
Subject: Merging Household+Female+Male files
Posted by [UAB_user](#) on Tue, 13 Jan 2015 21:11:20 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello,

Sorry for a double post, but I tried posting this in the "Nepal" section of the forum, but i noticed that several threads have gone unanswered since November and I got nervous my question would be forgotten....

I am looking at male migration in the 2011 Nepal DHS and I want to merge the female file (IR) with the household file (HR) and then merge this to the male file (MR). I know you have to match the male and female files on V001 + V002 + V003 and then match that new data set to the household file using V001 + V002, like what is recommended on the DHS web page, but I am not sure if I did it right since the final number for each merge do not seem to add up.

The Household file has 10826 observations and the female file has 12674 observations. When I merge them with the following SAS code I get 14588 observations.

```
data derek.full2011;  
merge derek.nphr60fl (in=a) derek.npir60fl (in=b) ; by V001 V002 ; run;
```

The male file has 4121 observations in it and when I merge it into the file above (14588 obs) using the following SAS code I get 18709 observations, which I think means none of the male file merged into my household + female files

```
data derek.full2011b;  
merge derek.full2011 (in=a) derek.npmr60fla (in=b); by V001 V002 V003;  
run;
```

Thank you for the help
Derek

Subject: Re: Merging Household+Female+Male files
Posted by [Liz-DHS](#) on Thu, 15 Jan 2015 16:19:12 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear Derek,
I have forwarded your message to one of our experts. We should have a response in a day or two. Thank you for posting again.

Subject: Re: Merging Household+Female+Male files
Posted by [Liz-DHS](#) on Thu, 15 Jan 2015 17:05:46 GMT

Dear Derek,

Here is a response from one of our experts, Mr. Albert Themme,

Quote:Hello Derek,

I am not familiar with SAS, but your first merge between household and individual women's data looks correct. The result of the merge should be all individual women with household information added to them (12674) and some cases with only household information and without individual woman data, because no women were interviewed in these households. The identification variables are indeed V001 (cluster), V002 (household) and V003 (woman's line number)

You then went on to merge this file with the male data file. Here the identification variables are MV001 (cluster), MV002 (household) and MV003 (man's line number) renamed to V001, V002 and V003. Since the line-numbers of all household members in the same household are unique, V003 and MV003 will never be the same. I am not sure what happens when you merge the man's data to the already merged household/woman's file, but believe that the man's data will simply be added to the household/woman's file.

Depending on your desired unit of analysis: if you want all women and men as unit of analysis with the household information added to them, then I would first add the women and man's data together ($10826 + 4121 = 14947$ cases) and then merge this with the household data. In this case you will get a file with all interviewed women and men as separate cases with the household information added. If you want to create couples files then variable V034 needs to be used (linked to MV003) and then the household data could be added.

It would be more difficult to create and analyze household information, with all data for interviewed women and men in that household added to the household information (in this case you would have 10826 cases).

Let us know if this is not clear or does not answer your questions.

Subject: Re: Merging Household+Female+Male files
Posted by [UAB_user](#) on Thu, 15 Jan 2015 21:34:42 GMT
[View Forum Message](#) <> [Reply to Message](#)

This was very helpful! thank you very much!

Derek
