Subject: Merging IR, KR and HR files

Posted by Aarthi1410 on Mon, 05 Aug 2024 16:46:10 GMT

View Forum Message <> Reply to Message

I am trying to merge datasets for DHS India 2015-16 and 2019-21. I need variables from all three files since there are mother's education levels employment status and ownership of land/house included as controls which I am planning to include in my analysis. I use the following codes to merge the three for both rounds. For 2015-16, the code works fine, however for 2019-21, when I merge the IR and KR files _merge == 3 is only 333 and when I finally merge the HR file the total observation for mother's education and employment is approximately 2000. Is there some other way to merge in a correct manner?

use "\${cleaned}\cleaned_household_2015",clear

gen cluster=hv001 gen hh=hv002 sort cluster hh save HH_2015temp.dta, replace

use "\${cleaned}\cleaned_women_2015", clear gen cluster=v001 gen hh=v002 gen mo_line=v003 sort cluster hh mo_line save women2015_temp.dta, replace

use "\${cleaned}\cleaned_children_2015", clear gen cluster=v001 gen hh=v002 gen mo line=v003

sort cluster hh mo_line
quietly merge cluster hh mo_line using women2015_temp.dta
tab _merge
keep if _merge==3
drop _merge

* Merge the child and mother with the household data sort cluster hh quietly merge cluster hh using HH_2015temp.dta tab _merge keep if _merge==3 drop _merge