
Subject: Merged Household and Domestic Violence dataset
Posted by [Aye Thiri Kyaw](#) on Thu, 04 Mar 2021 17:39:39 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi
This is Aye Thiri
My colleague and I: We have a merged dataset for Household and women individual data.

Question 1

Our question "Is there any way we can tell that those who answered the Household questionnaires is women or men"?

Question 2: Is there a way we can tell the age of the child for the women/men who answered the child discipline questions?

Kindest Regards

Subject: Re: Merged Household and Domestic Violence dataset
Posted by [Bridgette-DHS](#) on Sat, 06 Mar 2021 00:23:12 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

First question: The line number of the household respondent is given by hv003 for everyone in the household. The sex of that person is given by hv104. If you want to attach that to everyone in the household, you can use these lines in Stata, with the PR file:

```
gen temp=hv104 if hvidx==hv003
list hv001 hv002 hv003 hvidx hv104 temp if _n<=100, table clean
egen hhresp_hv104=min(temp), by(hv001 hv002)
drop temp
list hv001 hv002 hv003 hvidx hv104 hhresp* if _n<=100, table clean
```

The "list" lines will show you how the gen and egen lines work. You can extend this to copy other characteristics of the household respondent onto every every person in the household.

Second question: The relevant line number for the child who was selected for the child discipline module in the Myanmar survey was sh148. As I said the other day, sh148-sh161 were copied onto the record for every person in the household. You can restrict to the reference children just with "keep if hvidx==sh148" and then you have the age of the child as hv105. Alternatively, if you just want to make those children more identifiable but keep everyone in the PR file, you could do the following to make the sh variables ".", for NA, in the PR file:

```
local cdvars sh148 sh149 sh150 sh151 sh152 sh153 sh154 sh155 sh156 sh157 sh158 sh159
sh160 sh161
list hv001 hv002 hvidx hv105 sh148-sh152 if _n<=100, table clean
```

```
foreach cdvar of local cdvars {
  replace `cdvar'= . if hvidx~=sh148
}
list hv001 hv002 hvidx hv105 sh148-sh152 if _n<=100, table clean
```

Again, the "list" lines are just for checking.

Subject: Re: Merged Housedhold and Domestic Violence dataset

Posted by [smiedema](#) on Mon, 15 Mar 2021 18:16:47 GMT

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

Thank you Tom and Bridgette. This is very useful.

Unfortunately, the provided code does not work because there are multiple possible variables for hvidx, hv104 etc. When I tried to add the `_*`, the code does not run and I continue to receive "ambiguous abbreviation" error messages from Stata.

In addition, to add to my colleagues previous note, we are not only trying to determine whether the individual who answered the household survey was a woman, but also whether she was the woman who was selected for the domestic violence module. Do you have any insight into either of these issues?

Thank you very much for your assistance.

Best,
Stephanie

Subject: Re: Merged Housedhold and Domestic Violence dataset

Posted by [smiedema](#) on Mon, 15 Mar 2021 19:40:57 GMT

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

Hi Tom and Bridgette,

First - please ignore the question about code below. I used the PR file, and it worked. The second question still remains. We would like to know whether it is possible to restrict the sample to women who answered both the household survey (e.g. reported their own use of harsh child discipline against the selected child) and the domestic violence module questions. I do not see line items for the domestic violence module, so am not sure this is possible?

Thanks,
Stephanie

Subject: Re: Merged Housedhold and Domestic Violence dataset
Posted by [Bridgette-DHS](#) on Tue, 16 Mar 2021 19:09:40 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

The line number of the respondent, v003, matches her line number in the household survey, hvidx. The line number of the household respondent (in the PR file) is hv003. In the merged PR/IR file, if v003=hv003, then the woman was the household respondent. Hope this will answer the question.

Subject: Re: Merged Housedhold and Domestic Violence dataset
Posted by [smiedema](#) on Wed, 17 Mar 2021 17:41:44 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you very much for this response and clarification. It is much appreciated. Best, Stephanie
