
Subject: discrepancies after merging PR and IR/MR
Posted by [fabienne](#) on Wed, 05 Aug 2015 16:10:14 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello everyone

I am working with the DHS VI from Liberia.

I appended the women file (IR) and the men file (MR) and merged it to the household member file (PR), which resulted in 48,219 observations and 13,357 matches (merged on: v001 v002 v003).

To check the success of the merging process, I looked at the consistency of variables which are given in both the individual (IR/MR) datasets and the household (PR) dataset:

- hv105 and (m)v012 (Age of respondent)
- hv106 and (m)v106 (highest educational level)

Unfortunately, I found 2990 age and 691 education discrepancies. Commonly, the age discrepancies are small (1 year). Does this indicate a failed merge (which I do not believe) or is there a methodological explanation for this effect? Which age and education variable is more trustworthy and should be used for further analysis?

Thank you and best regards, Fabienne

Subject: Re: discrepancies after merging PR and IR/MR
Posted by [Bridgette-DHS](#) on Fri, 07 Aug 2015 14:52:20 GMT
[View Forum Message](#) <> [Reply to Message](#)

Response from Senior Data Processing Specialist, Nouredine Abderrahim:

You need to ensure that the file constructed from the men and the women's data is complete. You need also to keep in mind that the person's file does not include only the individuals who were interviewed but also members who were not interviewed such as children, non defacto members, etc.

Also, be aware that the data in the household roster (such as HV105, HV106, ..) might be different from the individual data (V012, MV012, V106, MV106) because the respondent to the household interview might be different from the respondent to the individual interview.

To check that the merge was done correctly, you need to see if the data of the women and men in the new file, matches the data in the individual data files (such as V106 in the new data file matches V106 in the individual data file, the same for MV106 in the men's data file).
