

Understanding Society User Support - Support #1920

Create a Schooling years variable

06/10/2023 02:39 PM - Pauline Bucher

Status:	Resolved	Start date:	06/10/2023
Priority:	Normal	% Done:	100%
Assignee:	Understanding Society User Support Team		
Category:	Data management		
Description			
Hi team,			
I need help with some Stata coding...I want to create a schooling years variable to get the number of years of education of an individual. I am only using wave 11 for my analysis. I have realised that both scend and feend only ask new entrants which I guess explain why there are so many -8 responses? From there I have 2 questions.			
1)how would you combine feend and scend to not have this divide and jsut know the "education" leaving age of an individual?			
2)More importantly- what would be the most efficient way to get the scend and feend values for those in Wave 11 that were not new entrants? Should I load the whole ukhls, drop all missing/-8 ? but from there how would I append/attach those values to my wave 11 dataset (how to match the pidp here?)			
Thanks a lot for your help in advance!			

History

#1 - 06/10/2023 02:50 PM - Pauline Bucher

Note: I am confused about a 3) thing

3)When recoding the -8 and -1 as misisng in wave 11 for both scend and feend, I only get 1081 observations.

How is that consistent? The way I understand, it means that there were only 1081 new entrants in wave 11 which does not sound right? Am I missing something?

I could work with a restricted sample of new entrants wave 11 but I doubt there were only 1081 and don't know why I have so few?

#2 - 06/10/2023 03:51 PM - Pauline Bucher

Here is the code I have tried for query 2) using the xwavedat file but I still get only 1081 observations...

```
use xwavedat, clear
keep pidp feend scend
rename feend felage
rename scend slage
sort pidp
recode slage (-1=.) (-8=.)
recode felage (-1=.) (-8=.)
save educ, replace
```

```
use k_indresp, clear
```

```
use pidp k_sex k_dvage k_scend k_feend k_country k_health k_disdif1 k_disdif2 k_disdif3 k_disdif4 k_disdif5 k_disdif6 k_disdif7 k_disdif8 k_disdif9
k_disdif10 k_disdif11 k_disdif12 k_disdif96 k_fimnlabgrs_dv k_dissev1 k_dissev2 k_dissev3 k_dissev4 k_dissev5 k_dissev6 k_dissev7 k_dissev8
k_dissev9 k_dissev10 k_dissev11 k_dissev12 using k_indresp, clear
```

```
rename k_sex sex
rename k_dvage age
rename k_scend slage
rename k_feend felage
rename k_health disabled
rename k_disdif1 mobility
rename k_disdif2 lifting
rename k_disdif3 dexterity
rename k_disdif4 continence
rename k_disdif5 hearing
rename k_disdif6 sight
rename k_disdif7 communication
```

```
rename k_disdif8 learning
rename k_disdif9 danger
rename k_disdif10 coordination
rename k_disdif11 personalcare
rename k_disdif12 other
rename k_disdif96 none
rename k_country nation
rename k_dissev1 mobilitysev
rename k_dissev2 liftingsev
rename k_dissev3 dexteritysev
rename k_dissev4 continencesev
rename k_dissev5 hearingsev
rename k_dissev6 sightsev
rename k_dissev7 communicationsev
rename k_dissev8 learningsev
rename k_dissev9 dangersev
rename k_dissev10 coordinationsev
rename k_dissev11 personalcaresev
rename k_dissev12 othersev
sort pidp
```

```
recode slage (-1=.) (-8=.)
recode felage (-1=.) (-8=.)
```

save data, replace

merge 1:1 pidp using educ, keepusing(felage slage)

drop if age==.

#3 - 06/13/2023 05:13 PM - Understanding Society User Support Team

- Category changed from *Derived variables* to *Data management*

- Status changed from *New* to *Feedback*

- % Done changed from 0 to 70

- Private changed from *Yes* to *No*

Hi,

- 1) I am afraid I can't help you with this one, that depends on your research problem and we leave such decisions to researchers.
- 2) Please link that information from the xwavedat datafile using pidp.

Hope it helps.

Best wishes,
Piotr
UKHLS User Support

#4 - 06/13/2023 07:11 PM - Understanding Society User Support Team

Sorry, I missed your further posts.

Re 2, the advice is the same, the most efficient way is to link that information from xwavedat, you don't need to use k_scend and k_feend at all, scend and feend from xwavedat includes the same information (for wave 11, that is), save for a few mismatches due additional cleaning performed in xwavedat. The -8 on k_scend and k_feend are people who were outside of the universe for these questions in wave 11, you can find the universe in the pdf of the questionnaires (<https://www.understandingsociety.ac.uk/documentation/mainstage/questionnaires>). Btw, please note that the number of respondents who were asked k_feend, k_scend or both is smaller than 1081 (count if k_feend >0 | k_scend >0 -> 702 cases), that's because there are some cases for which k_feend >0 & k_scend >0 is true.

I would also recommend recoding all possible missing values at once, that is, to recode the whole range from -9 to -1, e.g. mvdecode k_scend k_feend, mv(-9/-1). In this case this wouldn't make any difference, but it can be useful for other variables, you can also recode all variables in the dataset at once: mvdecode _all, mv(-9/-1)

Best wishes,
Piotr,
UKHLS User Support

#5 - 11/30/2023 01:09 PM - Understanding Society User Support Team

- Status changed from *Feedback* to *Resolved*

- % Done changed from 70 to 100