

Title

isid — Check for unique identifiers

[Syntax](#)[Remarks and examples](#)[Menu](#)[Also see](#)[Description](#)[Options](#)

Syntax

```
isid varlist [using filename] [, sort missok]
```

Menu

Data > Data utilities > Check for unique identifiers

Description

`isid` checks whether the specified variables uniquely identify the observations.

Options

`sort` specifies that the dataset be sorted by *varlist*.

`missok` indicates that missing values are permitted in *varlist*.

Remarks and examples

stata.com

▶ Example 1

Suppose that we want to check whether the mileage ratings (`mpg`) uniquely identify the observations in our `auto` dataset.

```
. use http://www.stata-press.com/data/r13/auto
(1978 Automobile Data)
. isid mpg
variable mpg does not uniquely identify the observations
r(459);
```

`isid` returns an error and reports that there are multiple observations with the same mileage rating. We can locate those observations manually:

```
. sort mpg
. by mpg: generate nob = _N
. list make mpg if nob > 1, sepby(mpg)
```

	make	mpg
1.	Linc. Mark V	12
2.	Linc. Continental	12
(output omitted)		
68.	Dodge Colt	30
69.	Mazda GLC	30
72.	Datsun 210	35
73.	Subaru	35



▷ Example 2

`isid` is useful for checking a time-series panel dataset. For this type of dataset, we usually need two variables to identify the observations: one that labels the individual IDs and another that labels the periods. Before we set the data using `tsset`, we want to make sure that there are no duplicates with the same panel ID and time. Suppose that we have a dataset that records the yearly gross investment of 10 companies for 20 years. The panel and time variables are `company` and `year`.

```
. use http://www.stata-press.com/data/r13/grunfeld, clear
. isid company year
```

`isid` reports no error, so the two variables `company` and `year` uniquely identify the observations. Therefore, we should be able to `tsset` the data successfully:

```
. tsset company year
panel variable:  company (strongly balanced)
time variable:  year, 1935 to 1954
delta: 1 year
```



□ Technical note

The `sort` option is a convenient shortcut, especially when combined with `using`. The command

```
. isid patient_id date using newdata, sort
```

is equivalent to

```
. preserve
. use newdata, clear
. sort patient_id date
. isid patient_id date
. save, replace
. restore
```



Also see

- [D] **describe** — Describe data in memory or in file
- [D] **ds** — List variables matching name patterns or other characteristics
- [D] **duplicates** — Report, tag, or drop duplicate observations
- [D] **lookfor** — Search for string in variable names and labels
- [D] **codebook** — Describe data contents
- [D] **inspect** — Display simple summary of data's attributes