Subject: Merging household and women's individual data Myanmar 2015-2016
Posted by smiedem on Wed, 06 Jan 2021 22:21:10 GMT
View Forum Message <> Reply to Message

Hi, I'm trying to merge the household and women's individual data using the Myanmar DHS 2015-2016 using code provided in previous forum posts. However, it appears as though only 4,942 of 12,885 women interviewed were matched with household data. Below is the code I used (saving the hh data file as mdhs_hh and the women's file as mdhs_women). I used the women's data set as the base file, as recommended on the DHS Program merging datasets webpage. Any advice for why so few of the women are merging with the household data file? I checked the MDHS report and I could find no indication of why this may be. Thanks for any advice you may have.

use "C:\Users\qlu4\Desktop\mdhs_hh.dta"
gen v001=hv001 //recode hh survey cluster number variable to match women's data set
gen v002 = hv002 //recode hh survey household number variable to match women's data set
gen v003 = hv003 //recode hh
keep v001 v002 v003 sh148 sh148f sh149 sh150 sh151 sh152 sh153 sh154 sh155 sh156 sh157 sh158 sh159 sh160 sh161 //keep only relevant variables
sort v001 v002 v003
save "C:\Users\qlu4\Desktop\mdhs_hh.dta", replace

use "C:\Users\qlu4\Desktop\mdhs_women.dta"
sort v001 v002 v003
merge 1:m v001 v002 v003 using "C:\Users\qlu4\Desktop\mdhs_hh.dta"

Result # of obs.

not matched   15,501
from master   7,943 (_merge==1)
from using   7,558 (_merge==2)

matched   4,942 (_merge==3)

I also tried merge 1:1 v001 v002 v003 using "C:\Users\qlu4\Desktop\mdhs_hh.dta", which gave me the same results...

Best,
Stephanie

Subject: Re: Merging household and women's individual data Myanmar 2015-2016
Posted by Bridgette-DHS on Thu, 14 Jan 2021 18:50:20 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Specialist, Kerry MacQuarrie:

Dear Stephanie,
You have renamed and matched the cluster and household number appropriately. However, hv003 does not match with v003. The variable hv003 refers to the line number of the person who answered the household questionnaire, not the line number of all individual women. You don't need it. Try the following:

use "C:\Users\qlu4\Desktop\mdhs_hh.dta"
gen v001=hv001 //recode hh survey cluster number variable to match women's data set
gen v002 = hv002 //recode hh survey household number variable to match women's data set
keep v001 v002 sh148 sh148f sh149 sh150 sh151 sh152 sh153 sh154 sh155 sh156 sh157 sh158 sh159 sh160 sh161 //keep only relevant variables
sort v001 v002
save "C:\Users\qlu4\Desktop\mdhs_hh.dta", replace

use "C:\Users\qlu4\Desktop\mdhs_women.dta"
sort v001 v002
merge 1:m v001 v002 using "C:\Users\qlu4\Desktop\mdhs_hh.dta"

In the Myanmar DHS, this matches 12,885 cases and leaves 3,095 cases in the HR file unmatched (ie households that did not have a corresponding eligible woman in the IR file)

File Attachments
1) mrg.png, downloaded 937 times

Subject: Re: Merging household and women's individual data Myanmar 2015-2016
Posted by anikhpg42@gmail.com on Fri, 22 Jan 2021 08:37:43 GMT
View Forum Message <> Reply to Message

Dear Kerry MacQuarrie,
For merging two recent datasets of BDHS, i.e. BDPR18 and BDIR18, I have applied the same command that you indicated. But, the result doesn't match. the error status showed "variables v001 v002 do not uniquely identify observations in the master data". What am I missing?

I have used the following commands:

clear all
set maxvar 30000
set more off

```
use "D:\data\BDPR18.DTA"
gen v001=hv001 //recode PR survey cluster number variable to match women's data set
gen v002 = hv002 //recode PR survey household number variable to match women's data set
keep v001 v002 hv005 hv022 hv021 sb333aa sb333ab sb318a sb335b sb327a
sort v001 v002
save "D:\data\BDPR18_MERGED.dta", replace

use "D:\data\BDIR18.DTA"
sort v001 v002
merge 1:m v001 v002 using "D:\data\BDPR18_MERGED.dta"
```

Subject: Re: Merging household and women's individual data Myanmar 2015-2016
Posted by Bridgette-DHS on Fri, 22 Jan 2021 14:29:56 GMT
View Forum Message <> Reply to Message

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

You need to include line number. In the PR file that's hvidx and in the IR file it's v003.

```
use "/Users/t/Desktop/Actual dataset/PR2.DTA"

rename hv001 v001
rename hv002 v002
rename hvidx v003
sort v001 v002 v003
save "/Users/Desktop/Actual dataset/PR2.DTA", replace

use "/Users/t/Desktop/Actual dataset/IR.DTA"
sort v001 v002 v003

merge m:1 v001 v002 v003 using PR2
```

Subject: Re: Merging household and women's individual data Myanmar 2015-2016
Posted by anikhpg42@gmail.com on Sat, 23 Jan 2021 01:54:32 GMT
View Forum Message <> Reply to Message

Thank you very much, Dr. Tom Pollum