Understanding Society User Support - Support #2157

Merging x_indresp and x_child, then merging across waves

10/02/2024 10:53 AM - Rehana Meeajan

Status:	Feedback		**Start date:**	10/02/2024
Priority:	Normal		**% Done:**	50%
Assignee:	Understanding Society User Support Team			
Category:	Data documentation			

Description

Hi,
(Using R)
I'm trying to merge variables from j_child and j_indresp, which I'm able to do using j_hidp - creating j_merge.
j_child: breastfeeding data variables
j_indresp: loneliness, famsup

I am having difficulty merging cross wave data from here.

I have tried renaming j_hidp to hidp,
if I then wish to merge across waves (j_merge with k_merge, l_merge etc), on inspecting the data I have duplicate pidp entries with the same breastfeeding data, with the same hidp.
Presumably different adults providing the same data about one infant's breastfeeding data and household support.

Merging by pidp doesn't work.

I am wondering if I instead need to merge across waves by j_pno (in j_indresp) and j_rapno (person providing breastfeeding data in j_child)
Would this be acceptable?
Would this be possible by renaming j_pno to "pno", and j_rapno to "pno" and merging across waves in this manner by full join?

Many thanks in advance for any thoughts or suggestions,
Rehana

History

#1 - 10/03/2024 06:24 AM - Understanding Society User Support Team

- Status changed from New to In Progress

- % Done changed from 0 to 30

Hello,

The individuals in the files j_child are hh members 0-15 year olds. Those in j_indresp are adult respondents who are 16+ year olds. So, if you try to merge these two files there will be no matched cases. Are you trying to merge children's information with that of their parents?

Best wishes,
Understanding Society User Support Team

#2 - 10/03/2024 08:44 AM - Rehana Meeajan

Understanding Society User Support Team wrote in #note-1:

> Hello,
>
> The individuals in the files j_child are hh members 0-15 year olds. Those in j_indresp are adult respondents who are 16+ year olds. So, if you try to merge these two files there will be no matched cases. Are you trying to merge children's information with that of their parents?
>
> Best wishes,
> Understanding Society User Support Team

Hi, thanks for the reply. Yes that's right, I'm hoping to merge j_indresp responses (about family support) with the parent's responses to breastfeeding questions in j_child

Would it be correct to merge j_indresp and j_child by j_hidp,

then keep only the data where j_pno (from j_indresp) matches j_rapno8 (person responding to breastfeeding questions in j_child).
My expectation is that this would only keep j_indresp responses from the breastfeeding parent (assuming that it was the breastfeeding parent who provided responses to the breastfeeding questions).

I have tried this, and it looks like I get similar breastfeeding data to looking solely at j_child, but with a few missing cases (suspect this will be because the person replying about breastfeeding questions didn't have matching responses to indresp)

#3 - 10/11/2024 10:38 AM - Understanding Society User Support Team

- Category set to Data documentation

- Status changed from In Progress to Feedback

- Assignee changed from Rehana Meeajan to Understanding Society User Support Team

- % Done changed from 30 to 50

- Private changed from Yes to No

Hello Rehana,

You are correct merging by j_hidp and j_rapno8 is the right approach. In the j_child file, the variable j_pno already exists, so you should rename it to something else, like pno_bf. Alternatively, you could rename the j_pno variable in indresp to j_rapno8.

You can also use the Pregnancy and Early Childhood (PEACH) data file (xwavepeach), which is a single cross-wave file that compiles key data reported by parents for all children under 10, along with information on pregnancy and parenting styles, from all waves of the main survey. The information is provided at the child level using the child's identifier, ensuring each row is uniquely identifiable. Additionally, the pidp identifier of the parent or caregiver who provided the information is included to facilitate linking to family circumstances.

If you only want data from Wave 10, you should filter by wave_newborn = 10 and merge pidp_newbornmother with pidp from indresp. You just need to rename pidp in indresp to match the PEACH identifier (pidp_newbornmother).

You can find more information about the PEACH data here.
https://www.understandingsociety.ac.uk/documentation/peach/

I hope this information is helpful.

Best wishes,
Roberto Cavazos
Understanding Society User Support Team