Create customizable tables

Zhao Xu

Principal Software Engineer
StataCorp LLC

2023 Chinese Stata Conference
Aug 14, 2023

Challenges in creating customized tables

Common challenges in creating a customized table:

- Need to copy Stata generated results to a table with predefined layout
  - Error-prone and time-consuming

Challenges in creating customized tables

Common challenges in creating a customized table:

- Need to copy Stata generated results to a table with predefined layout
  - Error-prone and time-consuming
- Need to tweak the table layout, result formats, cell styles, etc
  - More time-consuming

Challenges in creating customized tables

Common challenges in creating a customized table:

- Need to copy Stata generated results to a table with predefined layout
  - Error-prone and time-consuming
- Need to tweak the table layout, result formats, cell styles, etc
  - More time-consuming
- Need to update the table due to new or improved data
  - Redo step 1 and 2

Challenges in creating customized tables

Common challenges in creating a customized table:

- Need to copy Stata generated results to a table with predefined layout
  - Error-prone and time-consuming
- Need to tweak the table layout, result formats, cell styles, etc
  - More time-consuming
- Need to update the table due to new or improved data
  - Redo step 1 and 2
- Need to create similar tables with consistent styles frequently
  - Redo step 1, 2, and 3

Challenges in creating customized tables

Common challenges in creating a customized table:

- Need to copy Stata generated results to a table with predefined layout
  - Error-prone and time-consuming
- Need to tweak the table layout, result formats, cell styles, etc
  - More time-consuming
- Need to update the table due to new or improved data
  - Redo step 1 and 2
- Need to create similar tables with consistent styles frequently
  - Redo step 1, 2, and 3
- Need to export the table to different document formats (Word, PDF, Excel, etc.)
  - Headache

Challenges in creating customized tables

Efficient Solutions?

Efficient solutions

Tools for creating customizable tables:

- **dtable** - Create a table of summary statistics, a.k.a, Table 1 (Stata 18)

Efficient solutions

Tools for creating customizable tables:

- **dtable** - Create a table of summary statistics, a.k.a, Table 1 (Stata 18)
- **etable** - Create a table of estimation results (Stata 17)
- **table** - Create a table of frequencies, summaries, and command results (redesigned in Stata 17)
- **collect** - A suite of commands to create customizable tables (Stata 17)

Efficient solutions

Tools for creating reproducible reports with Stata results:

- **putdocx** - Create Word (.docx) documents
- **putpdf** - Create PDF (.pdf) documents
- **putexcel** - Create Excel (.xlsx/.xls) files
- **dyndoc** - Convert dynamic Markdown document to HTML or Word document
- **dyntext**/**docx2pdf**/**html2docx** - Utility commands

New customizable tables

Table 1: Summary Statistics

	Car origin		
	Domestic	Foreign	Total
N	52 (70.3%)	22 (29.7%)	74 (100.0%)
Price	6072 (3097.1)	6385 (2621.9)	6165 (2949.5)
Weight (lbs.)	3317 (695.4)	2316 (433.0)	3019 (777.2)
Mileage (mpg)	20 (4.7)	25 (6.6)	21 (5.8)
Repair record 1978			
1	2 (4.2%)	0 (0.0%)	2 (2.9%)
2	8 (16.7%)	0 (0.0%)	8 (11.6%)
3	27 (56.2%)	3 (14.3%)	30 (43.5%)
4	9 (18.8%)	9 (42.9%)	18 (26.1%)
5	2 (4.2%)	9 (42.9%)	11 (15.9%)

New customizable tables

Table 2: Linear regression results

	1		2		3	
Age of mother	7.97		7.04		10.42	
	[-11.88, 27.82]		[-12.54, 26.62]		[-8.97, 29.82]	
Weight at last menstrual period	4.18	*	4.02	*	3.73	*
	[0.74, 7.62]		[0.63, 7.41]		[0.37, 7.10]	
Smoked during pregnancy						
Smoker			-268.15	*	-229.90	*
			[-476.86,-59.43]		[-438.55,-21.26]	
Premature labor history (count)						
1					-457.23	**
					[-764.86,-149.60]	
2					-115.18	
					[-742.44,512.08]	
3					887.83	
					[-500.05,2275.71]	
Intercept	2216.22	***	2363.77	***	2363.79	***
	[1625.81,2806.63]		[1770.54,2956.99]		[1779.81,2947.76]	
Observations	189		189		189	
AIC	3026		3021		3017	
BIC	3035		3034		3039	
R2	0.0378		0.0701		0.1215	

*** p<.001, ** p<.01, * p<.05
Model 1: age lwt
Model 2: age lwt i.smoke
Model 3: lwt i.smoke i.ptl

New customizable tables

Table 3: Table of test results

	Normotensive	Hypertensive	Diff	pvalue	
Age (years)	42.17	54.97	12.81	0.0000	***
Height (cm)	167.72	167.55	-0.17	0.3661	
Weight (kg)	68.27	76.86	8.59	0.0000	***
Body Mass Index	24.20	27.36	3.16	0.0000	***
Systolic Blood Pressure	116.49	150.54	34.05	0.0000	***
Diastolic Blood Pressure	74.17	92.01	17.84	0.0000	***
Serum cholesterol (mg/dL)	208.73	229.88	21.15	0.0000	***
Serum triglycerides (mg/dL)	129.23	166.04	36.81	0.0000	***
High density lipids (mg/dL)	49.94	49.22	-0.73	0.0195	**
Hemoglobin (g/dL)	14.14	14.42	0.28	0.0000	***
Hematocrit (%)	41.65	42.44	0.79	0.0000	***
Serum iron (mcg/dL)	101.84	96.17	-5.67	0.0000	***
Serum albumin (g/dL)	4.68	4.65	-0.03	0.0001	***
Serum vitamin C (mg/dL)	1.05	1.02	-0.03	0.0070	***
Serum zinc (mcg/dL)	87.06	85.75	-1.32	0.0000	***
Serum copper (mcg/dL)	125.08	126.34	1.26	0.0674	*
Lead (mcg/dL)	13.88	14.93	1.06	0.0000	***

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

Export tables to different document formats

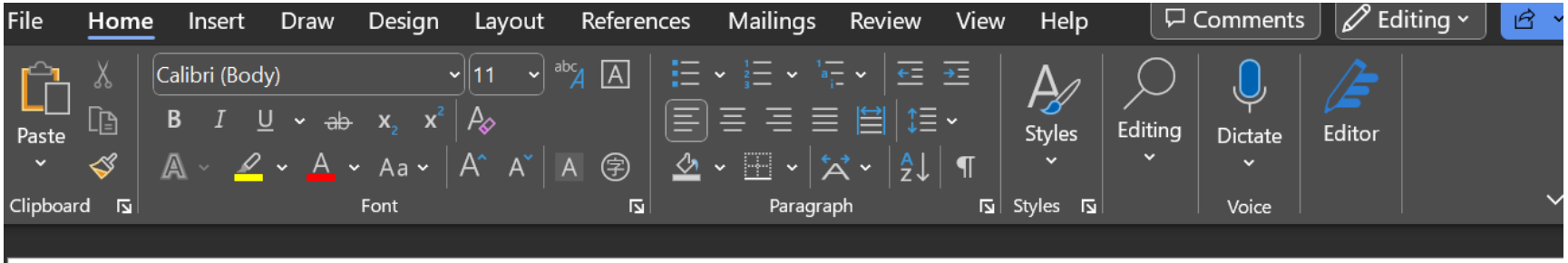


Table 1: Summary Statistics

	Car origin		
	Domestic	Foreign	Total
N	52 (70.3%)	22 (29.7%)	74 (100.0%)
Price	6072 (3097.1)	6385 (2621.9)	6165 (2949.5)
Weight (lbs.)	3317 (695.4)	2316 (433.0)	3019 (777.2)
Mileage (mpg)	20 (4.7)	25 (6.6)	21 (5.8)
Repair record 1978			
1	2 (4.2%)	0 (0.0%)	2 (2.9%)
2	8 (16.7%)	0 (0.0%)	8 (11.6%)
3	27 (56.2%)	3 (14.3%)	30 (43.5%)
4	9 (18.8%)	9 (42.9%)	18 (26.1%)
5	2 (4.2%)	9 (42.9%)	11 (15.9%)

Export tables to different document formats

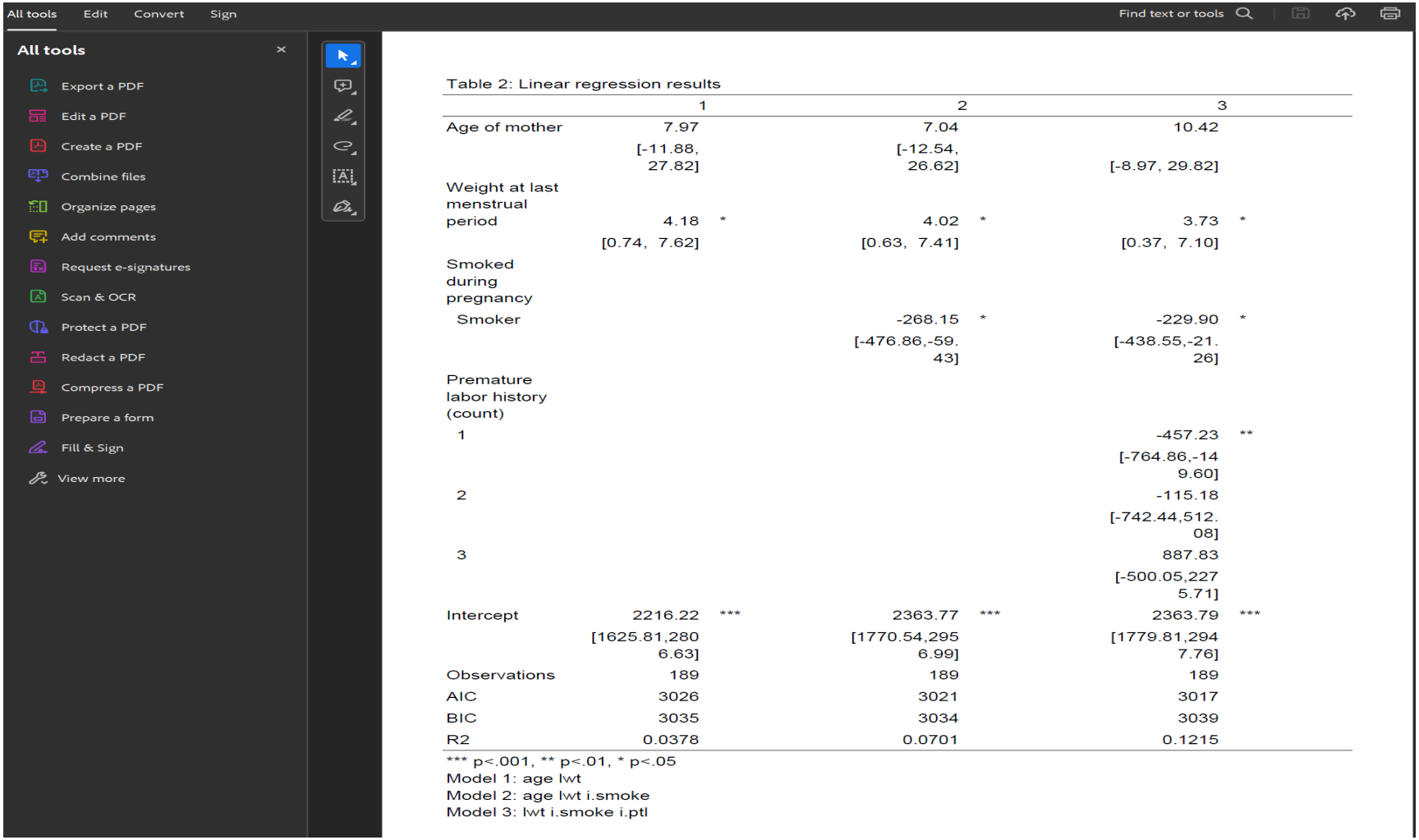
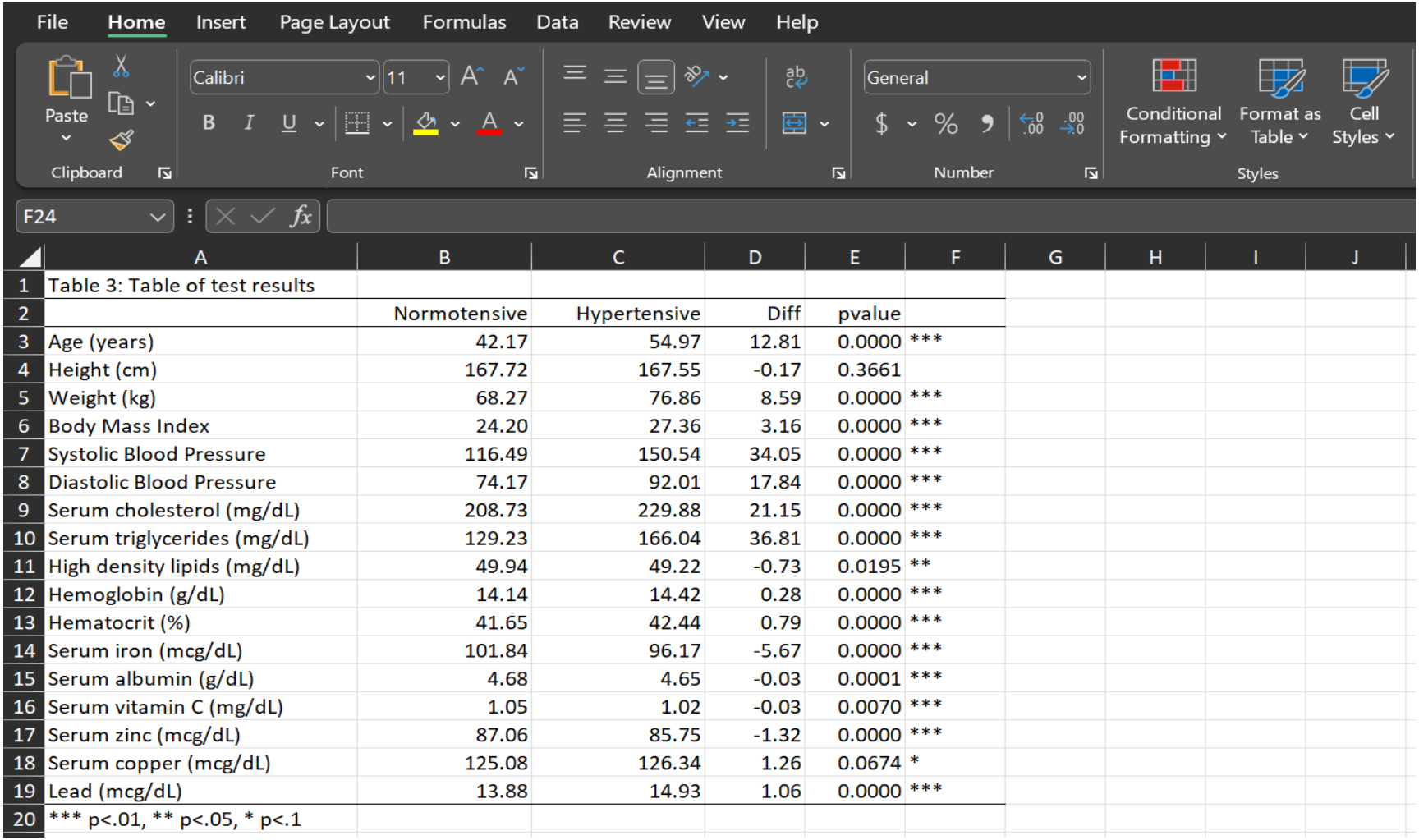


Table 2: Linear regression results

	1		2		3	
Age of mother	7.97		7.04		10.42	
	[-11.88, 27.82]		[-12.54, 26.62]		[-8.97, 29.82]	
Weight at last menstrual period	4.18	*	4.02	*	3.73	*
	[0.74, 7.62]		[0.63, 7.41]		[0.37, 7.10]	
Smoked during pregnancy						
Smoker			-268.15	*	-229.90	*
			[-476.86,-59.43]		[-438.55,-21.26]	
Premature labor history (count)						
1					-457.23	**
					[-764.86,-149.60]	
2					-115.18	
					[-742.44,512.08]	
3					887.83	
					[-500.05,2275.71]	
Intercept	2216.22	***	2363.77	***	2363.79	***
	[1625.81,2806.63]		[1770.54,2956.99]		[1779.81,2947.76]	
Observations	189		189		189	
AIC	3026		3021		3017	
BIC	3035		3034		3039	
R2	0.0378		0.0701		0.1215	

*** p<.001, ** p<.01, * p<.05
Model 1: age lwt
Model 2: age lwt i.smoke
Model 3: lwt i.smoke i.ptl

Export tables to different document formats



	A	B	C	D	E	F
1	Table 3: Table of test results					
2		Normotensive	Hypertensive	Diff	pvalue	
3	Age (years)	42.17	54.97	12.81	0.0000	***
4	Height (cm)	167.72	167.55	-0.17	0.3661	
5	Weight (kg)	68.27	76.86	8.59	0.0000	***
6	Body Mass Index	24.20	27.36	3.16	0.0000	***
7	Systolic Blood Pressure	116.49	150.54	34.05	0.0000	***
8	Diastolic Blood Pressure	74.17	92.01	17.84	0.0000	***
9	Serum cholesterol (mg/dL)	208.73	229.88	21.15	0.0000	***
10	Serum triglycerides (mg/dL)	129.23	166.04	36.81	0.0000	***
11	High density lipids (mg/dL)	49.94	49.22	-0.73	0.0195	**
12	Hemoglobin (g/dL)	14.14	14.42	0.28	0.0000	***
13	Hematocrit (%)	41.65	42.44	0.79	0.0000	***
14	Serum iron (mcg/dL)	101.84	96.17	-5.67	0.0000	***
15	Serum albumin (g/dL)	4.68	4.65	-0.03	0.0001	***
16	Serum vitamin C (mg/dL)	1.05	1.02	-0.03	0.0070	***
17	Serum zinc (mcg/dL)	87.06	85.75	-1.32	0.0000	***
18	Serum copper (mcg/dL)	125.08	126.34	1.26	0.0674	*
19	Lead (mcg/dL)	13.88	14.93	1.06	0.0000	***
20	*** p<.01, ** p<.05, * p<.1					

Features

With **dtable**, **etable**, **table**, and **collect** commands, now you can easily create tables of

- Descriptive (summary) statistics and test statistics comparing groups on summary statistics
- Estimation and postestimation results
- Results of hypothesis tests
- Results returned by any Stata commands
- Combinations of the above

Features

- Customize tables by changing
  - Table layout (one-way, two-way, or multiway)
  - String and numeric formats of results
  - Row headers and column headers
  - Significance stars
  - Table title and notes
  - Table and cell font, size, and color
  - Cell appearance styles, including borders, shading, margins, alignment, etc.
- **Save the table styles and formats in a style file** and reuse for other tables later.

Features

- Export tables to different formats

suffix	Output format
.docx	Microsoft Word
.pdf	PDF
.xlsx	Microsoft Excel 2007/2010 or newer
.xls	Microsoft Excel 1997/2003
.html	HTML 5 with CSS
.tex	LaTeX
.smcl	SMCL
.txt	plain text
.markdown	Markdown
.md	Markdown

- Include those tables in documents created by **putdocx**, **putexcel**, **putpdf**, and **dyndoc**

Create customizable tables with dtable

dtable allows you to easily create a table of descriptive statistics, commonly known as "Table 1".

- Various statistics are supported (mean, median, sd, skewness, kurtosis, quartiles, frequency, percentage, etc).
- Report statistics based on sample groups.
- Report statistics for equality tests of variables between groups.

Syntax:

```
dtable [varlist] [, by(varname)
    continuous([varlist_c][, statistics(cstats)])
    factor([varlist_f][, statistics(fstats)])
    export(filename) options]
```

A quick example

```
. webuse nhanes2l, clear
(Second National Health and Nutrition Examination Survey)

. dtable bpsystol age weight i.race i.hlthstat

----------------------------------------
                               Summary
----------------------------------------
N                                 10,351
Systolic blood pressure 130.882 (23.333)
Age (years)              47.580 (17.215)
Weight (kg)              71.898 (15.356)
Race
  White                   9,065 (87.6%)
  Black                   1,086 (10.5%)
  Other                     200 (1.9%)
Health status
  Excellent               2,407 (23.3%)
  Very good               2,591 (25.1%)
  Good                    2,938 (28.4%)
  Fair                    1,670 (16.2%)
  Poor                      729 (7.1%)
----------------------------------------
```

Statistics for different groups

```
. dtable bpsystol age weight i.race i.hlthstat, by(diabetes)
      nformat(%12.2f mean sd) title(Table 1)
      export(mytable1.docx, replace)

Table 1
------------------------------------------------------------------------
                                        Diabetes status
                          Not diabetic     Diabetic         Total
------------------------------------------------------------------------

N                          9,850 (95.2%)      499 (4.8%) 10,349 (100.0%)
Systolic blood pressure 130.09 (22.76) 146.65 (28.39)  130.89 (23.33)
Age (years)                46.92 (17.19)  60.69 (11.47)   47.58 (17.22)
Weight (kg)                71.66 (15.22)  76.67 (17.18)   71.90 (15.36)
Race
  White                    8,659 (87.9%)    404 (81.0%)   9,063 (87.6%)
  Black                    1,000 (10.2%)     86 (17.2%)   1,086 (10.5%)
  Other                      191 (1.9%)       9 (1.8%)      200 (1.9%)
Health status
  Excellent                2,383 (24.2%)      24 (4.8%)   2,407 (23.3%)
  Very good                2,546 (25.9%)      45 (9.0%)   2,591 (25.1%)
  Good                     2,805 (28.5%)    133 (26.7%)   2,938 (28.4%)
  Fair                     1,508 (15.3%)    162 (32.5%)   1,670 (16.2%)
  Poor                       594 (6.0%)     135 (27.1%)      729 (7.1%)
------------------------------------------------------------------------
(collection DTable exported to file mytable1.docx)
```

Descriptive statistics

By default, **mean** and **standard deviation** are reported for continuous variables, **frequencies** and **percentages** are reported for factor variables.

cstats	Description
mean	mean
sd	standard deviation
median	median
min	minimum value
max	maximum value
q1	first quartile
q2	second quartile
q3	third quartile
And more	...

fstats	Description
fvfrequency	frequency of each level
fvpercent	percentage within each level
fvproportion	proportion within each level
fvrawfrequency	unweighted frequency
fvrawpercent	unweighted percentage
fvrawproportion	unweighted proportion

Different descriptive statistics for different variables

```
.  dtable bpsystol age weight i.race i.hlthstat, by(diabetes, nototals)
       title(Table 1) continuous(bpsystol weight, stat(p25 p75))
       continuous(age, stat(min max))
       nformat(%12.0f min max) nformat(%12.2f p25 p75)

Table 1
------------------------------------------------
                            Diabetes status
                     Not diabetic    Diabetic
------------------------------------------------
N                     9,850 (95.2%)    499 (4.8%)
Systolic blood pressure 114.00 140.00 126.00 160.00
Age (years)                 20 74          22 74
Weight (kg)               60.56 80.85   64.41 86.98
Race
  White               8,659 (87.9%)   404 (81.0%)
  Black               1,000 (10.2%)    86 (17.2%)
  Other                  191 (1.9%)     9 (1.8%)
Health status
  Excellent           2,383 (24.2%)    24 (4.8%)
  Very good           2,546 (25.9%)    45 (9.0%)
  Good                2,805 (28.5%)   133 (26.7%)
  Fair                1,508 (15.3%)   162 (32.5%)
  Poor                   594 (6.0%)   135 (27.1%)
------------------------------------------------
```

Test variables for equality between groups

You can customize the descriptive table by testing each variable for equality between groups.

```
. dtable bpsystol age weight i.race i.hlthstat, by(diabetes, nototals tests)
      nformat(%12.2f mean sd) title(Table 1. Summary by group with a test.)
note: using test regress across levels of diabetes for bpsystol, age, and weight.
note: using test pearson across levels of diabetes for race and hlthstat.

Table 1. Summary by group with a test.
------------------------------------------------------------
                                Diabetes status
                        Not diabetic     Diabetic     Test
------------------------------------------------------------
N                        9,850 (95.2%)     499 (4.8%)
Systolic blood pressure 130.09 (22.76) 146.65 (28.39) <0.001
Age (years)              46.92 (17.19)  60.69 (11.47) <0.001
Weight (kg)              71.66 (15.22)  76.67 (17.18) <0.001
Race
  White                  8,659 (87.9%)    404 (81.0%) <0.001
  Black                  1,000 (10.2%)     86 (17.2%)
  Other                     191 (1.9%)       9 (1.8%)
Health status
  Excellent              2,383 (24.2%)      24 (4.8%) <0.001
  Very good              2,546 (25.9%)      45 (9.0%)
  Good                   2,805 (28.5%)    133 (26.7%)
  Fair                   1,508 (15.3%)    162 (32.5%)
  Poor                      594 (6.0%)    135 (27.1%)
------------------------------------------------------------
```

Test variables for equality between groups

ctest	Description
regress	main effects test from a linear regression
poisson	main effects test from a Poisson regression
lnormal	main effects test from a log-normal regression
kwallis	Kruskal-Wallis rank test

ftest	Description
pearson	Pearson's χ^2 test
fisher	Fisher's exact test
lrchi2	likelihood-ratio χ^2 test
gamma	Goodman and Kruskal's gamma
kendall	Kendall's τ_b
cramer	Cramér's V
svylr	survey-adjusted likelihood-ratio test
svywald	survey-adjusted Wald test
svyllwald	survey-adjusted log-linear Wald test

Test variables for equality between groups

```
. dtable bpsystol age weight i.race, by(diabetes, nototals)
      continuous(, test(kwallis)) factor(, test(pearson))
      sample(Sample, statistics(frequency) place(seplabels))
      sformat("n=%s" frequency)
      title(Table 1. Summary by group with a test.)
      note(Kruskal-Wallis test for continuous variables.)
      note(Pearson's test for factor variables.)

Table 1. Summary by group with a test.
------------------------------------------------------------
                                Diabetes status
                          Not diabetic       Diabetic
                             n=9,850           n=499
------------------------------------------------------------
Systolic blood pressure 130.088 (22.759) 146.651 (28.387)
Age (years)              46.918 (17.193)  60.687 (11.475)
Weight (kg)              71.658 (15.220)  76.670 (17.175)
Race
  White                   8,659 (87.9%)      404 (81.0%)
  Black                   1,000 (10.2%)       86 (17.2%)
  Other                      191 (1.9%)         9 (1.8%)
------------------------------------------------------------
Kruskal-Wallis test for continuous variables.
Pearson's test for factor variables.
```

Save and reuse table styles

```
. collect style save mydtable, replace
(style from DTable saved to file mydtable.stjson)

. set dtable_style mydtable

. sysuse auto, clear
(1978 automobile data)

. dtable mpg weight i.rep78, by(foreign)

Table 1. Summary by group with a test.
------------------------------------------------------------
                                  Car origin
                          Domestic             Foreign
                            n=52                 n=22
------------------------------------------------------------
Weight (lbs.)       3,317.115 (695.364) 2,315.909 (433.003)
Mileage (mpg)            19.827 (4.743)      24.773 (6.611)
Repair record 1978
  1                         2 (4.2%)             0 (0.0%)
  2                        8 (16.7%)             0 (0.0%)
  3                       27 (56.2%)            3 (14.3%)
  4                        9 (18.8%)            9 (42.9%)
  5                         2 (4.2%)            9 (42.9%)
------------------------------------------------------------
Kruskal-Wallis test for continuous variables.
Pearson's test for factor variables.
```

Create customizable tables with etable

etable allows you to easily create:

- A standard estimation table with current estimation results.
- A table using results from a **margins** command.
- A table using results from stored estimates by **estimates store**.
- Add model statistics and coefficient-specific statistics.
- Add customizable stars to label statistically significant results.

Syntax:

```
etable [, estimates(namelist) margins
    cstat(cstat) mstat(mstat) stars([statspec])
    export(filename) options]
```

A quick example

First, we fit a logistic regression model on the second National Health and Nutrition Examination Survey data (NHANES II).

```
. webuse nhanes2l, clear
(Second National Health and Nutrition Examination Survey)

. logistic diabetes i.agegrp bmi i.highbp

Logistic regression                                     Number of obs = 10,349
                                                        LR chi2(7)    = 416.01
                                                        Prob > chi2   = 0.0000
Log likelihood = -1791.7561                             Pseudo R2     = 0.1040

------------------------------------------------------------------------------
    diabetes | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
      agegrp |
      30-39  |   1.730448   .5895795     1.61   0.108     .8874554    3.374199
      40-49  |   4.259599   1.297735     4.76   0.000     2.344448    7.739213
      50-59  |   6.888277   1.993273     6.67   0.000     3.906582    12.14575
      60-69  |   10.88779   2.952805     8.80   0.000     6.398693    18.52629
        70+  |   15.25109   4.308098     9.65   0.000     8.767088    26.53057
             |
         bmi |    1.07177   .0091357     8.13   0.000     1.054014    1.089826
    1.highbp |   1.251171   .1290527     2.17   0.030     1.022161    1.531491
       _cons |   .0011192   .0003767   -20.19   0.000     .0005787    .0021646
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```

A quick example

By default, **etable** creates a table of coefficients, standard errors, and the number of observations.

```
. etable

------------------------------
                        diabetes
------------------------------
Age group
  30-39                   1.730
                         (0.590)
  40-49                   4.260
                         (1.298)
  50-59                   6.888
                         (1.993)
  60-69                  10.888
                         (2.953)
  70+                    15.251
                         (4.308)
Body mass index (BMI)     1.072
                         (0.009)
High blood pressure
  1                       1.251
                         (0.129)
Intercept                 0.001
                         (0.000)
Number of observations    10349
------------------------------
```

Display coefficient statistics

You can use **cstat()** option to specify which statistics to display in the table.

```
. etable, cstat(_r_b) cstat(_r_ci)

----------------------------------------
                            diabetes
----------------------------------------
Age group
  30-39                          1.730
                       [0.887     3.374]
  40-49                          4.260
                       [2.344     7.739]
  50-59                          6.888
                       [3.907    12.146]
  60-69                         10.888
                       [6.399    18.526]
  70+                           15.251
                       [8.767    26.531]
Body mass index (BMI)            1.072
                       [1.054     1.090]
High blood pressure
  1                              1.251
                       [1.022     1.531]
Intercept                        0.001
                       [0.001     0.002]
Number of observations           10349
----------------------------------------
```

Coefficient statistics

cstat	Description
_r_b	coefficient reported by estimation
_r_se	standard errors of _r_b
_r_z	test statistics for _r_b
_r_z_abs	absolute values of _r_z
_r_p	*p*-values for _r_b
_r_lb	lower bounds of confidence intervals for _r_b
_r_ub	upper bounds of confidence intervals for _r_b
_r_ci	confidence inters for _r_b
_r_crlb	lower bounds of credible intervals for _r_b
_r_crub	upper bounds of credible intervals for _r_b
_r_cri	credible intervals of Bayesian estimates

Display model statistics

You can use **mstat()** option to specify which model statistics to display in the table.

```
. etable, replay mstat(N) mstat(aic) mstat(bic)

----------------------------------------
                            diabetes
----------------------------------------
Age group
  30-39                          1.730
                       [0.887     3.374]
  40-49                          4.260
                       [2.344     7.739]
  50-59                          6.888
                       [3.907    12.146]
  60-69                         10.888
                       [6.399    18.526]
  70+                           15.251
                       [8.767    26.531]
Body mass index (BMI)            1.072
                       [1.054     1.090]
High blood pressure
  1                              1.251
                       [1.022     1.531]
Intercept                        0.001
                       [0.001     0.002]
Number of observations           10349
AIC                            3599.51
BIC                            3657.47
----------------------------------------
```

Model statistics

mstat	Description
N	number of observations
aic	Akaike's information criterion
bic	Schwarz's Bayesian information criterion
F	F statistic
chi2	χ^2
ll	log likelihood of fitted model
r2	R^2
r2_a	adjusted R^2
rank	rank of fitted model
scalar	any e() scalar

Customize table

You can further customize the table by formatting the results, labeling statistically significant results, adding a title and notes, and more.

```
. etable, replay cstat(_r_b, nformat(%4.2f)) cstat(_r_ci, cidelimiter(,) nformat(%6.2f))
    showstars showstarsnote stars(.05 "*" .01 "**" .001 "***", attach(_r_b))
    title("Table 2: Logistic Regression Model For Diabetes")

Table 2: Logistic Regression Model For Diabetes
----------------------------------------
                            diabetes
----------------------------------------
Age group
  30-39                      1.73
                      [0.89,  3.37]
  40-49                      4.26 ***
                      [2.34,  7.74]
  50-59                      6.89 ***
                      [3.91, 12.15]
  60-69                     10.89 ***
                      [6.40, 18.53]
  70+                       15.25 ***
                      [8.77, 26.53]
Body mass index (BMI)        1.07 ***
                      [1.05,  1.09]
High blood pressure
  1                          1.25 *
                      [1.02,  1.53]
Intercept                    0.00 ***
                      [0.00,  0.00]
Number of observations      10349
AIC                       3599.51
BIC                       3657.47
----------------------------------------
*** p<.001, ** p<.01, * p<.05
```

Tables with more than one estimation result

etable can be used to create a table comparing different estimation results. After fitting each model, you can save the results from each model with **estimates store** and specify which model results to include in the table with the **estimates()** option.

```
. quietly logistic diabetes i.agegrp bmi
. estimates store model1

. quietly logistic diabetes i.agegrp
. estimates store model2

. quietly logistic diabetes bmi
. estimates store model3

. etable, estimates(model1 model2 model3)

-------------------------------------------------
                     diabetes diabetes diabetes
-------------------------------------------------
Age group
  30-39                 1.770    2.017
                       (0.603)  (0.685)
  40-49                 4.452    5.251
                       (1.353)  (1.590)
  50-59                 7.405    9.076
                       (2.128)  (2.596)
  60-69                11.814   13.948
                       (3.172)  (3.735)
  70+                  16.841   19.494
                       (4.694)  (5.418)
Body mass index (BMI)   1.077            1.089
                       (0.009)          (0.008)
Intercept               0.001    0.007   0.005
                       (0.000)  (0.002) (0.001)
Number of observations  10349    10349   10349
-------------------------------------------------
```

Tables with more than one estimation result

etable provides a lot of customization options that can be used to create a publication-quality table quickly.

```
. etable, estimates(model1 model2 model3)
          column(index)
          cstat(_r_b, nformat(%4.2f))
          cstat(_r_ci, cidelimiter(,) nformat(%6.2f))
          showstars showstarsnote
          stars(.05 "*" .01 "**" .001 "***", attach(_r_b))
          mstat(N,   nformat(%8.0fc) label("Observations"))
          mstat(aic, nformat(%5.0f))
          mstat(bic, nformat(%5.0f))
          mstat(r2_p, nformat(%5.4f) label("Pseudo R2"))
          title(Table 2: Logistic Regression Model For Diabetes)
          note(Data Source: NHANES, 1981)
          note(Model 1: agegrp bmi)
          note(Model 2: agegrp)
          note(Model 3: bmi)
          export(mytable2.docx, replace)
```

Tables with more than one estimation result

Table 2: Logistic Regression Model For Diabetes

	1	2	3
Age group			
30-39	1.77	2.02 *	
	[0.91, 3.45]	[1.04, 3.92]	
40-49	4.45 ***	5.25 ***	
	[2.45, 8.08]	[2.90, 9.51]	
50-59	7.41 ***	9.08 ***	
	[4.22, 13.01]	[5.18, 15.90]	
60-69	11.81 ***	13.95 ***	
	[6.98, 20.00]	[8.25, 23.57]	
70+	16.84 ***	19.49 ***	
	[9.75, 29.08]	[11.31, 33.61]	
Body mass index (BMI)	1.08 ***		1.09 ***
	[1.06, 1.09]		[1.07, 1.10]
Intercept	0.00 ***	0.01 ***	0.01 ***
	[0.00, 0.00]	[0.00, 0.01]	[0.00, 0.01]
Observations	10,349	10,349	10,349
AIC	3602	3675	3892
BIC	3653	3718	3906
Pseudo R2	0.1028	0.0841	0.0279

*** p<.001, ** p<.01, * p<.05
Data Source: NHANES, 1981
Model 1: agegrp bmi
Model 2: agegrp
Model 3: bmi
(collection ETable exported to file mytable2.docx)

Create customizable tables with table

table was redesigned in Stata 17 to be more powerful and flexible. It can be used to create:

- Table of one-way, two-way, and multiway tabulations
- Table of summary statistics (**dtable**)
- Table of estimation results (**etable**)
- Table of results from other Stata commands

Tabulations

One-way tabulation:

```
table (rowvar) ()
table () (colvar)
```

Two-way tabulation:

```
table (rowvar) (colvar)
```

Multiway tabulation:

```
table (rowvars) (colvars) [(tabvars)]
```

One-way tabulation

```
. webuse nhanes2l, clear
(Second National Health and Nutrition Examination Survey)

. table (hlthstat) ()

--------------------------
              |  Frequency
--------------+-----------
Health status |
  Excellent   |      2,407
  Very good   |      2,591
  Good        |      2,938
  Fair        |      1,670
  Poor        |        729
  Total       |     10,335
--------------------------

. table () (hlthstat)

-----------------------------------------------------------------------
          |                        Health status
          |  Excellent   Very good    Good    Fair    Poor     Total
----------+------------------------------------------------------------
Frequency |      2,407       2,591   2,938   1,670     729    10,335
-----------------------------------------------------------------------
```

Two-way tabulation

```
. table (hlthstat) (region)

----------------------------------------------------
              |                  Region
              |     NE      MW       S       W    Total
--------------+-------------------------------------
Health status |
  Excellent   |    562     730     546     569    2,407
  Very good   |    558     721     651     661    2,591
  Good        |    631     735     807     765    2,938
  Fair        |    257     419     532     462    1,670
  Poor        |     77     167     317     168      729
  Total       |  2,085   2,772   2,853   2,625   10,335
----------------------------------------------------

. table (region) (hlthstat)

------------------------------------------------------------------
        |                      Health status
        |  Excellent   Very good    Good    Fair    Poor     Total
--------+---------------------------------------------------------
Region  |
  NE    |        562         558     631     257      77     2,085
  MW    |        730         721     735     419     167     2,772
  S     |        546         651     807     532     317     2,853
  W     |        569         661     765     462     168     2,625
  Total |      2,407       2,591   2,938   1,670     729    10,335
------------------------------------------------------------------
```

Two-way tabulation

```
. table (hlthstat) (region),
      statistic(frequency)
      statistic(percent, across(hlthstat))
      nototals
```

	Region			
	NE	MW	S	W
Health status				
Excellent				
Frequency	562	730	546	569
Percent	26.95	26.33	19.14	21.68
Very good				
Frequency	558	721	651	661
Percent	26.76	26.01	22.82	25.18
Good				
Frequency	631	735	807	765
Percent	30.26	26.52	28.29	29.14
Fair				
Frequency	257	419	532	462
Percent	12.33	15.12	18.65	17.60
Poor				
Frequency	77	167	317	168
Percent	3.69	6.02	11.11	6.40

Multiway tabulation

```
. table (agegrp) (sex highbp)

----------------------------------------------------------------------------------
          |                                  Sex
          |           Male                 Female                   Total
          |  High blood pressure      High blood pressure      High blood pressure
          |     0        1   Total       0        1   Total       0        1    Total
----------+-----------------------------------------------------------------------
Age group |
  20-29   |   825      291   1,116   1,103      101   1,204   1,928      392   2,320
  30-39   |   480      290     770     687      165     852   1,167      455   1,622
  40-49   |   336      274     610     434      228     662     770      502   1,272
  50-59   |   255      347     602     335      354     689     590      701   1,291
  60-69   |   568      801   1,369     625      866   1,491   1,193    1,667   2,860
  70+     |   147      301     448     180      358     538     327      659     986
  Total   | 2,611    2,304   4,915   3,364    2,072   5,436   5,975    4,376  10,351
----------------------------------------------------------------------------------

. table (agegrp) (sex highbp), nototals

-----------------------------------------------------
          |                  Sex
          |           Male                 Female
          |  High blood pressure      High blood pressure
          |     0          1              0          1
----------+------------------------------------------
Age group |
  20-29   |   825        291          1,103        101
  30-39   |   480        290            687        165
  40-49   |   336        274            434        228
  50-59   |   255        347            335        354
  60-69   |   568        801            625        866
  70+     |   147        301            180        358
-----------------------------------------------------
```

Multiway tabulation

```
. label define yesno 0 "No" 1 "Yes"
. label values highbp diabetes heartatk yesno
. label variable diabetes "Diabetes"
. table (sex agegrp) (diabetes), nototals
      statistic(percent, across(diabetes))
```

	Diabetes	
	No	Yes
Sex		
Male		
Age group		
20-29	99.64	0.36
30-39	99.61	0.39
40-49	97.38	2.62
50-59	94.68	5.32
60-69	91.96	8.04
70+	88.39	11.61
Female		
Age group		
20-29	99.09	0.91
30-39	97.88	2.12
40-49	96.07	3.93
50-59	94.19	5.81
60-69	91.42	8.58
70+	89.03	10.97

Results from other commands

table can also be used to create tables with results from other commands:

```
table (rowspec) (colspec) [,
        command(cmdspec) [command(cmdspec)] ...
]
```

cmdspec is ***explist*:] *command***. *explist* can be results stored in **r()** and **e()** by the *command*, or named expressions *name* = *exp*.

Results from other commands

```
. sysuse auto, clear
(1978 automobile data)

. ttest mpg, by(foreign)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
Domestic |      52    19.82692     .657777    4.743297    18.50638    21.14747
 Foreign |      22    24.77273     1.40951    6.611187    21.84149    27.70396
---------+--------------------------------------------------------------------
Combined |      74     21.2973    .6725511    5.785503     19.9569    22.63769
---------+--------------------------------------------------------------------
    diff |           -4.945804    1.362162               -7.661225   -2.230384
------------------------------------------------------------------------------
    diff = mean(Domestic) - mean(Foreign)                         t =  -3.6308
H0: diff = 0                                     Degrees of freedom =       72

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0003         Pr(|T| > |t|) = 0.0005          Pr(T > t) = 0.9997

. return list

scalars:
              r(level) =  95
                 r(sd) =  5.785503209735139
               r(sd_2) =  6.611186898567625
               r(sd_1) =  4.743297247514701
                 r(se) =  1.362162113622176
                r(p_u) =  .9997372920330829
                r(p_l) =  .0002627079669171
                  r(p) =  .0005254159338342
                  r(t) =  -3.630848447731832
               r(df_t) =  72
               r(mu_2) =  24.77272727272727
                r(N_2) =  22
               r(mu_1) =  19.82692307692308
                r(N_1) =  52
```

Results from other commands

We want to create a table with the command used, average mileage per gallon for domestic and foreign cars, the difference between the two group means, and the *p*-value.

Those values are stored in r(mu_1), r(mu_2), and r(p) returned by the **ttest** command.

```
. table (command) (result),
      command(Domestic = r(mu_1)
              Foreign  = r(mu_2)
              Diff = (r(mu_2)-r(mu_1))
              pvalue = r(p)
              : ttest mpg, by(foreign))

------------------------------------------------------------------------
                        |  Domestic    Foreign       Diff     pvalue
------------------------+-----------------------------------------------
ttest mpg, by(foreign) |  19.82692   24.77273   4.945804   .0005254
------------------------------------------------------------------------
```

Results from other commands

table can also create a table with results from multiple commands.

```
. local columns "Domestic = r(mu_1) Foreign  = r(mu_2) Diff = (r(mu_2)-r(mu_1)) pvalue = r(p)"
.
. table (command) (result),
      command(‘columns’: ttest mpg, by(foreign))
      command(‘columns’: ttest price, by(foreign))
      command(‘columns’: ttest weight, by(foreign))

------------------------------------------------------------------------------
                           |  Domestic    Foreign        Diff     pvalue
---------------------------+--------------------------------------------------
ttest mpg, by(foreign)     |  19.82692   24.77273    4.945804   .0005254
ttest price, by(foreign)   |  6072.423   6384.682    312.2587   .6801851
ttest weight, by(foreign)  |  3317.115   2315.909   -1001.206   2.62e-08
------------------------------------------------------------------------------
```

```
. table (colname result) (command),
      command(_r_b _r_se: regress price mpg weight)
      command(_r_b _r_se: regress price mpg weight i.foreign)
      nformat(%6.2f) sformat("(%s)" _r_se) style(table-reg3)

--------------------------------------------
                     |          1           2
---------------------+----------------------
      Mileage (mpg)  |    -49.51      21.85
                     |   (86.16)     (74.22)
                     |
      Weight (lbs.)  |     1.75        3.46
                     |    (0.64)      (0.63)
                     |
Car origin=Domestic  |                 0.00
                     |                (0.00)
 Car origin=Foreign  |               3673.06
                     |               (683.98)
                     |
          Intercept  |   1946.07     -5853.70
                     |  (3597.05)   (3376.99)
--------------------------------------------
```

Create customizable tables with collect

The **collect** suite of commands provide a convenient way to create customizable tables:

- Collect results from one or more Stata commands.
- Lay out the results into a one-way, two-way, or multiway table.
- Customize the table by changing formats of results, row and column headers, title and notes, etc.
- Export the finalized table to other document formats.
- Save table styles it a style file and reuse them to other tables.

See [TABLES] Customizable Tables for more information.

Create customizable tables with collect

Collect results from Stata commands:

```
. sysuse auto
(1978 automobile data)

. collect _r_b _r_ci: regress mpg weight price i.foreign

      Source |       SS           df       MS      Number of obs   =        74
-------------+----------------------------------   F(3, 70)        =     45.93
       Model |  1620.30716         3  540.102388   Prob > F        =    0.0000
    Residual |  823.152295        70  11.7593185   R-squared       =    0.6631
-------------+----------------------------------   Adj R-squared   =    0.6487
       Total |  2443.45946        73  33.4720474   Root MSE        =    3.4292

------------------------------------------------------------------------------
         mpg | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      weight |  -.0067758   .0009048    -7.49   0.000    -.0085805   -.0049712
       price |   .0000566   .0001922     0.29   0.769    -.0003268       .00044
             |
     foreign |
    Foreign  |  -1.855891   1.289063    -1.44   0.154    -4.426846    .7150641
       _cons |   41.95948   2.377726    17.65   0.000     37.21725     46.7017
------------------------------------------------------------------------------
```

Create customizable tables with collect

Lay out the results in a table:

```
. collect layout (colname) (result)

Collection: default
      Rows: colname
   Columns: result
   Table 1: 5 x 2

------------------------------------------------
              | Coefficient        95% CI
--------------+---------------------------------
Weight (lbs.) |   -.0067758 -.0085805 -.0049712
Price         |    .0000566 -.0003268    .00044
Domestic      |           0
Foreign       |   -1.855891 -4.426846  .7150641
Intercept     |    41.95948  37.21725   46.7017
------------------------------------------------
```

Create customizable tables with collect

Customize the table:

```
. collect style showbase off

. collect style cell, nformat(%6.2f)

. collect style cell result[_r_ci], sformat("[%s]") cidelimiter(", ")

. collect stars _r_p 0.01 "***" 0.05 "** " 0.1 "* " 1 " ", attach(_r_b)

. collect notes : "*** p<.01, ** p<.05, * p<.1"

. collect title "Table 1: Regression results"

. collect preview

Table 1: Regression results
------------------------------------------
              | Coefficient      95% CI
--------------+---------------------------
Weight (lbs.) |    -0.01*** [-0.01,  -0.00]
Price         |       0.00  [-0.00,   0.00]
Foreign       |      -1.86  [-4.43,   0.72]
Intercept     |    41.96*** [37.22,  46.70]
------------------------------------------
*** p<.01, ** p<.05, * p<.1
```

Create customizable tables with collect

Export the table to other document formats:

```
. collect export mytable3.docx, replace
(collection default exported to file mytable3.docx)
```

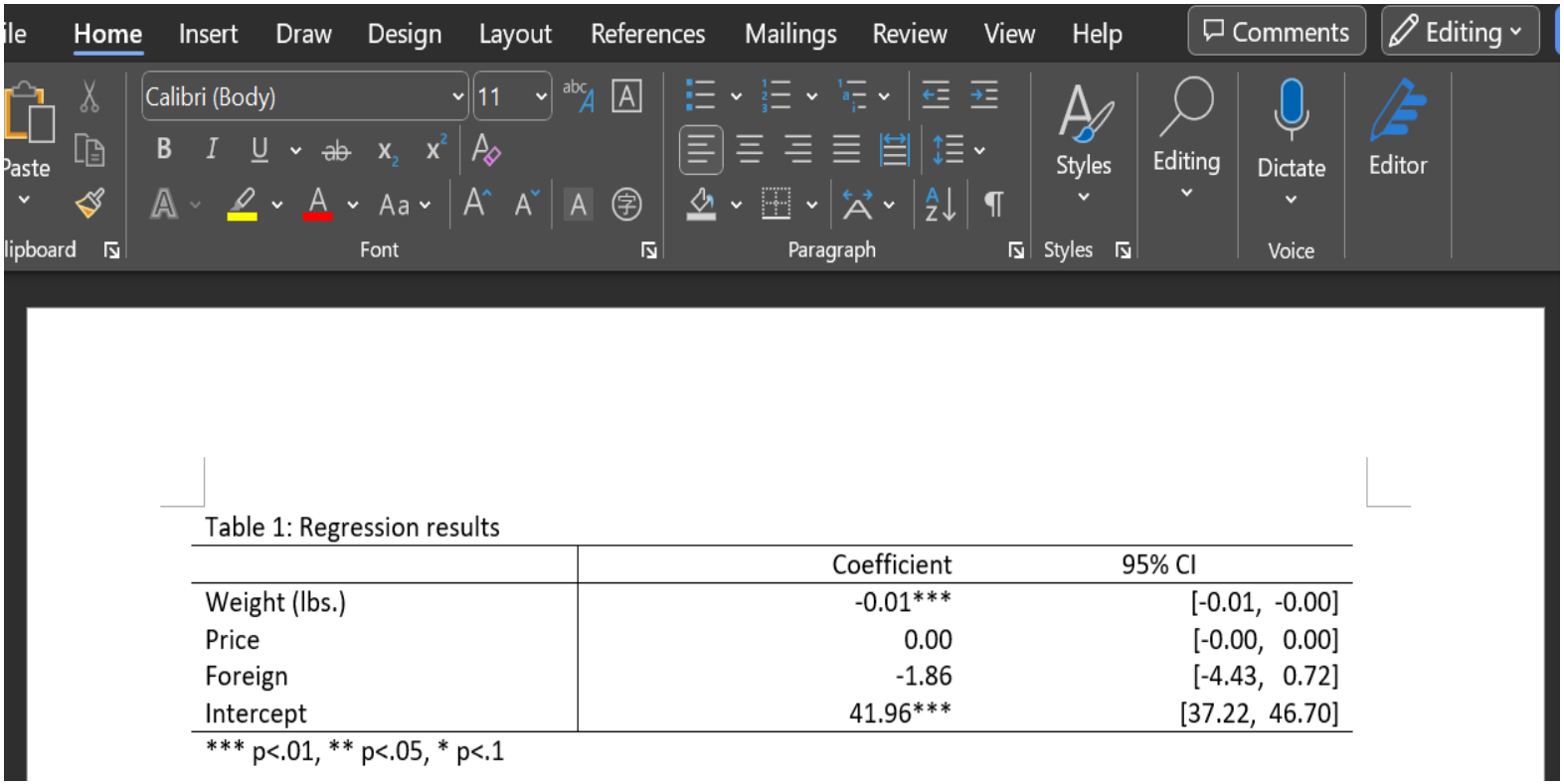


Table 1: Regression results

	Coefficient	95% CI
Weight (lbs.)	-0.01***	[-0.01, -0.00]
Price	0.00	[-0.00, 0.00]
Foreign	-1.86	[-4.43, 0.72]
Intercept	41.96***	[37.22, 46.70]

*** p<.01, ** p<.05, * p<.1

Create customizable tables with collect

Save and reuse table styles:

```
. collect style save mystyle.stjson
(style from default saved to file mystyle.stjson)

. quietly collect _r_b _r_ci: logit foreign mpg price
. collect layout (colname) (result)

Collection: default
      Rows: colname
   Columns: result
   Table 1: 3 x 2

------------------------------------------------
              | Coefficient        95% CI
--------------+---------------------------------
Mileage (mpg) |    .2338353   .1022338   .3654368
Price         |     .000266   .0000375   .0004945
Intercept     |   -7.648111 -11.65364 -3.642586
------------------------------------------------

. collect style use mystyle.stjson, replace

Collection: default
      Rows: colname
   Columns: result
   Table 1: 3 x 2

. collect preview

Table 1: Regression results
-----------------------------------------
              | Coefficient      95% CI
--------------+--------------------------
Mileage (mpg) |     0.23***    [0.10,   0.37]
Price         |     0.00**     [0.00,   0.00]
Intercept     |    -7.65*** [-11.65,  -3.64]
-----------------------------------------
*** p<.01, ** p<.05, * p<.1
```

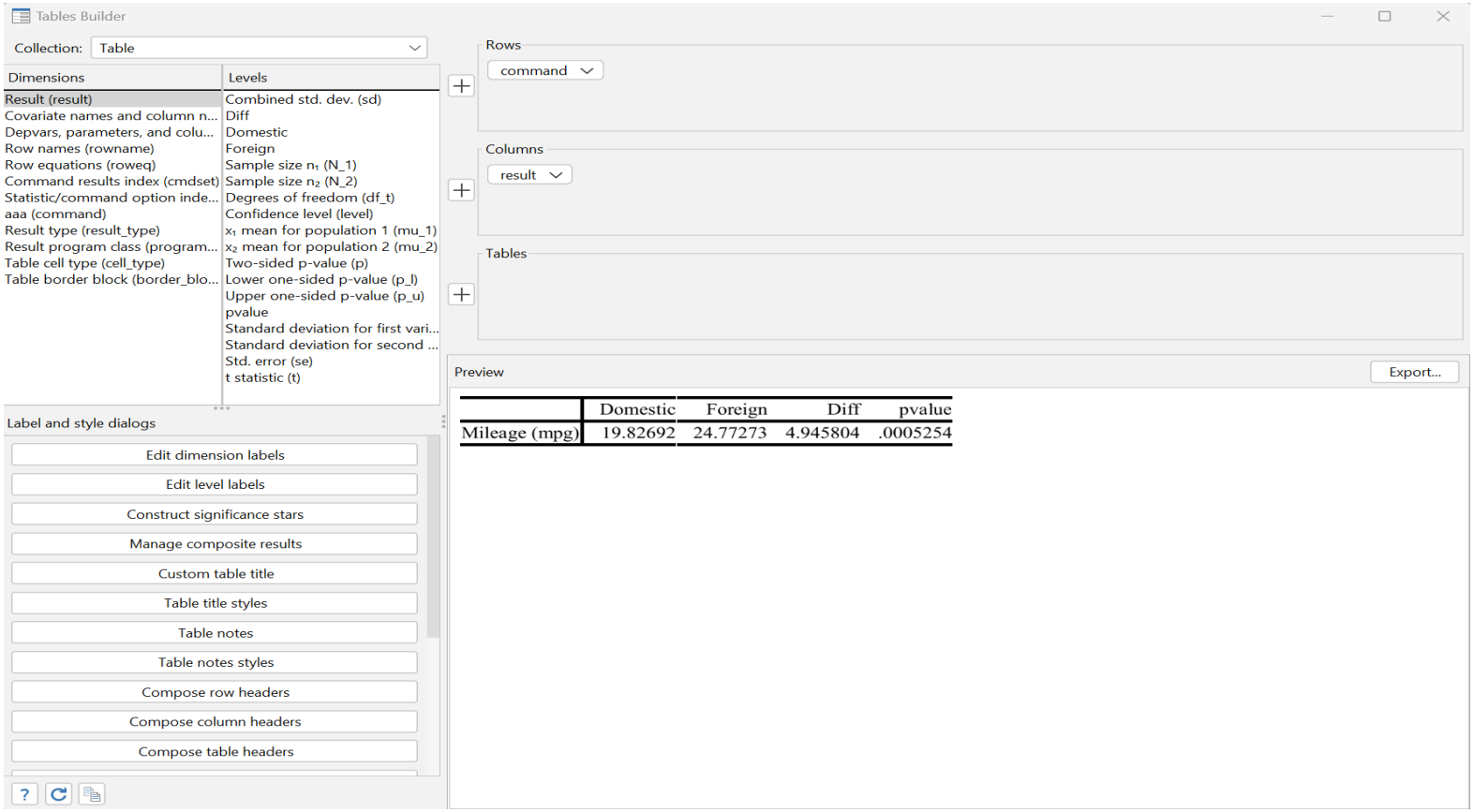
Customize tables using collect

collect can also be used to further customize tables created by **dtable**, **etable**, and **table**.

```
. table (command) (result),
      command(Domestic = r(mu_1)
              Foreign  = r(mu_2)
              Diff = (r(mu_2)-r(mu_1))
              pvalue = r(p)
              : ttest mpg, by(foreign))

------------------------------------------------------------------------
                       |  Domestic    Foreign       Diff     pvalue
-----------------------+------------------------------------------------
ttest mpg, by(foreign) |  19.82692   24.77273   4.945804   .0005254
------------------------------------------------------------------------

. collect label levels command 1 "Mileage (mpg)", modify

. collect preview

-----------------------------------------------------------
              |  Domestic    Foreign       Diff     pvalue
--------------+--------------------------------------------
Mileage (mpg) |  19.82692   24.77273   4.945804   .0005254
-----------------------------------------------------------
```

Customize tables using collect

Tables Builder:




```
. db tables
```

Create reproducible reports

With **dtable**, **etable**, **table**, and **collect** commands, you can easily create reproducible reports that combine text with summary statistics, regression results, graphs, and other Stata results.

You can combine those commands with **putdocx**, **putpdf**, **putexcel**, and **dyndoc** to create such reports.

Create reproducible reports

Font Paragraph Styles Editing

A sample report

Introduction

In this report, we will use the 1978 automobile data to make a monthly report. Firstly, we generated a summary report of the variables we are interested as follows:

Table 1: Summary statistics

	Car origin		
	Domestic	Foreign	Total
N	52 (70.3%)	22 (29.7%)	74 (100.0%)
Weight (lbs.)	3317.12 (695.36)	2315.91 (433.00)	3019.46 (777.19)
Repair record 1978			
1	2 (4.2%)	0 (0.0%)	2 (2.9%)
2	8 (16.7%)	0 (0.0%)	8 (11.6%)
3	27 (56.2%)	3 (14.3%)	30 (43.5%)
4	9 (18.8%)	9 (42.9%)	18 (26.1%)
5	2 (4.2%)	9 (42.9%)	11 (15.9%)

Predictive Margins

We ran a linear regression model for the outcome mpg using the predictors weight and rep78 and estimated the predictive margins of rep78.

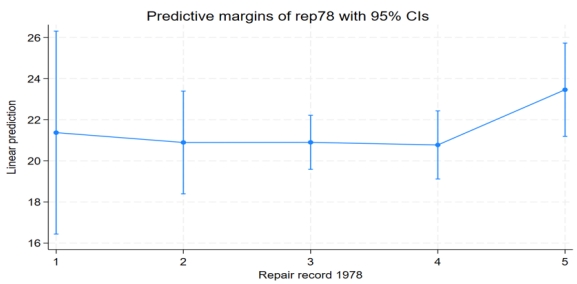


Figure 1: Predictive margins of repair record

Different regression models for automobile data

Lastly, we ran two different models to evaluate which one best fits the data.

Table 2: Regression Models For Automobile Data

	1		2	
Weight (lbs.)	-0.01	***	-0.01	***
	[-0.01, -0.00]		[-0.01, -0.00]	
Repair record 1978				
2	-0.48			
	[-6.00, 5.05]			
3	-0.47			
	[-5.57, 4.63]			
4	-0.60			
	[-5.81, 4.61]			
5	2.09			
	[-3.36, 7.53]			
Intercept	38.06	***	39.44	***
	[31.88, 44.24]		[36.22, 42.66]	
Observations	69		74	
AIC	374		395	
BIC	388		399	
Adjusted R2	0.6456		0.6467	

*** p<.001, ** p<.01, * p<.05
Data Source: 1978 automobile data
Model 1: rep78 weight
Model 2: weight

Create reproducible reports

```
putdocx clear
putdocx begin

sysuse auto, clear

// Title
putdocx paragraph, font(, 24) halign(center)
putdocx text ("A sample report")

// Introduction
putdocx paragraph, style(Heading1)
putdocx text ("Introduction")

putdocx textblock begin
In this report, we will use the 1978 automobile data to make a monthly report.
Firstly, we generated a summary report of the variables we are interested as
follows:
putdocx textblock end

dtable weight i.rep78, by(foreign)        ///
        nformat(%12.2f mean sd)                                  ///
        title("Table 1: Summary statistics")
putdocx collect

// Predictive margins of age group
putdocx paragraph, style(Heading1)
putdocx text ("Predictive Margins")

putdocx textblock begin
We ran a linear regression model for the outcome mpg using the predictors
weight and rep78 and estimated the predictive margins of rep78.
putdocx textblock end

regress mpg weight i.rep78
margins rep78
marginsplot

graph export margins.png, replace
```

Create reproducible reports

```
putdocx paragraph, halign(center)
putdocx image margins.png, height(3) width(6) linebreak
putdocx text ("Figure 1: Predictive margins of repair record"), bold

// Regression results of different models
putdocx paragraph, style(Heading1)
putdocx text ("Different regression models for automobile data")

putdocx textblock begin
Lastly, we ran two different models to evaluate which one best fits the
data.
putdocx textblock end

quietly regress mpg weight i.rep78
estimates store model1

quietly regress mpg weight
estimates store model2

etable, estimates(model1 model2)            ///
          column(index)                                          ///
          cstat(_r_b, nformat(%4.2f))                ///
          cstat(_r_ci, cidelimiter(,) nformat(%6.2f)) ///
          showstars showstarsnote                          ///
          stars(.05 "*" .01 "**" .001 "***", attach(_r_b)) ///
          mstat(N,   nformat(%8.0fc) label("Observations")) ///
          mstat(aic, nformat(%5.0f))                 ///
          mstat(bic, nformat(%5.0f))                 ///
          mstat(r2_a, nformat(%5.4f) label("Adjusted R2"))  ///
          title(Table 2: Regression Models For Automobile Data) ///
          note(Data Source: 1978 automobile data)        ///
          note(Model 1: rep78 weight)                        ///
          note(Model 2: weight)
putdocx collect

putdocx save myreport.docx, replace
```

Resources

- **dtable** - Create a table of descriptive statistics
  https://www.stata.com/manuals/rdtable.pdf
- **etable** - Create a table of estimation results
  https://www.stata.com/manuals/retable.pdf
- **table** - Table of frequencies, summaries, and command results
  https://www.stata.com/manuals/rtableintro.pdf
- **collect** - Stata Customizable Tables and Collected Results Reference
  https://www.stata.com/manuals/tables.pdf

Thank You!