

---

Subject: Re: HELP!: Analysis on youth-specific age group only

Posted by [user-rhs](#) on Wed, 19 Mar 2014 20:31:21 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

malayaka wrote on Wed, 19 March 2014 15:55: Is it possible for me to merge these two (or is it necessary?).

Depends on what you're trying to answer. If you are trying to see agreement in response between spouses, then definitely merge these on household unique identifier. If not, keeping them separate should be fine.

malayaka wrote on Wed, 19 March 2014 15:55: The women variables are different than that of the men, v102 and mv102 respectively.

As you continue to work with DHS data, you will find that DHS is very good about keeping variable naming conventions so that you can figure out whether the variable pertains to the woman (prefix v), man (prefix mv), household (prefix hv), child (prefix b), maternal questions (m), local variables (s), and so on.

malayaka wrote on Wed, 19 March 2014 15:55: how do I make this "new" dataset specifically only for 15-19, 15-24 and 20-24 year olds?

You can use the Stata command `-keep-` and the age variable to drop observations outside of your age range.

`keep if v012>=35` will delete everyone who is younger than or equal to 35 from your dataset

\*Important: Stata is case-sensitive. All built-in commands are in lowercase. Most user-written commands are also lowercase.

malayaka wrote on Wed, 19 March 2014 15:55: and be able to get the breakdown of my variables (education, fertility, etc) per wealth quintile as well as region? The `-tab-` command is used to get tabulations (and cross-tabulations). syntax is `tab rowvbl colvbl`

`tab v130 v106` gets you the tabulation between religion (row) and education (column):

```
. tab v130 v106
```

```
.....|.....highest.educational.level
...religion.|no.educat...primary..secondary.....higher.|.. ...Total
-----+-----+-----
...orthodox.|.....2,836.....2,582.....910.....667.|.. ...6,995.
...catholic.|.....77.....81.....7.....12.|.. ...177.
...protestant.|.....1,212.....1,366.....209.....149.|.. ...2,936.
...muslim.|.....3,974.....1,777.....266.....153.|.. ...6,170.
...traditional.|.....69.....22.....1.....1.|.. ...93.
...other.|.....107.....26.....1.....2.|.. ...136.
...99.|.....3.....4.....1.....0.|.. ...8.
-----+-----+-----
.....Total.|.....8,278.....5,858.....1,395.....984.|.. ...16,515
```

Be aware that missing variables are coded as 99 and you need to recode it into system missing (.) for Stata to recognize it as a missing value. UCLA has an excellent resource on how to get started on Stata (Link: <http://www.ats.ucla.edu/stat/stata/sk/> ), and I encourage you to spend some time on their site to figure out how to do what you need to do.

Stata has great documentation Stata commands. You simply need to type help and the command name you want to find out more about and a window will pop up showing you the syntax, the options for that command, and examples at the bottom of the page.

Again, StatCompiler (<http://www.statcompiler.com/>) might be able to get you the numbers you need without you having to write any commands, so try that first.

Good luck.

RHS

---