drop — Drop variables or observations

Syntax

Drop variables

\texttt{drop \ textit{varlist}}

Drop observations

\texttt{drop if \textit{exp}}

Drop a range of observations

\texttt{drop in range [if \textit{exp}]}

Keep variables

\texttt{keep \textit{varlist}}

Keep observations that satisfy specified condition

\texttt{keep if \textit{exp}}

Keep a range of observations

\texttt{keep in range [if \textit{exp}]}  

by is allowed with the second syntax of \texttt{drop} and the second syntax of \texttt{keep}; see \texttt{[D] by}.

Menu

\textbf{Drop or keep variables}

Data > Variables Manager

\textbf{Drop or keep observations}

Data > Create or change data > Drop or keep observations

Description

drop eliminates variables or observations from the data in memory.

keep works the same way as \texttt{drop}, except that you specify the variables or observations to be kept rather than the variables or observations to be deleted.
Warning: `drop` and `keep` are not reversible. Once you have eliminated observations, you cannot read them back in again. You would need to go back to the original dataset and read it in again. Instead of applying `drop` or `keep` for a subset analysis, consider using `if` or `in` to select subsets temporarily. This is usually the best strategy. Alternatively, applying `preserve` followed in due course by `restore` may be a good approach.

### Remarks and examples

You can clear the entire dataset by typing `drop _all` without affecting value labels, macros, and programs. (Also see [U] 12.6 Dataset, variable, and value labels, [U] 18.3 Macros, and [P] program.)

#### Example 1

We will systematically eliminate data until, at the end, no data are left in memory. We begin by describing the data:

```
. use http://www.stata-press.com/data/r13/census11
   (1980 Census data by state)
. describe
```

```
obs: 50 1980 Census data by state
vars: 15 6 Apr 2013 15:43
size: 3,300

variable name type     display       value
      label label
-------------------------
state     str14  %-14s     State
state2    str2   %-2s       Two-letter state abbreviation
region    int    %-8.0g     cenreg Census region
pop       long   %12.0gc    Population
poplt5    long   %12.0gc    Pop, < 5 year
pop5_17   long   %12.0gc    Pop, 5 to 17 years
pop18p    long   %12.0gc    Pop, 18 and older
pop65p    long   %12.0gc    Pop, 65 and older
popurban  long   %12.0gc    Urban population
medage    float  %9.2f     Median age
death     long   %12.0gc    Number of deaths
marriage  long   %12.0gc    Number of marriages
divorce   long   %12.0gc    Number of divorces
mrgrate   float  %9.0g     Migration rate
dvcrate   float  %9.0g     Divorce rate
```

Sorted by: region

We can eliminate all the variables with names that begin with `pop` by typing `drop pop*`:
Let's eliminate more variables and then eliminate observations:

```
. drop marriage divorce mrgrate dvcrate
. describe
```

Let's now drop the first 2 observations in our dataset:

```
. drop in 1/2
(2 observations deleted)
```
Finally, let’s get rid of everything:

```
. drop _all
. describe
Contains data
  obs: 0
  vars: 0
  size: 0
Sorted by:
```

Typing `keep in 10/1` is the same as typing `drop in 1/9`.

Typing `keep if x==3` is the same as typing `drop if x !=3`.

`keep` is especially useful for keeping a few variables from a large dataset. Typing `keep myvar1 myvar2` is the same as typing `drop` followed by all the variables in the dataset except `myvar1` and `myvar2`.

警惕：

除了删除变量和观测值外，`drop _all` 还会删除任何业务日历；参见[D] datetime business calendars.

参考文献


另外参见

[D] clear — Clear memory
[D] varmanage — Manage variable labels, formats, and other properties
[U] 11 Language syntax
[U] 13 Functions and expressions