Subject: Merging Mali 2001 HW to KR
Posted by Khandys Ag. on Sat, 14 Dec 2024 08:22:09 GMT
View Forum Message <> Reply to Message

Greetings,

I hope you are all doing well. I am attempting to merge the Mali 2001 HW dataset with the Mali 2001 KR dataset. I followed the merging procedure provided with the HW dataset and I was able to merge the dataset to the household recode file. However, when I try to merge it with the children recode dataset I get the following error in STATA:

variables v001 v002 b16 do not uniquely identify observations in the using data.

Similarly, I had the same issue with the Senegal 2005 HW dataset. I was able to troubleshoot my problem, by doing drop if _merge ==2 when I merged the HW recode onto the household member recode. I was then able to merge this household recode file with the HW observations onto the Sengal file, but this led to the following merge results :

Result	Number of obs	
Not matched	8,325	
from master	7,696	(_merge==1)
from using	629	(_merge==2)
Matched	3,248	(_merge==3)

If possible, I would greatly appreciate it if you could confirm whether my merging procedure for Senegal 2005 was done properly and provide guidance and how best to code the Mali 2001 datasets. Any guidance on how to merge these datasets would be greatly appreciated, as my previous procedure (without using the drop if _merge==1) seems to have worked on other datasets. I only experienced this issue with the Mali 2001 datasets and partially with the Senegal 2005 datasets.

Thank you for your help.

Subject: Re: Merging Mali 2001 HW to KR
Posted by Bridgette-DHS on Mon, 16 Dec 2024 13:38:52 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS staff member, Tom Pullum:

To merge a KR and HW file, you first drop from the KR file any children for whom b16 is 0 or a dot (NA). These are children who are living elsewhere or have died. You also have to extract the household id code from the woman's id code. In the Mali 2001 survey, the household id code is a 12-character string, and is the first 12 characters of the woman's id code (caseid) which is a 15-character string.

The following lines will do this for the Mali 2001 survey.

```
use "...MLKR41FL.DTA", clear
drop if b16==. | b16==0
gen hwhhid=substr(caseid,1,12)
gen hwline=b16
merge 1:1 hwhhid hwline using "...MLHW41FL.DTA"
```

You then get the distribution of _merge as follows:

```
. tab _merge

                _merge |      Freq.     Percent        Cum.
-----------------------+-----------------------------------
       master only (1) |          5        0.04        0.04
        using only (2) |      1,554       12.70       12.74
          matched (3) |     10,682       87.26      100.00
-----------------------+-----------------------------------
                 Total |     12,241      100.00
```

The 5 cases with _merge=1 are in the KR file but not the HW file. The 1,554 children with _merge=2 are in the HW file but not the KR file, and the 10,682 children with _merge=3 are in both files. If your goal is to add the anthropometry variables to the KR file, and then continue to analyze all the cases in the original KR file, you would drop the cases with _merge=2 and then drop _merge. You may want to add back in the children for whom b16 is 0 or a dot, who were dropped before the merge, depending on what other KR variables you are working with.

Please try something similar for the Senegal survey.