ///
        titlestyles(font(Arial Narrow, size(14) bold)) ///
        note("Data Source:  NHANES.") ///
        notestyles(font(Arial Narrow, size(10) italic))

putdocx table results = etable

putdocx save "mypaper.docx", replace
```

My Research Report

This report presents analysis of heart attacks using the NHANES data.

Results

There was a significant association between sex and suffering a heart attack. The log odds for women is -.889. There are 10349 sample members. The Pseudo R2 of the logistic regression model is .167.

heartatk	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
sex						
1. Male	0	(base)				
2. Female	-.8885162	.1017036	-8.74	0.000	-1.087852	-.6891808
rural						
0. Urban	0	(base)				
1. Rural	.1205816	.0982396	1.23	0.220	-.0719645	.3131277
agegrp						
1. 20–29	-3.604438	1.025455	-3.51	0.000	-5.614294	-1.594582
2. 30–39	-1.863896	.5492542	-3.39	0.001	-2.940415	-.7873779
3. 40–49	0	(base)				
4. 50–59	1.283435	.2577225	4.98	0.000	.7783081	1.788562
5. 60–69	1.870414	.2349129	7.96	0.000	1.409993	2.330835
6. 70+	2.172638	.2467651	8.80	0.000	1.688987	2.656288
_cons	-3.817548	.2320573	-16.45	0.000	-4.272372	-3.362724

Good Table Manners

A table should stand alone, complete and informative in itself

- Use informative titles
- Ensure all elements are clearly labelled
- Do not use variable names
- Provide any other necessary information in a table note

Photo by Samantha Gades on Unsplash

Always think from the reader's perspective

Automation will make you happier researcher!

Photo by Jacqueline Munguía on Unsplash

Producing Automated Tables using Stata

Monday
18/12/2023

Roxanne Connelly
University of Edinburgh
roxanne.connelly@ed.ac.uk

Banco De Portugal
Workshop

Photo by THE 9TH Coworking on Unsplash