

5 Opening and saving Stata datasets

How to load your dataset from disk and save it to disk

Opening and saving datasets in Stata works similarly to those tasks in other computer applications. There are a few differences, however. First, it is possible to save and open files from within Stata's Command window. Second, Stata allows just one dataset to be open and in use at any one time. It is possible to have many Viewers viewing many files, but only one dataset may be in use at any time. Keeping this in mind will make Stata's care in opening new datasets clear. This chapter outlines all the possible ways to open and save datasets.

A Stata dataset can be opened in a variety of ways, most of which are probably familiar to you from other applications:

- Double-click on a Stata data file, which is a file whose extension is `.dta`. Note: The file extension may not be visible, depending on what options you have set in your operating system.
- Select **File > Open...** or click on the **Open** button and navigate to the file.
- Select **File > Recent Datasets > filename**.
- Type `use filename` in the Command window. Stata will look for `filename` in the current working directory. If the file is located elsewhere, you will need to give its path. Be aware that if there is a space anywhere in the path or filename, you will need to put the filename inside quotation marks. See [U] 11.6 **File naming conventions**.
- Type `sysuse filename` in the Command window. Stata will look for `filename` in a series of directories called the `adopath`. Typically, this is for finding example datasets installed when you installed Stata, but it can also be used for easy access to your own datasets. For more information on the `adopath`, see [P] **sysdir**.
- Type `webuse filename` in the Command window. The `webuse` command is used to access datasets used in the Stata manuals; for example, `webuse lbw` loads the `lbw` dataset used in the documentation of the `logistic` command. For more information, see [D] **webuse**.

Because Stata has at most one dataset open at a time, opening a dataset will cause Stata to discard the dataset that is currently in memory. If there have been changes to the data in the currently open dataset, Stata will refuse to discard the dataset unless you force it to do so. If you open the file with any method other than the Command window, you will be prompted. If you use the Command window and the current data have changed, you will get the following error message:

```
. sysuse auto
no; data in memory would be lost
r(4);
```

These behaviors protect you from mistakenly losing data.

To save an unnamed dataset (or an old dataset under a new name):

1. select **File > Save As...**; or
2. type `save filename` in the Command window.

To save a dataset for use with Stata 11 or Stata 12 (Stata 11 can load Stata 12 datasets),

1. select **File > Save As...**, and select **Stata 12 Data (*.dta)** from the **Save as type** list; or
2. type `saveold filename` in the Command window.

To save a dataset that has been changed (overwriting the original data file),

1. select **File > Save**;
2. click on the **Save** button; or
3. type `save, replace` in the Command window.

Once you overwrite a dataset, there is no way to recover your original dataset. With important datasets, you may want to either keep a backup copy of your original `filename.dta` or save your changes to a dataset under a new name. This is no different from working with a word-processing document, except that recovering from an inadvertent save of a dataset is nearly impossible.

Important note: Changes you have made to a dataset are not permanent until you save them. You work with a copy of the dataset in memory, not with the data file itself. This should not be surprising, because it is the way that you work with almost all applications on your computer.

If you do not want to save your dataset, you can clear the dataset in memory and open a new dataset by typing `use filename, clear`.