Understanding Society User Support - Support #1119

fiseq_bh and pidp are not unique in the income dataset

12/16/2018 04:15 PM - Yanan Zhang

Status:	Resolved	Start date:	12/16/2018
Priority:	Normal	% Done:	100%
Assignee:	Yanan Zhang		
Category:	Data documentation		

Description

Greetings!

I am merging all the income datasets with the 18-wave BHPS and the 8-wave UKLHS. However, the row in income is not uniquely identified by fiseq_bh and pidp in wave B1, B3, B9, B10, B11, B15, B16, B17 and B18 (especially for the B17 and B18). Then I have tried to use fiseq_bh, pidp and ficode to identify rows. Still, there are about 360 rows are not uniquely identified by the three variables. Those repetitions are from the B17 and B18.

Could you please provide any suggestions on it? Do you have any documentations to explain how the rows are identified for each dataset?

Best wishes,

Yanan

History

#1 - 12/17/2018 10:34 AM - Stephanie Auty

- Category set to Data documentation

- Status changed from New to In Progress

- Assignee set to Stephanie Auty

- Target version set to BHPS

- % Done changed from 0 to 10

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 - 12/17/2018 10:34 AM - Stephanie Auty

- Private changed from Yes to No

#3 - 12/21/2018 10:34 PM - Stephanie Auty

ISER is now closed for the holidays and will reopen on 2nd January. We will get back to you soon afterwards.

With best wishes for the winter break.

#4 - 01/08/2019 01:38 PM - Stephanie Auty

- % Done changed from 10 to 20

Dear Yanan,

Thank you for bringing this issue to our attention. Some of our team members are looking into it and we will continue to update you with progress. This investigation may take some time due to the time passed since the data was collected.

Best wishes,
Stephanie

#5 - 01/09/2019 05:34 PM - Stephanie Auty

- Status changed from In Progress to Feedback

- Assignee changed from Stephanie Auty to Yanan Zhang

- % Done changed from 20 to 80

Dear Yanan,

We are working on documentation for the unique identifiers for each dataset.

In this case, pidp and fiseq_bh do not always uniquiely identify the row. We think that this may be because sometimes the script reset fiseq_bh for each type of income receipt. Ficode is a harmonised version of ficode_bh, which has some categories from ficode_bh combined for use with UKHLS data. If you use pidp, fiseq_bh and ficode_bh then these variables together will uniquiely identify the rows.

Best wishes,
Stephanie

#6 - 08/10/2022 11:28 AM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- % Done changed from 80 to 100