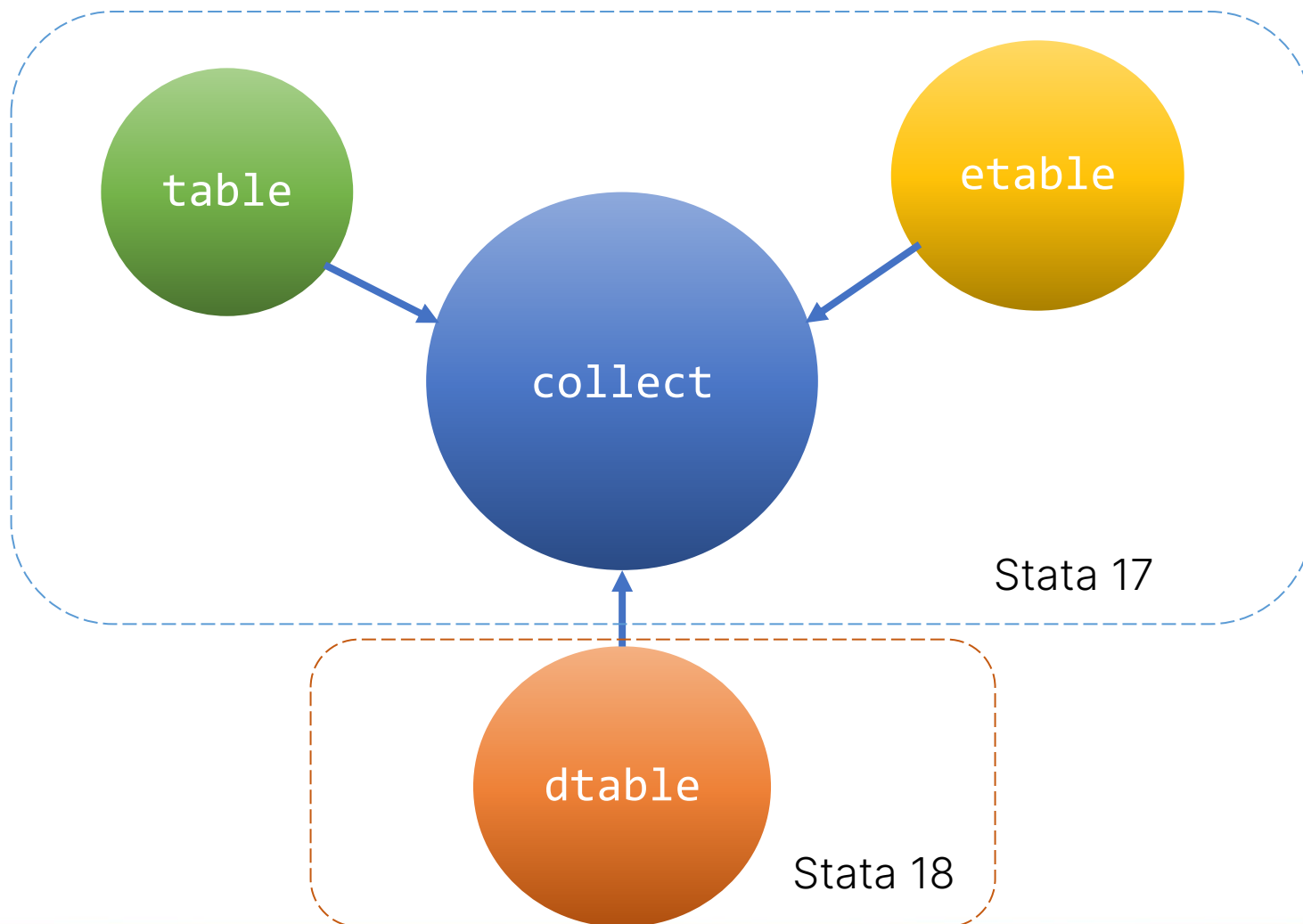


Tables of descriptive statistics

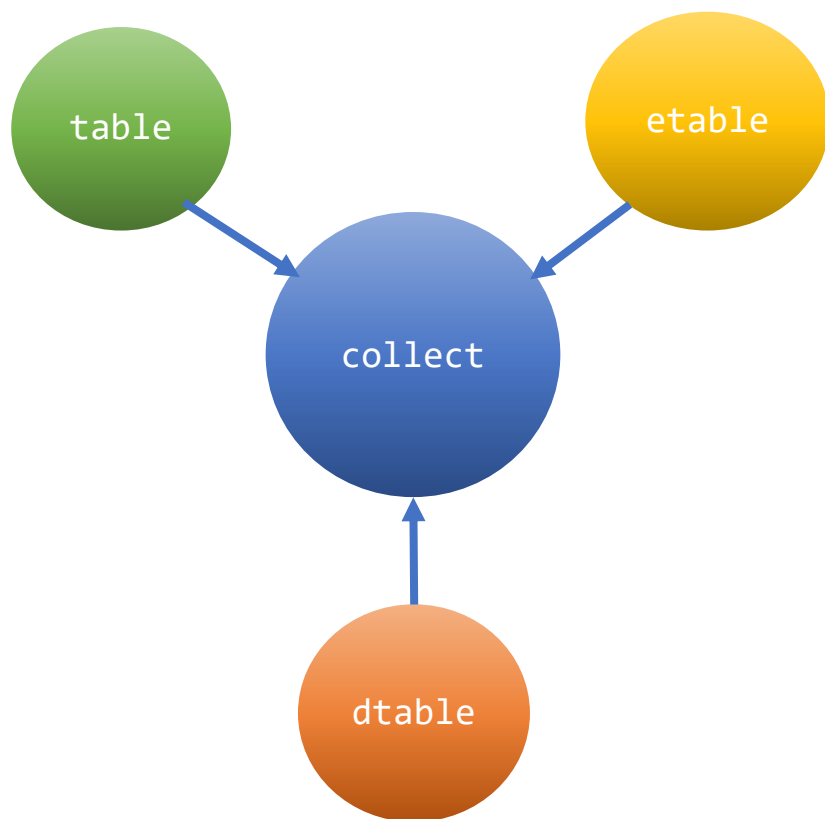
17 October 2023, 11:00 AM CDT

Mia Lv, StataCorp

Stata's table framework



Stata's table framework



- **collect**

commands to create, customize, and export tables

- **table**

tables of frequencies, summary statistics, and commands results

- **etable**

tables of estimation results

- **dtable**

tables of descriptive statistics

Convenience commands of
table and **collect**

Tables of descriptive statistics

- Present descriptive statistics of the data used in your research
- Often referred as “Table 1”

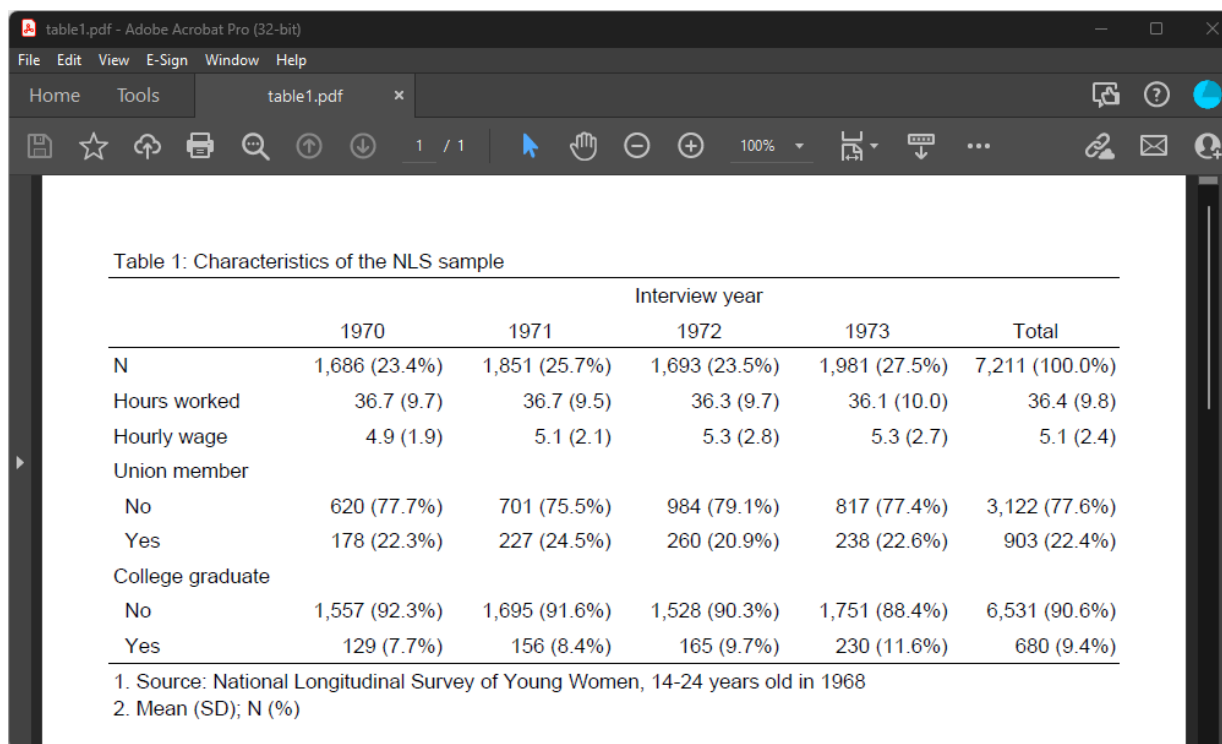


Table 1: Characteristics of the NLS sample

	Interview year				Total
	1970	1971	1972	1973	
N	1,686 (23.4%)	1,851 (25.7%)	1,693 (23.5%)	1,981 (27.5%)	7,211 (100.0%)
Hours worked	36.7 (9.7)	36.7 (9.5)	36.3 (9.7)	36.1 (10.0)	36.4 (9.8)
Hourly wage	4.9 (1.9)	5.1 (2.1)	5.3 (2.8)	5.3 (2.7)	5.1 (2.4)
Union member					
No	620 (77.7%)	701 (75.5%)	984 (79.1%)	817 (77.4%)	3,122 (77.6%)
Yes	178 (22.3%)	227 (24.5%)	260 (20.9%)	238 (22.6%)	903 (22.4%)
College graduate					
No	1,557 (92.3%)	1,695 (91.6%)	1,528 (90.3%)	1,751 (88.4%)	6,531 (90.6%)
Yes	129 (7.7%)	156 (8.4%)	165 (9.7%)	230 (11.6%)	680 (9.4%)

1. Source: National Longitudinal Survey of Young Women, 14-24 years old in 1968
2. Mean (SD); N (%)

More examples

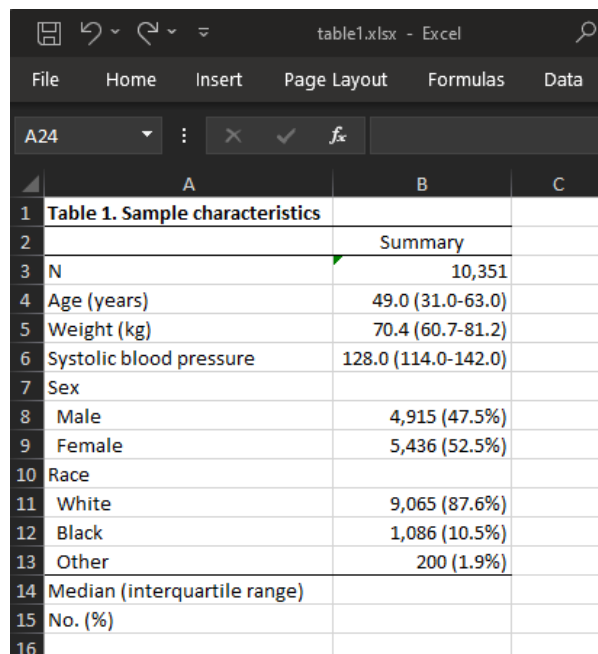


table1.xlsx - Excel

	A	B	C
1	Table 1. Sample characteristics		
2		Summary	
3	N	10,351	
4	Age (years)	49.0 (31.0-63.0)	
5	Weight (kg)	70.4 (60.7-81.2)	
6	Systolic blood pressure	128.0 (114.0-142.0)	
7	Sex		
8	Male	4,915 (47.5%)	
9	Female	5,436 (52.5%)	
10	Race		
11	White	9,065 (87.6%)	
12	Black	1,086 (10.5%)	
13	Other	200 (1.9%)	
14	Median (interquartile range)		
15	No. (%)		
16			

Table 1. Demographics

	Not diabetic (N=9,850)	Diabetic (N=499)
Age (years)	46.92 (17.19)	60.69 (11.47)
Weight (kg)	71.66 (15.22)	76.67 (17.18)
Systolic blood pressure	130.09 (22.76)	146.65 (28.39)
Sex		
Male	4,698 (47.7%)	217 (43.5%)
Female	5,152 (52.3%)	282 (56.5%)
Race		
White	8,659 (87.9%)	404 (81.0%)
Black	1,000 (10.2%)	86 (17.2%)
Other	191 (1.9%)	9 (1.8%)

Total sample: N = 10,349



table1 [Compatibility Mode] - Word

	Urban	Rural	All	P-value
N	79,965,794 (68.3%)	37,191,719 (31.7%)	117,157,513 (100.0%)	
Age (years)	41.81 ± 15.66	43.21 ± 15.11	42.25 ± 15.50	0.024
Weight (kg)	71.32 ± 15.37	73.14 ± 15.49	71.90 ± 15.43	<0.001
Systolic blood pressure	126.61 ± 21.44	127.68 ± 21.30	126.95 ± 21.40	0.406
Serum cholesterol (mg/dL)	213.30 ± 48.56	212.67 ± 48.22	213.10 ± 48.45	0.727

Table 1: Survey data summary

	Urban	Rural	All	P-value
N	79,965,794 (68.3%)	37,191,719 (31.7%)	117,157,513 (100.0%)	
Age (years)	41.81 ± 15.66	43.21 ± 15.11	42.25 ± 15.50	0.024
Weight (kg)	71.32 ± 15.37	73.14 ± 15.49	71.90 ± 15.43	<0.001
Systolic blood pressure	126.61 ± 21.44	127.68 ± 21.30	126.95 ± 21.40	0.406
Serum cholesterol (mg/dL)	213.30 ± 48.56	212.67 ± 48.22	213.10 ± 48.45	0.727

1. Mean ± SD; p-value from linear regression.
2. Statistics computed using the survey weights.
3. Tests adjusted for the survey design.

Tables of descriptive statistics before Stata 18

- Using **collect**
Examples:

<https://blog.stata.com/2021/06/24/customizable-tables-in-stata-17-part-3-the-classic-table-1/>

<https://www.statalist.org/forums/forum/general-stata-discussion/general/1719272-formatting-binary-and-categorical-variables-and-p-values-using-table>

- Community-contributed commands:
 - **table1** and **table1_mc**

Examples

- Example 1: The first try
- Example 2: The **by()**, **total()**, and **note()** options
- Example 3: The **sample()** option
- Example 4: The **continuous()** and **factor()** option
- Example 5: The **svy** and **subpop()** option
- Example 6: The **define()**, **nformat()**, and **sformat()** options
- Example 7: The **export()** option
- Example 8: Work together with **collect**
- Example 9: Using the dialog box

Example 1: The first try (1)

```
. sysuse auto, clear  
(1978 automobile data)  
. dtable price weight mpg i.rep78
```

```
-----  
                                Summary  
-----  
N                                74  
Price                6,165.257 (2,949.496)  
Weight (lbs.)        3,019.459 (777.194)  
Mileage (mpg)         21.297 (5.786)  
Repair record 1978  
  1                      2 (2.9%)  
  2                      8 (11.6%)  
  3                     30 (43.5%)  
  4                     18 (26.1%)  
  5                     11 (15.9%)  
-----
```


Example 1: The first try (2)

```
. sysuse auto, clear  
(1978 automobile data)  
. dtable price weight mpg i.rep78
```

Summary

N	74
---	----

Price	6,165.257 (2,949.496)
-------	-----------------------

Weight (lbs.)	3,019.459 (777.194)
---------------	---------------------

Mileage (mpg)	21.297 (5.786)
---------------	----------------

Repair record 1978	
--------------------	--

1	2 (2.9%)
---	----------

2	8 (11.6%)
---	-----------

3	30 (43.5%)
---	------------

4	18 (26.1%)
---	------------

5	11 (15.9%)
---	------------

Sample frequency statistics

Variable descriptive statistics

Example 1: The first try (3)

```
. sysuse auto, clear  
(1978 automobile data)  
. dtbl price weight mpg i.rep78
```

Summary

N 74

Price 6,165.257 (2,949.496)

Weight (lbs.) 3,019.459 (777.194)

Mileage (mpg) 21.297 (5.786)

Repair record 1978

1 2 (2.9%)

2 8 (11.6%)

3 30 (43.5%)

4 18 (26.1%)

5 11 (15.9%)

- Continuous variable default statistics:
mean (standard deviation)

Example 1: The first try (3)

```
. sysuse auto, clear  
(1978 automobile data)  
. dtbl price weight mpg i.rep78
```

Summary

N	74
Price	6,165.257 (2,949.496)
Weight (lbs.)	3,019.459 (777.194)
Mileage (mpg)	21.297 (5.786)

Repair record 1978

1	2 (2.9%)
2	8 (11.6%)
3	30 (43.5%)
4	18 (26.1%)
5	11 (15.9%)

- Continuous variable default statistics:

mean (standard deviation)

- **Factor** variable default statistics

frequency (percent)

Example 1: The first try (3)

```
. sysuse auto, clear  
(1978 automobile data)  
. dtbl price weight mpg i.rep78
```

```
-----  
                                Summary  
-----  
N                                74  
Price          6,165.257 (2,949.496)  
Weight (lbs.)   3,019.459 (777.194)  
Mileage (mpg)   21.297 (5.786)  
Repair record 1978  
1                2 (2.9%)  
2                8 (11.6%)  
3               30 (43.5%)  
4               18 (26.1%)  
5               11 (15.9%)  
-----
```

- Continuous variable default statistics:
mean (standard deviation)
- Factor variable default statistics
frequency (percent)
- Show **variable labels** by default instead of variable names

Example 1: The first try (4)

```
. dtable price weight mpg i.rep78, novarlabel
```

```
-----  
                        Summary  
-----  
N                                74  
price 6,165.257 (2,949.496)  
weight 3,019.459 (777.194)  
mpg    21.297 (5.786)  
rep78  
  1          2 (2.9%)  
  2          8 (11.6%)  
  3         30 (43.5%)  
  4         18 (26.1%)  
  5         11 (15.9%)  
-----
```

Show the variable names instead of variable labels.

Examples

Example 1: The first try

→ Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 2: The `by()`, `title()`, and `note()` options (1)

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

	Car origin		
	Domestic	Foreign	Total
N	52 (70.3%)	22 (29.7%)	74 (100.0%)
Price	6,072.423 (3,097.104)	6,384.682 (2,621.915)	6,165.257 (2,949.496)
Weight (lbs.)	3,317.115 (695.364)	2,315.909 (433.003)	3,019.459 (777.194)
Mileage (mpg)	19.827 (4.743)	24.773 (6.611)	21.297 (5.786)
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%)

Example 2: The `by()`, `title()`, and `note()` options (2)

```
. dtable price weight mpg i.rep78, by(foreign, nototal)
```

Car origin		
	Domestic	Foreign

N	52 (70.3%)	22 (29.7%)
Price	6,072.423 (3,097.104)	6,384.682 (2,621.915)
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%)

Example 2: The by(), title(), and note() options (3)

```
. dtable price weight mpg i.rep78, by(foreign, nototal tests)
```

note: using test regress across levels of foreign for price, weight, and mpg.

note: using test pearson across levels of foreign for rep78.

Not part of the table,
only displayed in
the result window

Car origin			
	Domestic	Foreign	Test

N	52 (70.3%)	22 (29.7%)	
Price	6,072.423 (3,097.104)	6,384.682 (2,621.915)	0.680
Weight (lbs.)	3,317.115 (695.364)	2,315.909 (433.003)	<0.001
Mileage (mpg)	19.827 (4.743)	24.773 (6.611)	<0.001
Repair record 1978			
1	2 (4.2%)	0 (0.0%)	<0.001
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%)	

Example 2: The `by()`, `title()`, and `note()` options (4)

```
. dtable price weight mpg i.rep78, by(foreign, nototal tests notestnotes)
```

Car origin			
	Domestic	Foreign	Test
N	52 (70.3%)	22 (29.7%)	
Price	6,072.423 (3,097.104)	6,384.682 (2,621.915)	0.680
Weight (lbs.)	3,317.115 (695.364)	2,315.909 (433.003)	<0.001
Mileage (mpg)	19.827 (4.743)	24.773 (6.611)	<0.001
Repair record 1978			
1	2 (4.2%)	0 (0.0%)	<0.001
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%)	

Example 2: The by(), title(), and note() options (5)

```
. replace foreign=. in 1/5
. dtable price weight mpg i.rep78, by(foreign, nototal tests notestnotes missing)
```

affected

	Car origin		.	Test
	Domestic	Foreign		
N	47 (63.5%)	22 (29.7%)	5 (6.8%)	
Price	6,180.340 (3,207.247)	6,384.682 (2,621.915)	5,058.000 (1,606.718)	0.667
Weight (lbs.)	3,324.255 (714.169)	2,315.909 (433.003)	3,250.000 (541.618)	<0.001
Mileage (mpg)	19.894 (4.904)	24.773 (6.611)	19.200 (3.114)	0.002
Repair record 1978				
1	2 (4.5%)	0 (0.0%)	0 (0.0%)	<0.001
2	8 (18.2%)	0 (0.0%)	0 (0.0%)	
3	24 (54.5%)	3 (14.3%)	3 (75.0%)	
4	8 (18.2%)	9 (42.9%)	1 (25.0%)	
5	2 (4.5%)	9 (42.9%)	0 (0.0%)	

Example 2: The by(), title(), and note() options (6)

```
. dtable price weight mpg i.rep78, by(foreign, nototal tests) title("Table 1: Descriptive stastics")  
note: using test regress across levels of foreign for price, weight, and mpg.  
note: using test pearson across levels of foreign for rep78.
```

Table 1: Descriptive stastics

	Car origin		
	Domestic	Foreign	Test
N	47 (68.1%)	22 (31.9%)	
Price	6,180.340 (3,207.247)	6,384.682 (2,621.915)	0.795
Weight (lbs.)	3,324.255 (714.169)	2,315.909 (433.003)	<0.001
Mileage (mpg)	19.894 (4.904)	24.773 (6.611)	0.001
Repair record 1978			
1	2 (4.5%)	0 (0.0%)	<0.001
2	8 (18.2%)	0 (0.0%)	
3	24 (54.5%)	3 (14.3%)	
4	8 (18.2%)	9 (42.9%)	
5	2 (4.5%)	9 (42.9%)	

Example 2: The by(), title(), and note() options (7)

```
. dtable price weight mpg i.rep78, by(foreign, nototal tests) title("Table 1: Descriptive stastics") ///  
> note("linear regression test performed for price weight mpg.") ///  
> note ("Pearson's chi-squared test for rep78.")
```

```
note: using test regress across levels of foreign for price, weight, and mpg.  
note: using test pearson across levels of foreign for rep78.
```

Not part of the table,
only displayed in the result window

Part of table

Table 1: Descriptive stastics

	Car origin		Test
	Domestic	Foreign	
N	47 (68.1%)	22 (31.9%)	
Price	6,180.340 (3,207.247)	6,384.682 (2,621.915)	0.795
Weight (lbs.)	3,324.255 (714.169)	2,315.909 (433.003)	<0.001
Mileage (mpg)	19.894 (4.904)	24.773 (6.611)	0.001
Repair record 1978			
1	2 (4.5%)	0 (0.0%)	<0.001
2	8 (18.2%)	0 (0.0%)	
3	24 (54.5%)	3 (14.3%)	
4	8 (18.2%)	9 (42.9%)	
5	2 (4.5%)	9 (42.9%)	

```
linear regression test performed for price weight mpg.  
Pearson's chi-squared test for rep78.
```

Multiple notes

Part of table

Example 2: The `by()`, `title()`, and `note()` options (8)

```
. dtable price weight mpg i.rep78, by(foreign, nototal tests) title("Table 1:  
Descriptive stastics") ///
```

```
> note("linear regression test performed for price weight mpg.") ///
```

```
> note ("Pearson's chi-squared test for rep78.") export(table.docx, replace)
```

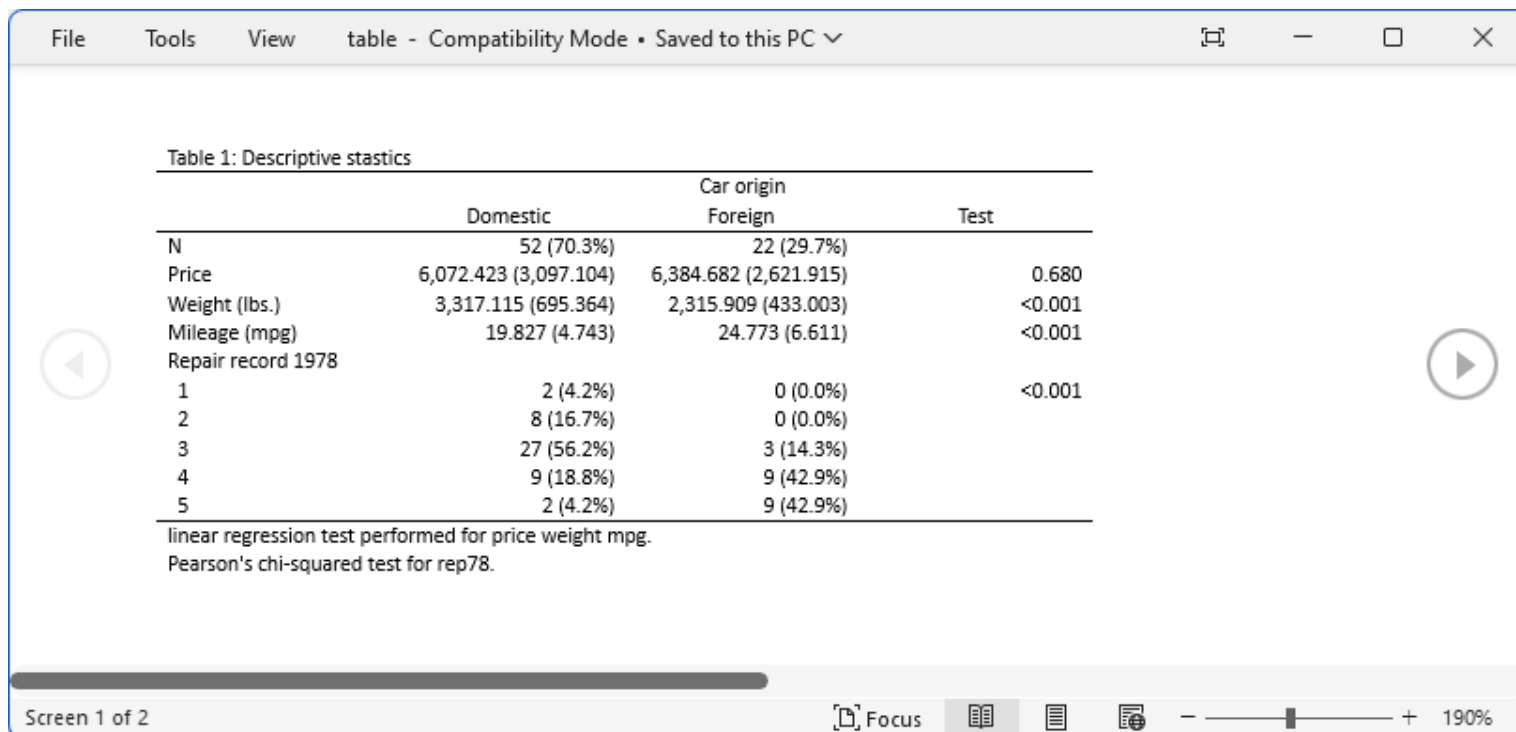


Table 1: Descriptive stastics

	Domestic	Car origin Foreign	Test
N	52 (70.3%)	22 (29.7%)	
Price	6,072.423 (3,097.104)	6,384.682 (2,621.915)	0.680
Weight (lbs.)	3,317.115 (695.364)	2,315.909 (433.003)	<0.001
Mileage (mpg)	19.827 (4.743)	24.773 (6.611)	<0.001
Repair record 1978			
1	2 (4.2%)	0 (0.0%)	<0.001
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%)	

linear regression test performed for price weight mpg.
Pearson's chi-squared test for rep78.

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

→ Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 3: The `sample()` option (1)

```
. webuse idu, clear
```

```
. dtable age ltime rtime i.needle, by(male)
```

Default row title

		Male		
		No	Yes	Total
ult row title				
N		76 (6.8%)	1,048 (93.2%)	1,124 (100.0%)
Age (in years)		28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1		22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1		11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles				
No		43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes		33 (43.4%)	369 (35.2%)	402 (35.8%)

Example 3: The `sample()` option (2)

```
. dtable age ltime rtime i.needle, by(male) sample("Sample frequency (percent)")
```

Custom row title

	No	Male Yes	Total
→ Sample frequency (percent)	76 (6.8%)	1,048 (93.2%)	1,124 (100.0%)
Age (in years)	28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1	22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1	11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles			
No	43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes	33 (43.4%)	369 (35.2%)	402 (35.8%)

Example 3: The `sample()` option (3)

```
. dtable age ltime rtime i.needle, by(male) ///  
> sample(, statistics(frequency proportion))
```

		Male			
		No	Yes	Total	
N		76 0.068	1,048 0.932	1,124	1.000
Age (in years)		28.776 (7.289)	31.656 (7.695)	31.462	(7.699)
Last time seronegative for HIV-1		22.129 (13.095)	24.323 (11.661)	24.175	(11.770)
First time seropositive for HIV-1		11.951 (10.055)	14.428 (9.170)	14.167	(9.258)
Shared needles					
No		43 (56.6%)	679 (64.8%)	722	(64.2%)
Yes		33 (43.4%)	369 (35.2%)	402	(35.8%)

Example 3: The `sample()` option (4)


```
. dtable age ltime rtime i.needle, by(male) ///  
> sample(, statistics(frequency proportion)) ///  
> sformat("N=%s" frequency) sformat("(%s)" proportion)
```

		Male		Total
		No	Yes	
N		N=76 (0.068)	N=1,048 (0.932)	N=1,124 (1.000)
Age (in years)		28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1		22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1		11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles				
No		43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes		33 (43.4%)	369 (35.2%)	402 (35.8%)

Example 3: The `sample()` option (5)

Supported sample frequency statistics:

<i>nstats</i>	Definition
<u>frequency</u>	frequency
<u>sumw</u>	sum of weights
<u>percent</u>	percentage
<u>proportion</u>	proportion
<u>rawpercent</u>	unweighted percentage
<u>rawproportion</u>	unweighted proportion



Affected by weights

- Without weights, the default sample frequency statistic is frequency.
- With weights, the default sample frequency statistic is sumw.
- With option `by()`, percent is added to the default sample frequency statistic.

Example 3: The `sample()` option (6)

```
. dtable age ltime rtime i.needle, by(male) ///  
> sample(, statistics(frequency proportion) place(item)) ///  
> sformat("N=%s" frequency) sformat("(%s)" proportion)
```

	Male		
	No	Yes	Total
N	N=76 (0.068)	N=1,048 (0.932)	N=1,124 (1.000)
Age (in years)	28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1	22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1	11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles			
No	43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes	33 (43.4%)	369 (35.2%)	402 (35.8%)

The default position

Example 3: The `sample()` option (7)

```
. dtable age ltime rtime i.needle, by(male) ///  
> sample(, statistics(frequency proportion) place(inlabels)) ///  
> sformat("N=%s" frequency) sformat("(%s)" proportion)
```

	Male In column header		
	No N=76 (0.068)	Yes N=1,048 (0.932)	Total N=1,124 (1.000)
Age (in years)	28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1	22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1	11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles			
No	43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes	33 (43.4%)	369 (35.2%)	402 (35.8%)

Example 3: The `sample()` option (8)

```
. dtable age ltime rtime i.needle, by(male) ///  
> sample(, statistics(frequency proportion) place(seplabels)) ///  
> sformat("N=%s" frequency) sformat("(%s)" proportion)
```

	No	Male Yes	Total
	N=76 (0.068)	N=1,048 (0.932)	N=1,124 (1.000)
Age (in years)	28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1	22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1	11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles			
No	43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes	33 (43.4%)	369 (35.2%)	402 (35.8%)

Example 3: The `sample()` option (9)

```
. dtable age ltime rtime i.needle, by(male) nosample
```

		Male	
		Yes	Total
		No	

Age (in years)	28.776 (7.289)	31.656 (7.695)	31.462 (7.699)
Last time seronegative for HIV-1	22.129 (13.095)	24.323 (11.661)	24.175 (11.770)
First time seropositive for HIV-1	11.951 (10.055)	14.428 (9.170)	14.167 (9.258)
Shared needles			
No	43 (56.6%)	679 (64.8%)	722 (64.2%)
Yes	33 (43.4%)	369 (35.2%)	402 (35.8%)

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

→ Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 4: The `continuous()` and `factor()` option (1)

1. For default statistics and tests, variables can be specified right after **`dtable`**. For example:

```
. dtable price mpg i.rep78 i.foreign
```

2. However, variables should be specified within **`continuous()`** or **`factor()`** if you want to customize their statistics or tests.

* Or you can use both 1 and 2!

```
. dtable x1 x2, continuous(x3 x4) factor(y1 y2)
```

* These two options are repeatable, and when multiple text styles, statistics, or tests apply to a variable, the rightmost specification is applied.

Example 4: The `continuous()` and `factor()` option (2)

- Syntax:

`continuous([varlist_c][, statistics(cstats) test(ctest) text_styles])`

and

`factor([varlist_f][, statistics(fstats) test(ftest) text_styles])`

- The supported statistics and tests can be found in **help dtable**

Example 4: The continuous() and factor() option (3a)

- Examples:

```
. webuse idu
. dtable, by(male, tests testnotes nototal) ///
sample(, statistic(frequency proportion)) ///
continuous(age, statistics( mean min max) test(kwallis)) ///
continuous(ltime rtime, statistics(mean skewness kurtosis) test(poisson)) ///
factor(needle, statistics(fvfrequency fvproportion)) ///
factor(jail inject, statistics(fvfrequency) test(fisher)) ///
sformat("(%s" fvproportion) nformat(%6.1f mean min max)
```

Example 4: The continuous() and factor() option (3b)

note: using test **kwallis** across levels of **male** for **age**.
 note: using test **poisson** across levels of **male** for **ltime** and **rtime**.
 note: using test **pearson** across levels of **male** for **needle**.
 note: using test **fisher** across levels of **male** for **jail** and **inject**.

	No				Male				Yes				Test	
N	76				0.068				1,048				0.932	
Age (in years)	28.8				18.0 46.0				31.7 17.0 52.0				0.002	
Last time seronegative for HIV-1	22.1	-0.305	2.017		24.3	-0.353	2.251		<0.001					
First time seropositive for HIV-1	12.0	0.951	2.285		14.4	0.749	3.024		0.020					
Shared needles														
No	43 (0.566)				679 (0.648)				0.149					
Yes	33 (0.434)				369 (0.352)									
Imprisoned at recruitment														
No					21				351 0.315					
Yes					55				697					
Injected drugs before recruitment														
No					47				659 0.902					
Yes					29				389					

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

→ Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 5: The `svy` and `subpop()` option (1)

- Work with complex survey design data (PSU, sampling weights, strata, etc) – **`svy`** option
 - Both summary statistics and test results are affected
- Perform subpopulation estimation on survey data
 - **`svy`** option together with **`subpop()`** option
 - Subpopulation estimation involves computing point and variance estimates for part of the population while consider the whole population size.
 - See the manual entry [\[SVY\] Subpopulation estimation](#) for more detailed information.

Example 5: The `svy` and `subpop()` option (2a)

```
. webuse nhanes21, clear
* declare the svy data
. svyset psu [pweight=finalwgt], strata(strata)
* the statistics will be computed using survey weights.
. dtable age bmi, by(heartatk, tests) svy ///
continuous(bpsystol tcresult, stat(median)) factor(sex)
```


Example 5: The `svy` and `subpop()` option (2b)

```
. dtable age bmi, by(heartatk, tests) svy continuous(bpsystol tcresult, stat(median)) factor(sex)
note: using test regress across levels of heartatk for age, bmi, bpsystol, and tcresult.
note: using test pearson across levels of heartatk for sex.
```

	No heart attack	Prior heart attack Had heart attack	Total	Test
N	113,647,835 (97.0%)	3,483,276 (3.0%)	117,131,111 (100.0%)	
Age (years)	41.695 (15.320)	60.491 (9.054)	42.254 (15.504)	<0.001
Body mass index (BMI)	25.235 (4.787)	26.604 (5.146)	25.276 (4.803)	<0.001
Systolic blood pressure	122.000	138.000	124.000	<0.001
Serum cholesterol (mg/dL)	207.000	231.000	208.000	<0.001
Sex				
Male	53,854,641 (47.4%)	2,304,839 (66.2%)	56,159,480 (47.9%)	<0.001
Female	59,793,194 (52.6%)	1,178,437 (33.8%)	60,971,631 (52.1%)	

Example 5: The `svy` and `subpop()` option (3)

* Perform subpopulation estimation on this survey data

```
. dtable age bmi, svy subpop(if heartatk==1) ///  
> continuous(bpsystol tcresult, stat(median)) factor(sex)
```

Summary

N	3,483,276
Age (years)	60.491 (11.248)
Body mass index (BMI)	26.604 (6.394)
Systolic blood pressure	138.000
Serum cholesterol (mg/dL)	231.000
Sex	
Male	2,304,839 (66.2%)
Female	1,178,437 (33.8%)

Example 5: The `svy` and `subpop()` option (3)

- Changes for tests allowed:
 - `kwallis` test is not allowed with weights or the `svy` option (continuous).
 - `fisher`, `lrchi2`, `gamma`, `kendall`, and `cramer` are not allowed with `aweight`s, `iweight`s, `pweight`s, or the `svy` option.
 - `svylr`, `svywald`, and `svyllwald` are allowed **only** with the `svy` option (factor).

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

→ Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 6: The `define()`, `nformat()`, and `sformat()` options (1)

- define composite statistics

```
. webuse idu, clear
```

(Modified Bangkok IDU Preparatory Study)

```
. dtable, continuous(age, statistics(mean minmax)) ///  
>    continuous(ltime rtime, statistics(mean variance))  ///  
>    define(minmax = min max, delimiter(-))
```

Summary

N		1,124
Age (in years)	31.462	<u>17.000-52.000</u>
Last time seronegative for HIV-1	24.175	138.534
First time seropositive for HIV-1	14.167	85.715

Example 6: The define(), nformat(), and sformat() options (2)

```
. dtable, continuous(age, statistics(mean minmax)) ///  
> continuous(ltime rtime, statistics(mean variance)) ///  
> define(minmax = min max, delimiter(-)) ///  
> nformat(%9.1f mean minmax) help format
```

Summary

N	1,124
Age (in years)	<u>31.5 17.0-52.0</u>
Last time seronegative for HIV-1	<u>24.2</u> 138.534
First time seropositive for HIV-1	<u>14.2</u> 85.715

Example 6: The `define()`, `nformat()`, and `sformat()` options (3)

```
. dtable, continuous(age, statistics(mean minmax)) ///  
> continuous(ltime rtime, statistics(mean variance)) ///  
> define(minmax = min max, delimiter(-)) ///  
> nformat(%9.1f mean minmax) ///  
> sformat("[%s]" minmax) sformat("(%s)" variance)
```

Summary

N	1,124
Age (in years)	31.5 [17.0-52.0]
Last time seronegative for HIV-1	24.2 (138.534)
First time seropositive for HIV-1	14.2 (85.715)

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

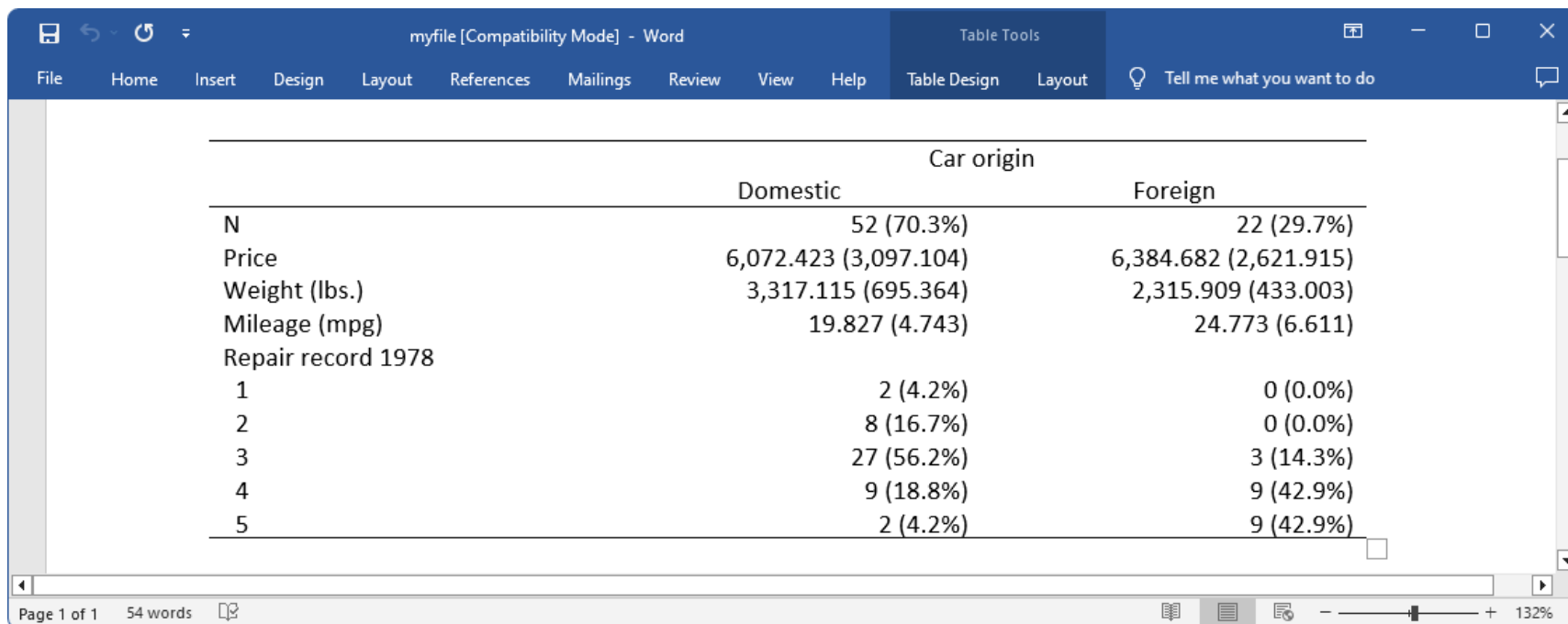
→ Example 7: The **export()** option

Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 7: The export() option (1)

```
. sysuse auto, clear  
. dtable price weight mpg i.rep78, by(foreign, nototal) ///  
> export(myfile.docx, replace)
```



	Car origin	
	Domestic	Foreign
N	52 (70.3%)	22 (29.7%)
Price	6,072.423 (3,097.104)	6,384.682 (2,621.915)
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%)

Example 7: The export() option (2)

Supported file format

docx	as(docx)	Microsoft Word
html	as(html)	HTML 5 with CSS
pdf	as(pdf)	PDF
xlsx	as(xlsx)	Microsoft Excel 2007/2010 or newer
xls	as(xls)	Microsoft Excel 1997/2003
tex	as(latex)	LaTeX
smcl	as(smcl)	SMCL
txt	as(txt)	plain text
markdown	as(markdown)	Markdown
md	as(markdown)	Markdown

Example 7: The export() option (3)

- The **alternative ways** to export:
 1. The **collect export** command (only needed if you make some further changes using **collect** after calling **dtable**)

```
dtable price weight mpg i.rep78, by(foreign, nototal)
* collect commands here
collect export myfile2.docx, replace
```

See example 8

Example 7: The export() option (4)

- The **alternative ways** ways to export:
2. insert the table obtained with **dtable** (or with **collect**) into a larger document

```
putdocx collect
```

```
putpdf collect
```

```
putexcel ul_cell = collect
```

- One example (in the end):

<https://blog.stata.com/2023/06/26/creating-tables-of-descriptive-statistics-in-stata-18-the-new-dtable-command/>

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

→ Example 8: Work together with **collect**

Example 9: Using the dialog box

Example 8: Work together with `collect` or `table` (1)

- Work with `collect`

```
. dtable age ltime rtime i.needle, by(male, nototal) ///  
> sample(Sample size) nformat(%6.2f mean sd) column(by(hide))
```

	No	Yes
Sample size	76 (6.8%)	1,048 (93.2%)
Age (in years)	28.78 (7.29)	31.66 (7.69)
Last time seronegative for HIV-1	22.13 (13.10)	24.32 (11.66)
First time seropositive for HIV-1	11.95 (10.06)	14.43 (9.17)
Shared needles		
No	43 (56.6%)	679 (64.8%)
Yes	33 (43.4%)	369 (35.2%)

Example 8: Work together with `collect` or `table` (2)

```
collect label levels male 0 "Female", modify
collect label levels male 1 "Male", modify
collect style cell male[0], warn shading( background(aqua))
collect style cell male[1], warn shading( background(lightyellow))
collect style row split, binder(`":'')
collect style header needle, title(label)
collect style cell var[0.needle], warn border( top, width(1))
collect style putdocx, layout(autofitcontents)
collect export ex8.docx, replace
```

	Female	Male
Sample size	76 (6.8%)	1,048 (93.2%)
Age (in years)	28.78 (7.29)	31.66 (7.69)
Last time seronegative for HIV-1	22.13 (13.10)	24.32 (11.66)
First time seropositive for HIV-1	11.95 (10.06)	14.43 (9.17)
Shared needles:No	43 (56.6%)	679 (64.8%)
Shared needles:Yes	33 (43.4%)	369 (35.2%)

Example 8: Work together with collect or table (3)

```
. collect style save mydtable, replace
. collect label save mylabel, replace

. clear all
. webuse idu
. dtable age ltime rtime i.needle, by(male, nototal) style(mydtable) label(mylabel) ///
> export(ex8b.docx, replace)
```

	Female	Male
Sample size	76 (6.8%)	1,048 (93.2%)
Age (in years)	28.78 (7.29)	31.66 (7.69)
Last time seronegative for HIV-1	22.13 (13.10)	24.32 (11.66)
First time seropositive for HIV-1	11.95 (10.06)	14.43 (9.17)
Shared needles:No	43 (56.6%)	679 (64.8%)
Shared needles:Yes	33 (43.4%)	369 (35.2%)

Example 8: Work together with `collect` or `table` (2)

- Work with `table` and `collect` (more advanced)

https://www.stata.com/support/faqs/reporting/combine-multiple-tables/#ex_3b

Examples

Example 1: The first try

Example 2: The **by()**, **total()**, and **note()** options

Example 3: The **sample()** option

Example 4: The **continuous()** and **factor()** option

Example 5: The **svy** and **subpop()** option

Example 6: The **define()**, **nformat()**, and **sformat()** options

Example 7: The **export()** option

Example 8: Work together with **collect**

→ Example 9: Using the dialog box

Example 9: Using the dialog box

- Click menu **Statistics > Summaries, tables, and tests > Table of descriptive statistics** to open the dialog box for **dtable**.

. db dtable

dtable - Table of descriptive statistics

Main if/in Weights Survey Sample Continuous Factors Composite Formats Title Notes Export Options

Variables: (optional)

Summary header title: (leave empty for default)

Summary header styles

☐ Describe variables by group

Variable that defines groups:

Group variable title:

Group variable levels and labels:

Allow numeric missing groups:

Tests across groups:

Show note for each test across groups:

Total sample statistics:

Test header styles Total header styles

OK Cancel Submit

More learning resources

- [dtable's PDF manual entry](#)
- [dtable's feature webpage](#)
- [Stata Blog article about dtable](#)
- [dtable's video tutorial](#)
- [Web training: Customize reproducible tables using Stata](#)
(7-10, November 2023)

Get help



- Technical support: tech-support@stata.com
- Statalist forum: www.statalist.org

Thank you!

