Subject: Merging from multiple countries
Posted by no7321 on Tue, 30 Oct 2018 22:47:11 GMT
View Forum Message <> Reply to Message

Hi,

I created a pooled dataset with the most recent IR files of 6 South Asian countries (Afghanistan, Bangladesh, India, Maldives, Nepal, Pakistan)

I realized that I also needed a variable from the household files (how households treated their water, i.e. boiled etc.) and so wanted to merge HR data files with the pooled IR data set I created. I first created a pooled HR dataset with all of the same countries and then executed the following code:

```
/*merge HR to IR*/
use "E:\South Asia\dta\sasiahr.dta"
gen v000=hv000
gen v001=hv001
gen v002=hv002
gen v003=hv003
sort v000 v001 v002 v003
save "E:\South Asia\dta\sasiahr.dta", replace

use "E:\South Asia\dta\sasia.dta"
sort v000 v001 v002 v003
merge m:m v000 v001 v002 v003 using "E:\South Asia\dta\sasiahr.dta"
```

However, there are still many individual women from the IR dataset that are not matched to household records. The results from the merge show the following:

```
 Result                      # of obs.
    -----------------------------------------
    not matched                   935,707
    from master                   521,319  (_merge==1)
    from using                    414,388  (_merge==2)

    matched                       259,242  (_merge==3)
```

I'm wondering why I have 521,319 women from my master file (the IR file) not merged with the HR files? I thought all women interviewed would have a household record I could match them to.

Thanks for the help.