

## Reproducible and automated reporting

- **Reproducible reports in Word, Excel, HTML, and PDF**
  - Markdown with dynamic document features creates HTML and Word documents
  - Create customized Word, Excel, and PDF documents
- **Automated reports**
  - Update reports as data change
  - Rerun one command or do-file to automatically update your entire document
- **Create reports that intermix**
  - Formatted text and tables
  - Summary statistics
  - Regression results
  - Stata graphics and other images in SVG, PNG, EPS, TIFF, and EMF formats
- **Customize tables of Stata results**
  - Create tables of results from any Stata command
  - Customize layout, fonts, headers, formats, titles, footnotes, and more
  - Reuse your style in future tables
  - Export to Word, Excel, LaTeX, PDF, HTML, and Markdown
  - Incorporate tables in complete reports

### Truly reproducible

Stata makes reproducible research easy. With Stata's integrated version control, you can use the **version 18** command, and any code you run today will produce the same results in whatever release of Stata you are running 10, 20, or more years from now. Stata handles backward compatibility for you. With the **datasignature** command, you can verify that your data have not changed. And when these tools for reproducibility are paired with the scripts that create your reports, you can easily create reproducible reports. Rerun your commands at any time and re-create your report.

### Customizable tables

With the **table** command, you can easily create cross-tabulations, tables of summary statistics, and tables of results from other Stata commands while customizing numeric formats, table layout, and more. Additionally, you can create a table with descriptive statistics, commonly known as a "Table 1", and export the results in a single step with the new **dtable** command. And when you want to report results from one or more regression commands—whether linear regressions, panel-data models, survival models, or other regression models—the **etable** command makes this easy. With the **collect** suite of commands, you can collect results from Stata commands and then customize your layout, labels, headers, fonts, titles, and other features of the table. When you have a style you like, you can save it and apply it to tables you create in the future.

### Dynamic and automated documents

Whether you prefer reporting in HTML, Word, Excel, or PDF, you can easily make your reports dynamic. Say that you need to run the same report each month, updating the results using new data. Simply rerun the commands that created the report with the updated dataset. All Stata results in the report, including graphs and tables, are updated automatically.

The screenshot shows the Stata Tables Builder interface. On the left, there are various options for customizing the table, such as 'Table title styles', 'Table row styles', and 'Table column headers'. The main window displays a table with the following data:

|                     | Nonunion     | Union       | Var     |
|---------------------|--------------|-------------|---------|
| College graduate    | 1,181 (78%)  | 313 (68%)   | p<0.001 |
| Mar                 | 316 (22%)    | 148 (32%)   |         |
| NonMARRA            | 441 (31%)    | 186 (23%)   | p<0.001 |
| MSMA                | 591 (42%)    | 382 (84%)   |         |
| Central City        | 383 (27%)    | 173 (38%)   |         |
| 300 nearest Census  | 44.4 (4.4)   | 7.89 (6.1)  | p<0.001 |
| Local Census method | 37.26 (10.2) | 38.66 (8.1) | p<0.001 |
| Hourly wage         | 9.70 (4.10)  | 9.67 (4.1)  | p<0.001 |

Below the table, there is a preview window showing a regression table with coefficients and confidence intervals:

|                        | 1                       | 2                       | 3                       | 4                       |
|------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
|                        | 0.05***<br>[0.02, 0.07] | 0.07***<br>[0.05, 0.09] | 0.09***<br>[0.08, 0.11] | 0.09***<br>[0.08, 0.10] |
|                        | -0.10*<br>[-0.20, 0.00] | 0.07**<br>[0.01, 0.13]  | 0.01<br>[-0.07, 0.10]   | 0.05<br>[-0.02, 0.12]   |
| nvalue x kstock        | 0.00***<br>[0.00, 0.00] | 0.00***<br>[0.00, 0.00] | 0.00***<br>[0.00, 0.00] | 0.00***<br>[0.00, 0.00] |
| Year fixed effects     | YES                     | NO                      | YES                     | NO                      |
| Company fixed effects  | YES                     | YES                     | NO                      | NO                      |
| R-squared              | 0.67                    | 0.84                    | 0.85                    | 0.85                    |
| Number of observations | 200                     | 200                     | 200                     | 200                     |

At the bottom right, there is a legend: \*p<0.05, \*\*p<0.01, \*\*\*p<0.001.

## Markdown to HTML or Word

Convert dynamic Markdown documents to HTML and Word using **dyndoc**. To start, create a text file that intermixes the Markdown text-formatting language with the Stata commands that create the output you want. Include dynamic tags (terms in `<< ... >>`) to give Stata instructions about executing commands and incorporating results. Then create your HTML or Word document by typing

`. dyndoc report.md`

or

`. dyndoc report.md, docx`

Want to change the style? Use different CSS, and automatically change the look and feel of your document.

Table 1: Log wage regression results

|                        | Model1              | Model2              |
|------------------------|---------------------|---------------------|
| Hours worked           | 0.003 **<br>(0.000) | 0.003 **<br>(0.000) |
| College graduate       | 0.383 **<br>(0.008) | 0.383 **<br>(0.008) |
| Age                    | 0.010 **<br>(0.000) | 0.010 **<br>(0.001) |
| Union member           | 0.190 **<br>(0.007) | 0.133 **<br>(0.038) |
| Union member x Age     |                     | 0.002<br>(0.001)    |
| Intercept              | 1.209 **<br>(0.020) | 1.222 **<br>(0.021) |
| Number of observations | 19193               | 19193               |

Regression results

We fit two regression models. First, we model log wage on hours worked, an indicator of college graduation, age, and an indicator of union membership. In the second model, we add an interaction between age and union membership.

Table 1: Log wage regression results

|                        | Model1              | Model2              |
|------------------------|---------------------|---------------------|
| Usual hours worked     | 0.003 **<br>(0.000) | 0.003 **<br>(0.000) |
| College graduate       | 0.383 **<br>(0.008) | 0.383 **<br>(0.008) |
| Age                    | 0.010 **<br>(0.000) | 0.010 **<br>(0.001) |
| Union member           | 0.190 **<br>(0.007) | 0.133 **<br>(0.038) |
| Union member x Age     |                     | 0.002<br>(0.001)    |
| Intercept              | 1.209 **<br>(0.020) | 1.222 **<br>(0.021) |
| Number of observations | 19193               | 19193               |

```
1 version 18
2
3 use nlswork, clear
4
5 // Create a document
6 putdocx begin
7
8 // Add headers and paragraph text
9 putdocx paragraph, style(Heading1)
10 putdocx text ("Data")
11
12 putdocx textblock begin
13 We use data from the National Longitudinal Survey,
14 focusing on young women 14-16 years of age in 1968.
15 putdocx textblock end
16
17 putdocx paragraph, style(Heading1)
18 putdocx text ("Regression results")
19
20 putdocx textblock begin
21 We fit two regression models. First, we model log wage on hours
22 worked, an indicator of college graduation, age, and an indicator
23 of union membership. In the second model, we add an interaction
24 between age and union membership.
25 putdocx textblock end
26
27 // Fit regression models and create table of results
28 regress ln_wage hours i.collgrad age i.union
29 estimates store Model1
30 regress ln_wage hours i.collgrad age i.union i.union#c.age
31 estimates store Model2
32
33 etable, estimates(Model1 Model2) col(estimates) ///
34 showstars showstarsnote title("Table 1: Log wage regression results")
35
36 collect style putdocx, layout(autofitcontents)
37 putdocx collect
38
39 putdocx save report2.docx, replace
40
```

report.md

```
1 <<dd_version: 2>>
2 <link rel="stylesheet" type="text/css" href="stmarkdown.css">
3
4 ##Study of High Blood pressure
5
6 We use data from the National Health and Nutrition Examination
7 Survey to study the relationship between high blood pressure
8 and other variables
9
10 ~~~~
11 <<dd_do:quietly>>
12 webuse nhanes2, clear
13 <</dd_do>>
14
15 ~~~~
16 ##Summary statistics
17
18 ~~~~
19 <<dd_do:quietly>>
20 table highbp, stat(fvfrequency sex)
21 stat(fvpercent sex)
22 stat(mean weight age)
23 stat(sd weight age)
24 collect recode fvfrequency = col1
25 fvpercent = column2
26 mean = column1
27 sd = column2
28 collect layout (var) (highbp#result[co
29 collect label levels highbp 0 "No" 1 "Y
30 collect style header result, level(hide
31 collect style row stack, nobinder space
32 collect style cell border_block, border
33 collect style cell var[sex]#result[col
34 collect style cell var[sex]#result[co
35 collect style cell var[weight age]#resu
36 collect style cell var[weight age]#resu
37 collect export sumtab.html, replace
38 <</dd_do>>
39 ~~~~
```

report.html

### Study of High Blood pressure

We use data from the National Health and Nutrition Examination Survey to study the relationship between high blood pressure and other variables

#### Summary statistics

|             | No            | Yes           | Total         |
|-------------|---------------|---------------|---------------|
| Sex         |               |               |               |
| Male        | 2,611 (43.7%) | 2,304 (22.7%) | 4,915 (47.5%) |
| Female      | 3,264 (56.2%) | 2,079 (47.3%) | 5,343 (52.5%) |
| Weight (kg) | 68.3 (13.0)   | 75.9 (16.2)   | 71.9 (15.4)   |
| Age (years) | 43.2 (16.8)   | 52.0 (14.9)   | 47.6 (17.5)   |

#### Logistic regression results

We fit a logistic regression model of high blood pressure on age, sex, the interaction between age and sex, and weight.

|                      | Odds ratio | Std. error | p-value | 95% CI      |
|----------------------|------------|------------|---------|-------------|
| Female               | 0.27       | 0.04       | 0.00    | 0.20 - 0.36 |
| Age (years)          | 1.04       | 0.00       | 0.00    | 1.04 - 1.04 |
| Age (years)          | 1.00       | 0.00       |         |             |
| Female x Age (years) | 1.02       | 0.00       | 0.00    | 1.02 - 1.03 |
| Weight (kg)          | 1.04       | 0.00       | 0.00    | 1.04 - 1.05 |
| Intercept            | 0.00       | 0.00       | 0.00    | 0.00 - 0.01 |

#### Interaction plot

Expected probability of high blood pressure

## Customized Word, PDF, and Excel files

With **putdocx**, **putpdf**, and **putexcel**, you can create customized Word, PDF, and Excel files that incorporate results from Stata commands, graphs, formatted text, and tables. For example,

Begin a document

`. putdocx begin`

Add text

`. putdocx textblock begin`

`...`

`. putdocx textblock end`

Add a graph

`. putdocx image`

Add a customized table

`. putdocx collect`

Save a document

`. putdocx save`

With a short script composed of similar commands, you can easily create a complete report.