
Subject: merging household and individual data
Posted by [DHS user](#) on Tue, 14 Jun 2016 14:09:01 GMT
[View Forum Message](#) <> [Reply to Message](#)

I have tried to merge the Nigeria 2013 household and individual data (which contains the appended male and female recode) but It doesn't merge.

I have asked some colleagues and was told it's impossible to get the living arrangement of youth aged 15-24 in a household as the question was only asked for children below age 15. I was also told that I could use the PR recode alone for the analysis without having to link with individual files but I am struggling with that.

I would like to get some help on this issue.

Subject: Re: merging household and individual data
Posted by [Bridgette-DHS](#) on Tue, 14 Jun 2016 14:11:05 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

All kinds of merges ARE possible with those files. Please confirm that what you want to construct is a file with the MR and IR data merged onto the PR file for 2013. Is that correct? If so, tell me and I will send the lines for the merges.

The parental survival and coresidence questions are only asked about children under 18, i.e. ages 0-17. These are given as the variables hv111-hv114 in the household file. It is true that for children under 18 you do not need to do any merging at all; you can just use the PR file.

People age 18 and above can be matched with their mother if their mother is alive and age 15-49 and in the household. This is because in a woman's birth history, b16 (for each child) is the line number of the child if the child lives in the same household as the mother. However, if the child and the mother are NOT in the same household, then it could be because the mother is older than 49, or has died, or is in a different household. It is impossible to say which.

There is no way to link with the father unless the father is the household head and the child is coded "child" for hv101. If the father is in the household but is not the household head, then you cannot identify him as the father, and you cannot tell whether the father has died or is in a different household. It is impossible to say which.

There have been some other surveys, for example some MICS surveys, in which the parental survival and coresidence questions go higher than age 17.

Subject: Re: merging household and individual data

Posted by [DHS user](#) on Tue, 14 Jun 2016 14:12:46 GMT

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

I want to construct a file with the MR and IR merged into PR file but it may not help me as I am interested in youth aged 15-24.

This leads to the second question and your suggestion. Using the PR file. How do I link the youth to the mother's in the PR file. I have attempted to do that but really confused. I am interested in education completed in single years for youth aged 15-24 . When I tab, I get 178,894. I then tab sex and get 30,221. Does this mean there are 30,221 males and females in the households. How do I link these youth to the households and also operationalize their living arrangement.

I hope I have been able to present my issues in the simplest form.

I would appreciate your response.

Subject: Re: merging household and individual data

Posted by [Bridgette-DHS](#) on Tue, 14 Jun 2016 14:14:58 GMT

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

Another response from Tom Pullum:

What I had in mind is given below. I said this would require the IR file. It COULD be done with the IR file, but it is easier with the BR file, and merging it with the PR file.

The file that is produced has a variable "line_number_of_mother". This is equivalent to hv112 if the child and mother are in the same household and the child's age is <18. However, it extends to ages 18 and above.

If the child is age 18+ and the child's mother is NOT in the same household as the child, then it is impossible to tell whether the mother is still alive. That information is provided with hv111 for children age <18. Let me know if you have questions. Cheers--Tom

* 2013 Nigeria DHS NG6A

* linking mothers and children in the same household

* The goal is to add a variable to each child in the PR file that will match

* hv112 for children age 0-17 but will continue for ages 18+

* Use BR file, merge with PR

* merge the mother's line number with the child

set more off

use e:\DHS\DHS_data\BR_files\NGBR6AFL.dta, clear

```
keep v001 v002 v003 b16
keep if b16>0 & b16<90
```

```
gen line_number_of_mother=v003
```

```
ren v001 hv001
ren v002 hv002
ren b16 hvidx
keep hv001 hv002 hvidx line_number_of_mother
sort hv001 hv002 hvidx
save e:\DHS\DHS_data\scratch\temp_child.dta, replace
```

* prepare the PR file, just keep the minimum variables needed for testing

```
use e:\DHS\DHS_data\PR_files\NGPR6AFL.dta, clear
```

```
keep hv001 hv002 hvidx hv104 hv105 hv111 hv112
```

```
sort hv001 hv002 hvidx
```

```
merge hv001 hv002 hvidx using e:\DHS\DHS_data\scratch\temp_child.dta
ren _merge merge_child
```

```
summarize hv105 if line_number_of_mother>0 & line_number_of_mother<.
```

* line number of mother should match hv112 for children under 18

```
correlate hv112 line_number_of_mother if hv105<18 & hv112>0 & hv112<90
```

```
tab line_number_of_mother if hv105>=18
```

```
tab hv105 if hv105>=18 & line_number_of_mother>0 & line_number_of_mother<.
```

Subject: Re: merging household and individual data
Posted by [DHS user](#) on Tue, 14 Jun 2016 14:16:10 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks again. I have run the codes but still a bit confused. My question may seem out of place. Would I be merging this new file with the IR and MR or can I go ahead to use the PR recode. This is because the PR does not have some variables I am interested in (number of siblings, religion). I am also confused as to how to use the variables. For e.g, can I assume hv105 to be the age of respondents?

Subject: Re: merging household and individual data
Posted by [Bridgette-DHS](#) on Tue, 14 Jun 2016 14:17:16 GMT
[View Forum Message](#) <> [Reply to Message](#)

Another response from Tom Pullum:

What I described is based on a merge of the BR and PR files. You can keep whatever variables you want from those files as part of the "keep" lines. You can merge variables from the IR and MR files onto the PR file, either before or after the merge with the BR file.

The general practice is to prefer age from the IR and MR files, i.e. v112 and mv112, over age in the PR file, i.e. hv105. hv105 comes from the household respondent and may not be as accurate as age that comes from the interview with the person. Similarly, for a child, b8 would be preferred to hv105, because b8 comes from the mother.
