Understanding Society User Support - Support #960

pid and pidp not uniquely identifying rows

04/16/2018 03:10 PM - Gabriele Dente

Status:	Resolved		**Start date:**	04/16/2018
Priority:	Normal		**% Done:**	100%
Assignee:	Gabriele Dente			
Category:				

Description

Greetings,

I'm trying to use the BHPS armonised files, and wanted to merge files in the same waves (I need to merge indresp, income and jobhist). The problem is that the pid and pidp variables, that are suppose to uniquely identify each individual, do not actually do it. What I mean is, each individual, identified through the pid or pidp variables, appears in multiple rows, so when I try to merge the files I get the error saying that "variable pidp(or pid) does not uniquely identify observations in the using data".
Any help on the matter would be really appreciated!

History

#1 - 04/16/2018 05:06 PM - Stephanie Auty

- Status changed from New to In Progress

- Assignee set to Stephanie Auty

- Target version set to X M

- % Done changed from 0 to 10

- Private changed from Yes to No

Many thanks for your enquiry. The Understanding Society team is looking into it and we will get back to you as soon as we can.

Best wishes,
Stephanie Auty - Understanding Society User Support Officer

#2 - 04/18/2018 09:13 AM - Alita Nandi

- Assignee changed from Stephanie Auty to Gabriele Dente

- % Done changed from 10 to 90

Hello Gabriele,

The row in indresp is uniquely identified by pid/pidp but the rows in files income & jobhist are uniquely identified by pidp + other variables.
For bw_jobhist_bh it is pidp/pid PLUS bw_jspno_bh
For bw_income_bh it is pidp/pid PLUS bw_fiseq_bh

Hope this helps.

Best wishes,
Alita

#3 - 04/18/2018 03:55 PM - Gabriele Dente

That helps a lot.
Thank you so much for your help!

Best regards,
Gabriele

#4 - 08/14/2018 04:56 PM - Stephanie Auty

- Status changed from In Progress to Resolved

- % Done changed from 90 to 100