Understanding Society User Support - Support #917

how to merge xwavedat with data from all the other seven waves

02/12/2018 10:05 AM - Nico Ochmann

Status:	Resolved	**Start date:**	02/12/2018
Priority:	High	**% Done:**	100%
Assignee:	Nico Ochmann		
Category:	Data management		

Description

Dear Alita,

this seems an obvious procedure, but I obtain a fairly high number of non-matches (60,792), which kind of concerns me.
This is what I do and I cannot detect an obvious mistake.
Your suggestions are highly appreciated, as usual.
Best. Nico

```
use pidp hhorig sex birthy feend_dv ukborn plbornc_all using "$Stata11_se/xwavedat_protect", clear
save  $Mergeddata_master2016/xwavedat, replace

foreach w in a b c d e f g {

use  pidp  `w'_istrtdaty `w'_jbhrs `w'_qfhigh_dv  `w'_dvage `w'_marstat  `w'_jbstat   ///
        `w'_paygu_dv `w'_fimnlabgrs_dv  `w'_jbsize `w'_jbsect  `w'_jbsemp  `w'_nnatch  ///
        `w'_gor_dv `w'_urban_dv `w'_jshrs   `w'_jbnssec8_dv ///
         using "$Stata11_se/`w'_indresp_protect", clear

gen wave = strpos("abcdefg","`w'")
    renpfix `w'_

save $Mergeddata_master2016/`w'wave, replace
    }

use  $Mergeddata_master2016/awave, replace
foreach w in b c d e f g{

append using $Mergeddata_master2016/`w'wave.dta
    }

save $Mergeddata_master2016/abcdefg_long, replace

merge m:1 pidp using $Mergeddata_master2016/xwavedat, force generate(_wemerge_2)

Result                           # of obs.
    -----------------------------------------
    not matched                        60,792
        from master                         0  (_wemerge_2==1)
        from using                     60,792  (_wemerge_2==2)

matched                           334,897  (_wemerge_2==3)
    -----------------------------------------
```

History

#1 - 02/12/2018 10:12 AM - Alita Nandi

- Assignee changed from Alita Nandi to Nico Ochmann

- % Done changed from 0 to 90

- Private changed from Yes to No

In this latest release xwavedat includes BHPS cases as well. If you look at the hhorig for these cases you will see that they are exclusively BHPS samples. If this is not the case let me know.
Best wishes,
Alita

#2 - 02/12/2018 10:12 AM - Alita Nandi

- Status changed from New to Feedback

#3 - 03/14/2018 10:24 AM - Nico Ochmann

Dear Alita,

I have notified you a few weeks ago, but my reply might have gone missing. At any rate, No this is not the case. For some reason, other cases from hhorig are included.

Best wishes.

Nico

#4 - 03/14/2018 10:44 AM - Alita Nandi

Hi Nico,

Could you please provide the frequency distribution of HHORIG for _m==2 cases?

Thanks,
Alita

#5 - 03/14/2018 11:58 AM - Nico Ochmann

Hi Alita,

thank you very much for working on this issue now.
Here we go:

tab hhorig if _wemerge_2==2

Sample origin, household Freq. Percent Cum.

ukhls gb 2009-10 18,519 30.46 30.46
ukhls ni 2009-10 1,165 1.92 32.38
bhps gb 1991 18,063 29.71 62.09
bhps sco 1999 3,393 5.58 67.67
bhps wal 1999 3,465 5.70 73.37
bhps ni 2001 3,943 6.49 79.86
ukhls emboost 2009-10 5,641 9.28 89.14
ukhls iemb 2015 3,751 6.17 95.31
ECHP - SCPR 1,308 2.15 97.46
ECHP - ONS 1,179 1.94 99.40
ECHP - NI 365 0.60 100.00

Total 60,792 100.00

#6 - 03/14/2018 12:19 PM - Alita Nandi

Hi Nico,

Your code is fine and the data is fine.

As you know XWAVEDAT includes everyone who has ever been enumerated in the study. With the latest release this includes BHPS sample members who are not part of UKHLS as well (these can be identified by xwdat_dv = 2 in XWAVEDAT). So, these wmerge =2 cases are (i) enumerated children and non-responding adults in Understanding Society (ii) BHPS sample members who were never part of Understanding Society).

tab _merge xwdat_dv

```
| Study enumerated in: UKHLS, BHPS
|             or both
               _merge | in UKHLS   in BHPS x    in both |     Total
----------------------+---------------------------------+----------
       using only (2) |    31,051     26,906      2,835 |    60,792
          matched (3) |   275,156          0     59,741 |   334,897
----------------------+---------------------------------+----------
                Total |   306,207     26,906     62,576 |   395,689
```

Best wishes,
Alita

#7 - 03/14/2018 02:44 PM - Nico Ochmann

Alita,

you are fantastic, thanks for clearing that up.

Have a great day.

Nico

#8 - 08/14/2018 04:07 PM - Stephanie Auty

- Status changed from Feedback to Resolved

- % Done changed from 90 to 100

#9 - 08/22/2023 01:01 PM - Understanding Society User Support Team

- Category changed from Data analysis to Data management