Title

intro — Introduction to data-management reference manual

Description

This entry describes this manual and what has changed since Stata 11. See the next entry, [D] **data management**, for an introduction to Stata's data-management capabilities.

Remarks

This manual documents most of Stata's data-management features and is referred to as the [D] manual. Some specialized data-management features are documented in such subject-specific reference manuals as [MI] *Stata Multiple-Imputation Reference Manual*, [TS] *Stata Time-Series Reference Manual*, [ST] *Stata Survival Analysis and Epidemiological Tables Reference Manual*, and [XT] *Stata Longitudinal-Data/Panel-Data Reference Manual*.

Following this entry, [D] **data management** provides an overview of data management in Stata and of Stata's data-management commands. The other parts of this manual are arranged alphabetically. If you are new to Stata's data-management features, we recommend that you read the following first:

[D] **data management** — Introduction to data-management commands
[U] **12 Data**
[U] **13 Functions and expressions**
[U] **11.5 by varlist: construct**
[U] **21 Inputting and importing data**
[U] **22 Combining datasets**
[U] **23 Working with strings**
[U] **25 Working with categorical data and factor variables**
[U] **24 Working with dates and times**
[U] **16 Do-files**

You can see that most of the suggested reading is in [U]. That is because [U] provides overviews of most Stata features, whereas this is a reference manual and provides details on the usage of specific commands. You will get an overview of features for combining data from [U] **22 Combining datasets**, but the details of performing a match-merge (merging the records of two files by matching the records on a common variable) will be found here, in [D] **merge**.

Stata is continually being updated, and Stata users are always writing new commands. To ensure that you have the latest features, you should install the most recent official update; see [R] **update**.

What's new

This section is intended for previous Stata users. If you are new to Stata, you may as well skip it.

1. **Automatic memory management**, which means that you no longer have to `set memory` and never again will you be told that there is no room because you set too little! Stata automatically adjusts its memory usage up and down according to current requirements.

The memory manager is tunable. We recommend the default settings. See [D] **memory** if you are interested.

Old do-files can still `set memory`. Stata merely responds, "`set memory` ignored".

2. **Excel files, importing and exporting.** And the new import preview tool lets you see the data before you import them. See [D] **import excel**.
3. **EBCDIC files, importing.** And you can convert between EBCDIC and ASCII formats; see [D] **infile (fixed format)** and [D] **filefilter**.
4. **ODBC connection strings, importing and exporting** and **ODBC support for Oracle Solaris.** See [D] **odbc**.
5. **PDF files, exporting of graphs and logs**. You can directly create PDFs from your Stata results. See [G-2] **graph export** and [R] **translate**.
6. **Business dates** allow you to define your own calendars so that they display correctly and lags and leads work as they should. You could create file `lse.stbcal` that recorded the days the London Stock Exchange is open (or closed) and then Stata would understand format `%tblse` just as it understands the usual date format `%td`. Once you define a calendar, Stata deeply understands it. You can, for instance, easily convert between `%tblse` and `%td` values. See [D] **datetime business calendars**.
7. **Improved documentation for date and time variables**. Anyone who has ever been puzzled by Stata's date and time variables, which is to say, anyone who uses them, should see [D] **datetime**, [D] **datetime translation**, and [D] **datetime display formats**.
8. **Renaming groups of variables** is now easy using `rename`'s new syntax that is 100% compatible with its old syntax. You can change names, swap names, renumber indices within variable names, and more. See [D] **rename group**.
9. **New functions**,
   a. **Tukey's Studentized range**, cumulative and inverse, `tukeyprob()` and `invtukeyprob()`.
   b. **Dunnett's multiple range**, cumulative and inverse, `dunnettprob()` and `invdunnettprob()`.
   c. **New date conversion functions** `dofb()` and `bofd()` convert between business dates and standard calendar dates. See [D] **datetime business calendars**.

   See [D] **functions**.
10. **New Stata commands getmata and putmata** make it easy to transfer your data into Mata, manipulate them, and then transfer them back to Stata. `getmata` and `putmata` are especially designed for interactive use. See [D] **putmata**.
11. **New Stata commands import sasxport, export sasxport, and import sasxport, describe** replace existing commands `fdause`, `fdasave`, and `fdadescribe`. `fdause`, `fdasave`, and `fdadescribe` are understood as synonyms. See [D] **import sasxport**.
12. **xshell** support for Mac. See [D] **shell**.

For a complete list of all the new features in Stata 12, see [U] **1.3 What's new**.

Also see

[U] **1.3 What's new**

[R] **intro** — Introduction to base reference manual