
Subject: merging HIV results and couples data

Posted by [lberes](#) on Mon, 11 Apr 2016 05:08:20 GMT

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

A 2014 forum post gives extremely helpful advice on how to merge the HIV data and couples data in STATA. The basic code is below. My question is on step 8. I am wondering why drop if `_merge==2` allows us to 'keep only women'. I thought that any combination of `v001`, `v002` and `v003` that came from either the couples or the HIV data could only be a woman. How does dropping that combination if it only came from the HIV dataset - and was not in the couples data set - exclude men? In other words, what does drop if `_merge==2` do for us in this code? Thank you!

code from post in Nov 2014 by Trevor-DHS

* Step 1: open AR file

use "xxAR61FL.DTA", clear

* Step 2: rename identifying variables

renvars hivclust hivnumb hivline / v001 v002 v003

* Step 3: sort by identifying variables

sort v001 v002 v003

* Step 4: save results

save "xxAR61FL_mergeprep.DTA", replace

* Step 5: open IR file

use "xxCR61FL.DTA", clear

* Step 6: sort by identifying variables

sort v001 v002 v003

* Step 7: merge!

merge v001 v002 v003 using "xxAR61FL_mergeprep.DTA"

* Step 8: Keep only women

drop if `_merge==2`

Then rename the added hiv variables to something unique for women, e.g.

rename hiv* w_hiv*

and repeat steps 1-8 above using mv003 instead of v003 throughout to merge the men's hiv test result and then finally rename the hiv variables to be for men, e.g.

rename hiv* m_hiv*