
Subject: Issue merging NFHS 2 women and household dataset

Posted by [Mrinal](#) on Tue, 30 Jan 2018 08:45:33 GMT

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

I am trying to merge the women and the household dataset of NFHS 2. However, after merging the results are not making any sense!! NFHS 2 women dataset has 90,303 observations and household dataset contains 92,486 observations but after merging the merged file shows 104,979 observation, I am trying to understand the result, is it correct or the merging has some issue?? I am also getting an error saying

Quote: variables v001 v002 do not uniquely identify observations in the master data
variables v001 v002 do not uniquely identify observations in
D:\Dropbox\stata\nfhs\IAHR42FL_sort.DTA

Complete result of merging and codes used to merge is given below for reference.

"_merge" result

tab _merge

_merge	Freq.	Percent	Cum.
-----+-----			
2	5,722	5.45	5.45
3	99,257	94.55	100.00
-----+-----			
Total	104,979	100.00	

Merging codes

****Merging household on women dataset****

****Round 2****

use "D:\Desktop\dhs\data\nfhs\2\IAHR42FL.DTA", clear

gen int v001 = hv001

gen int v002 = hv002

sort v001 v002

save "D:\Dropbox\stata\nfhs\IAHR42FL_sort.DTA", replace

use "D:\Desktop\dhs\data\nfhs\2\IAIR42FL.DTA", clear

sort v001 v002

merge v001 v002 using "D:\Dropbox\stata\nfhs\IAHR42FL_sort.DTA"

save "D:\Desktop\dhs\data\nfhs\2\IA_HR_IR_42FL.DTA", replace

Thanks

Mrinal
