

Can I use the 'by' and 'sort' commands together in one command in Stata?

Authored by
stats writer

June 30, 2024

RECOMMENDED CITATION

stats writer (2024). *Can I use the 'by' and 'sort' commands together in one command in Stata?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=163023>

The 'by' and 'sort' commands are both commonly used in Stata for data manipulation and analysis. The 'by' command allows for performing operations on subsets of data based on a specific variable, while the 'sort' command rearranges the data in a specified order. It is possible to combine these two commands in one command in Stata, allowing for the sorting of data by subsets of a variable. This can be useful for analyzing and visualizing data by specific categories.

Can I do by and sort in one command? | Stata FAQ

When you want to use by var:

command you need to first sort var before using the by, however, it is possible to the by and the sort into a single command.

Let's illustrate use of the process using the high school and beyond dataset.

**use <https://stats.idre.ucla.edu/stat/stata/notes/hsb2>,
clear
(highschool and beyond (200 cases))**

If we wanted summary statistics for each ses group, we would need to use sort and then by as illustrated below.

sort ses

by ses: summarize read write

-> **ses = low**

Variable | Obs Mean Std. Dev. Min Max

read | 47 48.2766 9.342987 28 68
write | 47 50.61702 9.490391 31 65

-> **ses = middle**

Variable | Obs Mean Std. Dev. Min Max

read | 95 51.57895 9.425609 31 73
write | 95 51.92632 9.106044 31 67

-> **ses = high**

Variable | Obs Mean Std. Dev. Min Max

read | 58 56.5 10.85834 34 76

write | 58 55.91379 9.442874 33 67

Here is one way to accomplish this with in a single step.

by ses, sort: summarize read write

-> ses = low

Variable | Obs Mean Std. Dev. Min Max

-----+-----

read | 47 48.2766 9.342987 28 68

write | 47 50.61702 9.490391 31 65

-> ses = middle

Variable | Obs Mean Std. Dev. Min Max

-----+-----

read | 95 51.57895 9.425609 31 73

write | 95 51.92632 9.106044 31 67

-> ses = high

Variable | Obs Mean Std. Dev. Min Max

```
-----+-----
read | 58 56.5 10.85834 34 76
write | 58 55.91379 9.442874 33 67
```

Alternatively, it is possible to combine by and sort into a single bysort

command. The bysort command looks like this:

bysort ses: summarize read write

-> ses = low

Variable | Obs Mean Std. Dev. Min Max

```
-----+-----
read | 47 48.2766 9.342987 28 68
write | 47 50.61702 9.490391 31 65
```

-> ses = middle

Variable | Obs Mean Std. Dev. Min Max

```
-----+-----  
read | 95 51.57895 9.425609 31 73  
write | 95 51.92632 9.106044 31 67
```

```
-----  
-----  
-> ses = high
```

Variable | Obs Mean Std. Dev. Min Max

```
-----+-----  
read | 58 56.5 10.85834 34 76  
write | 58 55.91379 9.442874 33 67
```

For More Information

For more information, see the help or reference manual about the sort command and by.