


13 Using the Do-file Editor—automating Stata

The Do-file Editor

Stata comes with an integrated text editor called the Do-file Editor, which can be used for many tasks. It gets its name from the term *do-file*, which is a file containing a list of commands for Stata to run (called a batch file or a script in other settings). See [U] 16 **Do-files** for more information. Although the Do-file Editor has advanced features that can help in writing such files, it can also be used to build up a series of commands that can then be submitted to Stata all at once. This feature can be handy when writing a loop to process multiple variables in a similar fashion or when doing complex, repetitive tasks interactively.

To get the most from this chapter, you should work through it at your computer. Start by opening the do-file editor, either by clicking on the **Do-file Editor** button, , or by typing `doedit` in the Command window and pressing *Return*.

The Do-file Editor toolbar

The Do-file Editor has seven buttons. Many of the buttons share a similar purpose with their look-alikes in the main Stata toolbar.



If you ever forget what a button does, hover the mouse pointer over a button, and a tooltip will appear.



Open: Open a do-file from disk in a new tab in the Do-file Editor.



Save: Save the current do-file to disk.



Print: Print the contents of the Do-file Editor.



Find: Open the *Find* dialog for finding text.



Show: Toggle display of invisible characters.



Zoom: Change the display size of the text.



Execute (do): Run the commands in the do-file, showing all commands and their output. If text is highlighted, the button becomes the **Execute selection (do)** button and will run only the selected lines, showing all output. We will refer to this as the **Do** button.

Using the Do-file Editor

Suppose that we would like to analyze fuel usage for 1978 automobiles in a manner similar to what we did in [GSM] 1 [Introducing Stata—sample session](#). We know that we will be issuing many commands to Stata during our analysis and that we want to be able to reproduce our work later without having to type each command again.

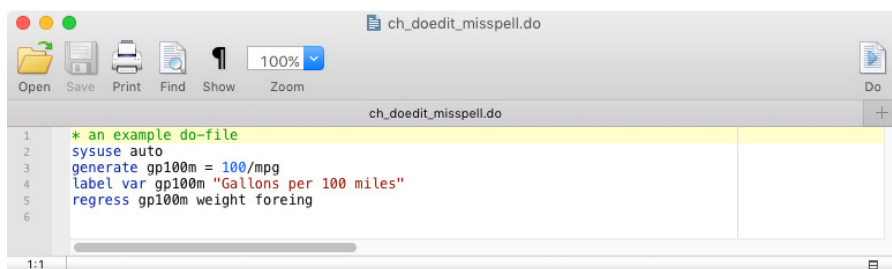
We can do this easily in Stata: simply save a text file containing the commands. When that is done, we can tell Stata to run the file and execute each command in sequence. Such a file is known as a Stata *do-file*; see [U] [16 Do-files](#).

To analyze fuel usage of 1978 automobiles, we would like to create a new variable containing gallons per mile. We would like to see how that variable changes in relation to vehicle weight for both domestic and imported cars. Performing a regression with our new variable would be a good first step.


To get started, click on the **Do-file Editor** button to open the Do-file Editor. After the Do-file Editor opens, type the commands below into the Do-file Editor. Purposely misspell the name of the `foreign` variable on the fifth line. (We are intentionally making some common mistakes and then pointing you to the solutions. This will save you time later.)

```
* an example do-file
sysuse auto
generate gp100m = 100/mpg
label var gp100m "Gallons per 100 miles"
regress gp100m weight foreing
```

Here is what your Do-file Editor should look like now:



You will notice that the color of the text changes as you type. The different colors are examples of the Do-file Editor's syntax highlighting. The colors and text properties of the syntax elements can be changed by right-clicking in the Do-file Editor, selecting **Preferences...** and then clicking on the **Syntax highlighting** tab in the resulting window.

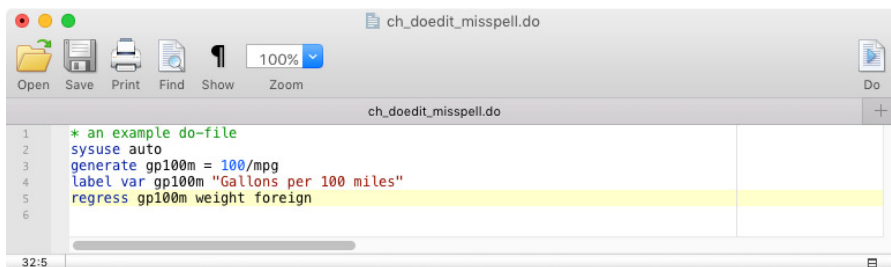
Click on the **Do** button, , to execute the commands. Stata executes the commands in sequence, and the results appear in the Results window:

```

. do /tmp/SD00001.000000
. * an example do-file
. sysuse auto
(1978 Automobile Data)
. generate gp100m = 100/mpg
. label var gp100m "Gallons per 100 miles"
. regress gp100m weight foreing
variable foreing not found
r(111);
.
end of do-file

```

The `do "/tmp/..."` command is how Stata executes the commands in the Do-file Editor. Stata saves the commands to a temporary file and issues the `do` command to execute them. Everything worked as planned until Stata saw the misspelled variable. The first three commands were executed, but an error was produced on the fourth. Stata does not know of a variable named `foreing`. We need to go back to the Do-file Editor and change the misspelled variable name to `foreign` in the last line:



We click on the **Do** button again. Alas, Stata now fails on the first line—it will not overwrite the dataset in memory that we changed.

```

. do /tmp/SD00001.000000
. * an example do-file
. sysuse auto
no; data in memory would be lost
r(4);
.
end of do-file

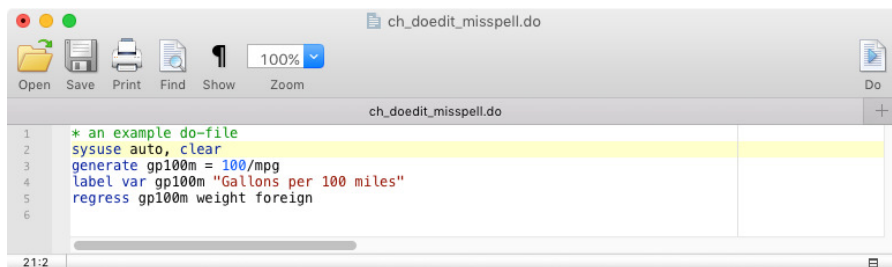
```

We now have a choice for what we should do:

- We can put a `clear` command in our do-file as the very first command. This automatically clears out Stata's memory before the do-file tries to load the `auto` dataset. This is convenient but dangerous because it defeats Stata's protection against throwing away changes without warning.
- We can type a `clear` command in the Command window to manually clear the dataset and then process the do-file again. This process can be aggravating when building a complicated do-file.

Here is some advice: Automatically clear Stata's memory while debugging the do-file. Once the do-file is in its final form, decide the context in which it will be used. If it will be used in a highly automated environment (such as when certifying), the do-file should still automatically clear Stata's

memory. If it will be used rarely, do not clear Stata's memory. This decision will save much heartache. We will add a `clear` option to the `sysuse` command to automatically clear the dataset in Stata's memory before the do-file runs:



The do-file now runs well, as clicking on the **Do** button shows:

```
. do /tmp/SD00001.000000
. * an example do-file
. sysuse auto, clear
(1978 Automobile Data)
. generate gp100m = 100/mpg
. label var gp100m "Gallons per 100 miles"
. regress gp100m weight foreign
```

Source	SS	df	MS			
Model	91.1761694	2	45.5880847	Number of obs = 74		
Residual	28.4000913	71	.400001287	F(2, 71) = 113.97		
Total	119.576261	73	1.63803097	Prob > F = 0.0000		
				R-squared = 0.7625		
				Adj R-squared = 0.7558		
				Root MSE = .63246		

gp100m	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
weight	.0016254	.0001183	13.74	0.000	.0013896	.0018612
foreign	.6220535	.1997381	3.11	0.003	.2237871	1.02032
_cons	-.0734839	.4019932	-0.18	0.855	-.8750354	.7280677

```
.
end of do-file
```

You might want to select **File > Save as...** to save this do-file while the Do-file Editor is in front. Later, you could select **File > Open...** to open it and then add more commands as you move forward with your analysis. By saving the commands of your analysis in a do-file as you go, you do not have to worry about retyping them with each new Stata session. Think hard about removing the `clear` option from the first command.

After you have saved your do-file, you can execute the commands it contains by typing `do filename`, where the *filename* is the name of your do-file.

The File menu

When the Do-file Editor is in front, the **File** menu's items apply to do-files. You may choose any of these menu items: create a **New > Do-file, Open...** an existing file, **Save** the current file, save the current file under a new name with **Save as...**, or **Print** the current file. There are also buttons on the Do-file Editor's toolbar that correspond to these features.

Finally, you can create a **New > Project...** to keep track of collections of files used in a project. These can be do-files, data files, graph files, or any other files you like. For more information on the Project Manager, see [P] [Project Manager](#).

The Edit menu

The **Edit** menu of the Do-file Editor includes the standard **Undo**, **Redo**, **Cut**, **Copy**, **Paste**, and **Find** capabilities. There are several other **Edit** menu features that you might find useful:


- You can select **Insert file...** to insert the contents of another file at the current cursor position in the Do-file Editor.
- You can select the current line with **Select line**.
- You can delete the current line with **Delete line**.
- **Find > Go to line...** will allow you to jump to a specific line number. The line numbers are displayed at the left of the Do-file Editor window.
- **Advanced** leads to a submenu with some programmer's friends:
 - **Shift right** indents the selection by one tab.
 - **Shift left** unindents the selection by one tab.
 - **Re-indent** indents the selection according to its nesting within blocks and programs.
 - **Toggle comment** toggles `//`-style comments at the start of the selected lines.
 - **Add block comment** puts a `/*` before and a `*/` after the selected region, commenting it out.
 - **Remove block comment** undoes the above.
 - **Make selection uppercase** converts the selection to all capital letters.
 - **Make selection lowercase** converts the selection to all lowercase letters.
 - **Convert to UTF-8...** converts the current file to UTF-8 encoding.
 - **Convert line endings to macOS/Unix format (\n)** converts the line endings for the current file to macOS/Unix format.
 - **Convert line endings to Windows format (\r\n)** converts the line endings for the current file to Windows format.

Matching and balancing of parentheses (), braces { }, and brackets [] are also available from the **Edit** menu. When you select **Edit > Find > Match brace**, the Do-file Editor looks at the character immediately to the left and right of the cursor. If either is one of the characters that the editor can match, the editor will find the matching character and place the cursor immediately in front of it. If there is no match, the menu item will be inactive.

When you select **Edit > Find > Balance braces**, the Do-file Editor looks to the left and right of the current cursor position or selection and creates a selection that includes the narrowest level of matching characters. If you select **Balance braces** again, the editor will expand the selection to include the next level of matching characters. If there is no match, the cursor will not move. Balancing braces is useful for working with complicated expressions or blocks of code defined by loops or `if` commands. See [P] [foreach](#), [P] [forvalues](#), [P] [while](#), and [P] [if](#) for more information.

Balance braces is easier to explain with an example. Type `(now (is the) time)` in the Do-file Editor. Place the cursor between the words `is` and `the`. Select **Edit > Find > Balance braces**. The Do-file Editor will select `(is the)`. If you select **Balance braces** again, the Do-file Editor will select `(now (is the) time)`.

Text in Stata strings can include Unicode characters and is encoded as UTF-8 (see [U] 12.4.2 [Handling Unicode strings](#)). However, you may have do-files, ado-files, or other text files that you used with Stata 13 or earlier, and those files contain characters other than plain ASCII such as accented characters, Chinese, Japanese, or Korean (CJK) characters, Cyrillic characters, and the like. If you open a file that is not encoded in UTF-8, Stata prompts you to specify the encoding for the file so that it can convert the file to UTF-8. If you cancel the conversion or choose the wrong encoding, you can try the conversion again later using **Convert to UTF-8**. The conversion to UTF-8 can be undone by using **Edit > Undo** and is not permanent until you save the do-file. For Stata datasets with characters not encoded in UTF-8 or for bulk conversion of multiple Stata files, you should use the `unicode translate` command.

Editing tip: You can split the Do-file Editor window vertically by clicking on the **Split window** button in the status bar . This is useful for looking at two widely separated places in a file at the same time. To go back to a single pane view, click on the close button in either of the split panes.

The View > Do-file Editor menu

You have already learned about the **Do** button. Selecting **View > Do-file Editor > Execute (do)** is equivalent to clicking on the **Execute (do)** button.

Selecting **View > Do-file Editor > Execute (do) to bottom** will send all the commands from the current line through the end of the contents of the Do-file Editor to the Command window. This method is a quick way to run a part of a do-file.

Selecting **View > Do-file Editor > Execute (include)** is similar to clicking on the **Execute (do)** button with one major difference: local macros defined in the current session can be expanded in the commands being executed.

Do is equivalent to Stata's `do` command, whereas **Execute (include)** is equivalent to Stata's `include` command. See [U] 16 [Do-files](#) for a complete discussion.

You can also preview files in the Viewer by selecting **View > Do-file Editor > Show file in Viewer**. This feature is useful when working with files that use Stata's SMCL tags, such as when writing help files or editing log files.

Saving interactive commands from Stata as a do-file

While working interactively with Stata, you might decide that you would like to rerun the last several commands that you typed interactively. From the Review window, you can send highlighted commands or even the entire contents to the Do-file Editor. You can also save commands as a do-file and open that file in the Do-file Editor. You can copy a command from a dialog (rather than submit it) and paste it into the Do-file Editor. See [GSM] 6 [Using the Data Editor](#) for details. Also see [R] [log](#) for information on the `cmdlog` command, which allows you to log all commands that you type in Stata to a do-file.

Projects

For advanced users managing many files as part of a project, Stata has a Project Manager that uses the Do-file Editor. For more information on the Project Manager, see [P] [Project Manager](#).