Subject: Re: Strange Issues w/ Data Formatting from DHS
Posted by tednoel on Wed, 01 May 2024 16:02:18 GMT
View Forum Message <> Reply to Message

Hi, thanks so much for the response. I've been a bit stumped by this merging process because there are three different datasets (IR, PR, and Wealth Index) that I have to combine and I'm a bit confused as to which one should be my base for merging. I was going to make the IR my base for merging because I've seen code that allows for the renaming of the cluster number, household number, and respondent's line number such as this:

use "/Users/tbear/Desktop/THESIS DATA/Tanzania_1999/Tanzania_1999_dta/TZIR41FL.dta", clear

* keep the variables you want
keep v0*
sort v001 v002 v003
save e:/Users/tbear/Desktop/THESISDATA/Tanzania_1999/Tanzania_199 9_dta/TZIR41FL.dta, replace

* Prepare PR file and merge
use "/Users/tbear/Desktop/THESIS DATA/Tanzania_1999/Tanzania_1999_dta/TZPR41FL.dta", clear
* reduce to women who are eligible for the IR file
keep if hv117==1

* keep the variables you want
keep hv0* sa33 sh*

rename hv001 v001
rename hv002 v002
rename hv003 v003

sort v001 v002 v003
merge v001 v002 v003 using ***Not entirely clear what I should be using here
tab _merge
******

BUT the problem is the wealth index only has the hhid variable that I can use to merge- and the IR file does not have this, only the PR file does. Should I be using the PR file as my base, merging the IR file, and then appending the wealth index?

Thank you so much for all of your help.