

Making Working & Publication-style Tables in Stata

See also the course's documents 'Making Contingency Tables in Stata' and 'Editing Stata estab and outreg2 Tables in MS-Word'.

Gary Klass, "How to Construct Bad Charts and Graphs"
<http://lilt.ilstu.edu/gmklass/pos138/datadisplay/badchart.htm>

Gary Klass, "Presenting Data: Tabular and graphic display of social indicators"
<http://lilt.ilstu.edu/gmklass/pos138/datadisplay/>

Jon Kolko, "Edward Tufte" (on principles of data presentation)
http://facultypages.scad.edu/~jkolko/IACT370_InformationArchitecture/IACT370_05_TuftePrinciples.pdf

Note: Copying tables directly from the Stata Results-window to MS-Word works well if MS-Word is formatted to display Courier 11 point or less (which is a fixed-space font). If necessary you can then change to a proportionally spaced font (such as Times New Roman) in MS-Word, but doing so may involve considerable formatting (see the course web document 'Editing estab and outreg2 Tables in MS-Word').

Here are examples of Stata table-making commands, which you can obtain by using 'findit,' 'help,' Stata listserv, 'UCLA Resources for Learning Stata', and 'ssc whatsnew.'

- tabstat, statsmat, tablemat, fsum, and tabform: summary tables (see also 'findit ...' for svytabs and svytabform).
- tabout: publication-style summary and contingency tables (see also svytabs and svytabform).
- mkcorr, corrtab, and makematrix: publication-style correlation matrices.
- The combination of eststo and esttab (which is a spin-off from estout; see below): not only displays publication-style regression tables in the Stata results window, but also can directly produce them in Word or Excel, and output them to LaTeX.
- outreg2: can directly create regression tables in Word or Excel, and output them to LaTeX.

*Formatting 'by group':
By-group as row variable*

	Ave. Income	% Poverty
Community1	12,636	34.3
Community2	15,151	31.1
Community3	16,204	29.5

By-group as column variable

	Community1	Community 2	Community 3
Ave. income	12,236	15,151	16,204
% poverty	34.3%	31.1%	29.5%

- **tabstat**

help tabstat

```
tabstat read write math science, stats(mean med sd min max) format(%9.2f)
```

```
tabstat read write math science, stats(mean med sd min max) format(%9.2f)  
col(stats)
```

```
tabstat read write math science, stats(mean med sd min max) format(%9.2f)  
by(female) col(stats)
```

Note: 'semean' displays standard error, 'cv' displays coefficient of variation.

- **statsmat** (findit & download; basically the same as tabstat, but displays & exports table as matrix)

findit statsmat

help statsmat

```
statsmat science, by(female) f(%9.0f) s(n min q max) matrix(m1)
```

```
statsmat science, by(female) f(%9.0f) s(n min q max) xpose matrix(m2)
```

- **tablemat** (findit & download; permits more flexible use of by-variables and labeling than tabstat & statsmat; displays the same statistics as tabstat, but displays only one statistic at a time)

findit tablemat

help tablemat

```
tablemat read write math, st(mean) by(ses) name(table) format(%9.2f)
```

```
tablemat read write math, st(mean) by(female ses prog) name(table) format(%9.2f)
```

- **fsum** (findit & download; permits display of # missing & lower-upper confidence intervals, which tabstat, statsmat, & tablemat do not)

findit fsum

help fsum

```
fsum read write math, s(n miss mean sd lci uci) format(%9.2f)
```

- **tabform** (findit & download; permits display of confidence intervals, which tabstat & tablemat do not; tests & displays results of equality of means, which tabstat, statsmat, tablemat, & fsum do not)

findit tabform
help tabform

tabform read write math using table.xls, by(ses) sd ci dec(2) sdbracket cibrace nototal

Open Excel, navigate to table.xls, and follow the wizard instructions to display the table.

tabform displays only the mean if other statistics are not requested.

tabform read write math using table.xls, by(ses) sd ci bdec(2) sdbracket cibrace nototal vertical

- **Publication-style summary & contingency tables (tabout):**

findit tabout
help tabout

use hsb2, clear

```
la var schtyp "School Type"
la def st 0 "Public" 1 "Private"
la val schtyp st
tab schtyp
```

```
la var ses "SES"
la def ses 1 "Low" 2 "Middle" 3 "High"
la val ses ses
tab ses
```

tabout ses schtyp using sesst, cell(col) format(1) clab(Col_%) stats(chi2) replace

Select MS-Word file 'sesst' (located in current folder), and edit in Word or copy/paste to Excel.

Some options:

cell: freq, cell, row, col, cum [all can be specified]

format: # decimal points; specify in order of types of cell entries specified

clab: table's title

layout: how the columns are laid out stats: types of statistics specified

- **Publication-style correlation matrix (mkcorr):**

findit mkcorr
help mkcorr

mkcorr read write math science socst, log(hsb2corr) sig means cdec(2) mdec(2)
casewise

Note: The default is pairwise (pwcrr). You must specify 'casewise' if you want the equivalent results of 'corr'.

Select file 'hsb2corr' (located in current folder) in Word, and edit in Word or copy/paste to Excel.

Options:

sig: display significance test p-values

cdec(): # decimal points displayed in correlation coefficients

means: display descriptive numerical summary

mdec(): # decimal points displayed in descriptive numerical summary

casewise: display casewise results (instead of default pairwise results)

- **Publication-style correlation matrix (corrtab):**

findit corrtab

help corrtab

corrtab read write math science, obs sig bonf vsort(read) format(%9.2f)

corrtab read write math science, cwd obs sig bonf vsort(read) format(%9.2f)

corrtab read write math science, spearman obs sig bonf vsort(read) format(%9.2f)

Note: pairwise (the equivalent of pwcrr in Stata) is the default unless cwd (casewise, equivalent to 'corr' in Stata) is not specified. Spearman may also be specified. corrtab is designed for a maximum of eight variables.

- **Publication-style correlation matrix (makematrix) :**

findit makematrix

help makematrix

makematrix, from(r(rho)) format(%9.2f): corr read-socst

makematrix, from(r(rho)) col(socst) format(%9.2f): corr read-socst

- **Working or publication-style OLS regression table (esttab):**

findit esttab

help esttab

help eststo

la var math "Math"

la var read "Read"

la var female "Female"

reg science math read female

eststo

[eststo refers to 'estimates store']

```
reg science math read female schtyp
eststo
```

```
esttab, se starlevels(+ .10 * .05 ** .01) b(%9.2f) r2(%9.2f) ar2(%9.2f) aic bic
nodepvars nomtitles title("OLS Science Models") addnotes("Note: Put your notes
here.")
```

Or to directly produce an MS-Word table:

```
esttab using science.rtf, se b(%9.2f) starlevels(+ .10 * .05 ** .01) r2(%9.2f)
ar2(%9.2f) aic bic nodepvars nomtitles title("OLS Science Models") addnotes("Note:
Put your notes here.")
```

Then click on the link in the Stata results window.

Note: 'csv' instead of 'rtf' creates an Excel document. See 'help esttab' regarding 'drop,' 'keep,' 'order'; formatting numbers; 'compress' (which squeezes the column space); 'wide' (to display standard errors next to coefficients); 'ci' (to display confidence intervals); 'beta' (to display standardized coefficients); optional specification of 'replace' or 'append'; and other options.

Note: see 'help eststo'.

- **Publication-style logistic regression table (esttab):**

```
la var white "White"
la var science "Science"
```

```
xi3:logistic white math science i.prog, nolog
eststo
```

```
esttab, eform se b(%9.3f) starlevels(+ .10 * .05 ** .01) scalars(ll df_m aic bic)
sfmt(%9.3f %9.0f %9.3f) nodepvars title("Logistic Whites vs. Others Model")
addnotes("Note: Put your notes here.")
```

Or:

```
estab using whites.rtf, eform se b(%9.3f) starlevels(+ .10 * .05 ** .01) scalars(ll
df_m aic bic) sfmt(%9.3f %9.0f %9.3f) nodepvars title("Logistic Whites vs. Others
Model") addnotes("Note: Put your notes here.")
```

Note: 'eform' to display odds ratios

- **Publication-style multinomial logit regression table (esttab):**

```
tab prog
tab prog, nlabel
mlogit prog math read science female, base(2) rrr nolog
eststo
```

```
esttab m1, eform se unstack b(%9.3f) starlevels(+ .10 * .05 ** .01) scalar(ll chi2
df_m aic bic) scalars(ll df_m aic bic) sfmt(%9.3f %9.0f %9.3f) nodepvars nomtitles
title("Multinomial Logit Model") addnotes("Note: Put your notes here.")
```

Or:

```
estab using whites.rtf, eform se unstack b(%9.3f) starlevels(+ .10 * .05 ** .01)
scalar(ll chi2 df_m aic bic) scalars(ll df_m aic bic) sfmt(%9.3f %9.0f %9.3f)
nodepvars nomtitles title("Multinomial Logit Model") addnotes("Note: Put your notes
here.")
```

Note: eform, as well as 'unstack' to display sub-equations in separate columns.

- **Using esttab in profile.do to make easy regression tables**

You can routinely make easy-to-read, full-feature regression tables for exploratory analysis if you include the following commands in profile.do (specifying any function keys you wish):

```
* eststo
global F1 = "eststo "
```

```
* esttab OLS regression table
global F2 = "esstab , se r2 ar2 aic bic starlevels(+ .10 * .05 ** .01 *** .001)
nodepvars nomtitles compress"
```

```
* esttab logistic regression table
global F3 = "estabb , eform se starlevels(+ .10 * .05 ** .01 *** .001) scalars(ll
chi2 df_m aic bic) nodepvars nomtitles compress"
```

```
* esttab multinomial logistic regression table
global F4 = "esstab , eform se unstack starlevels(+ .10 * .05 ** .01 *** .001)
scalars(ll chi2 df_m aic bic) nodepvars nomtitles compress"
```

Here's how to use the first of these commands to make an OLS regression working table, given that you've inserted the above commands into profile.do:

```
use hsb2, clear
regress science read write
F1
```

```
regress science read write math female
F1
```

F2 [or whatever function key you may have programmed, to produce the table]

Here's the result:

	(1)	(2)
read	0.435*** (0.0633)	0.261*** (0.0680)
write	0.315*** (0.0684)	0.303*** (0.0747)
math		0.282*** (0.0755)
female		-3.570*** (1.059)
_cons	12.51*** (3.160)	9.291** (3.123)
N	200	200
R-sq	0.456	0.527
adj. R-sq	0.450	0.518
AIC	1367.9	1343.7
BIC	1377.8	1360.2

Standard errors in parentheses

+ p<.10, * p<.05, ** p<.01, *** p<.001

▪ **Publication-style OLS regression table (outreg2):**

```
reg science math read female
outreg2 using m1, replace [Click 'seeout' in Stata results window.]
outreg2 using m1s, word replace [Click 'ols_science.rtf'.]
outreg2 using m1, excel replace [Click 'ols_science.xml'.]
outreg2 using m1, alpha(0.001, 0.01, 0.05) symbol(***, **, *) adjr bdec(2)
    ctitle(Science) addnote(Do not try this at home.) onecol long replace
outreg2 using m1, ci replace
outreg2 using m1, beta replace [betas in parentheses, in place of se's]
outreg2 read female, replace [Displays only the specified explanatory variables.]
```

Click on 'seeout' to see Stata display & 'ols.RTF' to see Word display.

Table>Format>Hide gridlines

```
reg science math read
outreg2 using m1, onecol long replace
reg science math read female prog
outreg2 using m1, onecol long append [to display results of both models]
Click on 'seeout' to see STATA display & 'ols.RTF' to see Word display.
Table>Format>Hide gridlines
```

Some options (after ','):

word: output text to a Word file, which is saved in the current directory

excel: output text to an Excel file

onecol: suppresses multi-column format if relevant

long: accompanies onecol, forcing Word or Excel to adopt one-column format

replace: replaces previous model or format

append: appends additional models (instead of 'replace')

ci: specifies confidence intervals instead of coefs

beta: specifies standardized coefs

bdec(): specifies # decimal places for regression coefs (default=3)

tdec(): if t-values are displayed, specifies # decimal places (default=3)

rdec(): if r2 or adj r2 is displayed, specifies # decimal places (default=3)

adec(): if additional stats are displayed (see 'e' and addstat')

bfmt(): specifies format for regress coefs – f, fixed; g, general

symbol(): specifies significance values; default is symbol(***, **, *)

alpha(): specifies significance levels; default is alpha(0.000, 0.01, 0.05)

adjr2: specifies adjusted r2

nor2: specifies not to display r2

e(): specify added stats, e.g., e(ll df_m chi2 aic); or e(all); does not display publication-style stat names (see 'addstat'). After estimating a model, type 'ereturn list' to obtain options and codes.

addstat(): to add other stats along with stat names. E.g., addstat(Log likelihood, e(ll), DF, e(df_m), chi2, e(chi2), AIC, e(aic)). After estimating a model, type 'ereturn list' to obtain options and codes.

title(): specifies title

ctitle(): specifies column title

addnote(): to add notes

Note: title and note options give the same output whether or not the text is enclosed within quotation marks. outreg2 gives an error message if addnote is formatted as follows: addnote(Source: my data). So don't include a colon in the text.

E.g.:

reg science math read

outreg2 using m1, bdec(2) bfmt(f) alpha(0.01, 0.05, 0.10) adjr2 addstat(AIC, e(aic))

title(OLS) addnote(Study data) word replace

Click on 'seeout' to see STATA display & 'm1.RTF' to see Word display.

Table>Format>Hide gridlines

Note: To display more than three levels of significance - alpha(0.001 0.01, 0.05, 0.10) symbol(***, **, *, +). To display two or more models in the same table – outreg2 using m1, append

- **Publication-style logistic regression table (outreg2)**

See the course document 'Editing Stata's esttab or outreg2 regression tables in MS-Word'.

```
logistic female math read white
outreg2 using m1, eform nor2 replace
outreg2 using m1, eform nor2 e(ll df_m chi2 aic) word replace
outreg2 using m1, eform nor2 addstat(Log likelihood, e(ll), DF, e(df_m), Chi2,
e(chi2)) word replace
```

Note: eform specifies odds-ratios. See other outreg2 options above. To display more than three levels of significance - alpha(.001 .01, .05, .10) symbol(***, **, *, +). To display two or more models in the same table – outreg using m1, append

- **Publication-style multinomial logit regression table (outreg2)**

```
mlogit ses math read, base(1) nolog
outreg2 using m1, eform nor2 addstat(ll, e(ll), df_m, e(df_m), chi2, e(chi2)) word
replace
```

```
mlogit ses math read write science, base(1) nolog
outreg2 using m1, eform nor2 e(ll df_m chi2) append
```

Note: See other outreg2 options above.