

Understanding Society User Support - Support #382

*_sex (not *_sex_cr) in egoalt file

06/12/2015 01:42 PM - Dave Griffiths

Status:	Closed	Start date:	06/12/2015
Priority:	Normal	% Done:	100%
Assignee:	Dave Griffiths		
Category:	Derived variables		
Description			
Hello, I've been looking at gender in the egoalt file and it appears that ego gender is based on *_sex, rather than *_sex_cr.			
I've found 54 cases with different *_sex and *_sex_cr values. In all cases, the ego-alt file contains data corresponding to *_sex, not *_sex_cr (thus, the miscoded data, rather than actual gender).			
My Stata syntax file spotting this problem is below (inresp gender only shown for years when _sex and _sex_cr are inconsistent).			
Dave			
global path1 "C:\data\under_society\data\"			
global path9 "C:\temp\"			
foreach wave in a b c d {			
use pidp `wave' _sex `wave' _sex_cr using ///			
\$path1\`wave'_indresp.dta, clear			
drop if `wave' _sex `wave' _sex_cr			
sort pidp			
save \$path9\sex`wave'.dta, replace			
}			
use \$path9\sexa.dta, clear			
foreach wave in b c d {			
sort pidp			
merge 1:1 pidp using \$path9\sex`wave'.dta			
drop _merge			
}			
save \$path9\sex_cr.dta, replace			
foreach wave in a b c d {			
use pidp `wave' _esex ///			
using \$path1\`wave'_egoalt.dta, clear			
duplicates drop			
save \$path9\ego`wave'.dta, replace			
}			
use \$path9\egoa.dta, clear			
foreach wave in b c d {			
sort pidp			
merge 1:1 pidp using \$path9\ego`wave'.dta			
drop _merge			
}			
sort pidp			
merge 1:1 pidp using \$path9\sex_cr.dta			
keep if _merge3			
list			

History

#1 - 06/12/2015 01:47 PM - Dave Griffiths

For some reason, the double equals signs aren't showing.

They should be in the sixth row
drop if `wave' _sex EQUALS EQUALS `wave' _sex_cr

And the penultimate row

keep if _merge EQUALS EQUALS 3

I suspect there are html reasons why displaying two equal signs together aren't showing.

#2 - 06/12/2015 02:36 PM - Redmine Admin

- *Category set to Derived variables*
- *Assignee set to Dave Griffiths*
- *Target version set to X M*
- *% Done changed from 0 to 50*

You are correct about the current source of the sex variable. We have so far only provided corrected sex and age variables for adult respondents based on their latest enumeration and retrospectively for the dates they gave an interview in previous waves. The latter is of course only important for age. The idea of using the latest enumeration for rule-based qc is that gender and date of birth are checked by interviewers each wave and that miscoding is thus expected to diminish over time. In the forthcoming November release we plan to extend the definitions of these variables to all, which we hope will be useful in cases such as the one you present. We also hope to at least provide some more guidance on this at that point and in the longer term incorporate this information into the checks of the relationships on the egoalt files.

FYI some links to Textile formatting can be found around here: <https://www.understandingsociety.ac.uk/support/projects/support/wiki>

On behalf of the team, Jakob

#3 - 06/23/2015 10:15 AM - Redmine Admin

- *Status changed from New to Closed*
- *% Done changed from 50 to 100*