Subject: Merging IR and PR - Indonesia 1987
Posted by albena on Tue, 31 Oct 2023 09:41:51 GMT
View Forum Message <> Reply to Message

Dear all,

I would like to merge the PR data file with the IR data file for Indonesia 1987. I am struggling with setting up the variables for merging in the PR data. In the IR data those are straightforward - v001 (sprov), v002 and v003.

Here is my Stata code to illustrate what I have so far with the PR data file.

use "..../IDHH01FL.dta", clear

gen id = _n

gen wt1 = 1 //generate self-weighting as the hhweight from PR cannot be combined with the wmweight from IR

keep hhage_* hhsex_* hhlno_* hhevm_* hhyear hhprov hhreg hhkeyer hhsamp hhsampno wt1 id
// no variable on whether hh member slept there last night -> difficult to reconstruct the sample of eligible women for the women's interview

foreach num of numlist 1/9 {

rename hhage_0`num' hhage_`num'

}

foreach num of numlist 1/9 {

rename hhsex_0`num' hhsex_`num'

}

foreach num of numlist 1/9 {

rename hhlno_0`num' hhlno_`num'

}

foreach num of numlist 1/9 {

rename hhevm_0`num' hhevm_`num'

```
}
reshape long hhage_ hhsex_ hhlno_ hhevm_, i(id) j(member)
drop if hhage_ ==. & hhsex_ ==. & hhlno_ ==. & hhevm_==.
keep if hhsex_ ==2
keep if hhage_ >=15 & hhage_ <=49
rename hhprov prov
rename hhsamp v002
rename hhlno_ v003
sort prov v002 v003
order prov v002 v003 member
save "..../IndonesiaPR1987.dta", replace

***** the IR data
use ".../IDIR01FL.dta", clear
keep v001 v002 v003 v005 v007 v010 v012 v013 v511 v806 sprov
gen wt=v005/1000000
rename v012 current_age
rename sprov prov
sort prov v002 v003
save ".../IndonesiaIR1987.dta", replace

**** Merge PR (master) with IR (using)
use ".../IndonesiaPR1987.dta", clear
merge 1:1 prov v002 v003 using "..../IndonesiaIR1987.dta"
```

This is the where I get an error message that I have duplicates in the variables for merging from the master data. I am pretty sure that I might have the wrong variables from the PR data for merging, but at this point I don't really which other ones to try or what other combination I could have.

Could you help me with merging these two datasets?

Thank you!

Albena

Subject: Re: Merging IR and PR - Indonesia 1987
Posted by Bridgette-DHS on Tue, 31 Oct 2023 14:27:12 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS staff member, Tom Pullum:

Many examples of merges have been given on the forum. At the end of this post I will paste the lines that I would normally use to do this kind of merge. I would reshape the HR file, to construct the equivalent to a PR file, then reduce the PR file to women (perhaps to women age 15-49) and then merge using the cluster ID, household ID, and line number. Unfortunately, in this HH file I cannot identify the cluster ID and household ID (the line number is hhlno). There are only a few unsubscripted variables in the HH file. I suspect that the ID variables have something to do with the three "*samp*" variables but I can't figure it out.

I don't think the approach you are using, without ID codes, will be reliable. For one thing, age in the household survey and age in the women's survey may not be exactly the same. There can be other inconsistencies that will result in ambiguous matches.

This is an interesting challenge, but the Indonesia 1987 DHS is one of our oldest surveys and we can't provide more support for it.

```
* Specify a workspace
cd e:\DHS\DHS_data\scratch

use "...IDHH01FL.DTA", clear
* must find the cluster id and household id

gen cluster=?hhnisamp
gen hh=?
```

```
rename *_0* *_*
keep cluster hh hhsex* hhage*

reshape long hhsex_ hhage_ .... ,i(cluster hh) j(line)
rename *_ *
keep if hhsex==2
sort cluster hh line
save ID01temp.dta, replace

use "...IDIR01FL.DTA", clear
gen cluster=v001
gen hh=v002
gen line=v003
sort cluster hh line
merge 1:1 cluster hh line using ID01temp.dta

tab _merge
keep if _merge==3
drop _merge
```