Subject: Re: urban-rural variable

Posted by Trevor-DHS on Fri, 24 Apr 2015 19:21:50 GMT

View Forum Message <> Reply to Message

It wasn't very clear from your explanation which survey and datasets you are using, but it appears that you are using the HIV tests results data, matched with other datasets. However it appears that there is a problem with the matching of your data. I would expect the number of women and men to be roughly equal, but you have twice as many men as women. Below is a short piece of code for matching with the persons recode file, and for checking the sex distribution:

cd "C:\Data\DHS Stata"

use "IAPR52FL.dta"

clonevar hivclust = hv001

clonevar hivnumb = hv002

clonevar hivline = hvidx

merge 1:1 hivclust hivnumb hivline using "IAAR51FL.dta"

keep if _merge==3

tab hv104

tab hv025

tab sh025

tab sh025 hv025

From this, I received the following output:

. tab hv104

sex of household member	•	Percent	Cum.
male female	51,641 54,016	48.88 51.12	48.88 100.00
Total	105,657	100.00	

. tab hv025

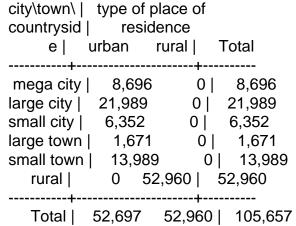
type of place of residence	Freq.	Percent	Cum.
urban rural	52,697 52,960	49.88 50.12	49.88 100.00
Total	105,657	100.00	

. tab sh025

city\town\c | ountryside | Freq. Percent Cum.

mega city	8,696	8.23	8.23
large city	21,989	20.81	29.04
small city	6,352	6.01	35.05
large town	1,671	1.58	36.64
small town	13,989	13.24	49.88
rural	52,960	50.12	100.00
Total	10E 6E7	100.00	
Total	105,657	100.00	

. tab sh025 hv025



As you can see in my tables (unweighted) there are about 49% men and 51% women, which is roughly what I would expect.

Also, as you can see the recoding of urban and rural makes sense in the last table.