Subject: Merging datasets in SPSS
Posted by rufus.benaud11 on Fri, 17 Mar 2017 01:45:43 GMT
View Forum Message <> Reply to Message

Hi, I've been trying to understand and work with DHS dataset on SPSS. My research question is to explore predictors of breastfeeding initiation, exclusive breastfeeding and bottle feeding in Ethiopia.

I have some queries around merging datasets in terms of bringing one variable from one dataset to another. For example, I want to bring a variable "Antenatal care" from household dataset to children's dataset. Given the IDs are different in both datasets, I am quite confused how to work this out. Can you please advise on this.

Thanks.

Subject: Re: Merging datasets in SPSS
Posted by Bridgette-DHS on Mon, 20 Mar 2017 21:16:40 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Research Associate, Cameron Taylor:

Thanks for the question! First you should check out our Merging Datasets page on the DHS Program website. http://preview.dhsprogram.com/data/Merging-Datasets.cfm
This page overviews the unique case identifiers in each data file, matching variables, and steps for merging datasets. However, before embarking on merging two datasets I would first make sure that a merge is necessary. The KR (kids file) already includes antenatal care information from the woman's file. Additionally, I am not sure why antenatal care would be in the household dataset (HR file) or the peoples recode file (PR) but if you are sure that you need to merge two datasets here is some guidance.

For example if you were going to merge the Peoples Recode (PR) and Kids Recode (KR) files in SPSS here is some guidance:

Using syntax
1) Open the PR file
2) Rename unique identifiers (cluster, household, and line numbers) in the using file to match the master file
• rename variables (hv001=v001).
• rename variables (hv002=v002).
• rename variables (hvidx=b16).
3) Sort PR file on these unique identifiers
• sort cases by v001(a) v002(a) b16(a).
4) Save PR file under a temporary name
5) Open KR file
6) Sort cluster, household, and line numbers in KR file
• Sort cases by v001(a) v002(a) b16(a).

Then using drop downs

Move v001 v002 b16 from "excluded" to "key variables" box
Check "Match cases on key variables"
Check "Indicate case source as variable"

Paste to syntax file! Highlight and run

You will then see the variable source01. This variable created during merging has a value of 0 for cases from the active dataset and a value of 1 for cases from the external data file. In our merge example KR is the active dataset and PR is the external dataset.

Question: So which children could be in source01=0?
Answer: A child who is in the KR file but not the PR file does not live in the household with their mother perhaps the child has died (check b5). Or they are older than 59 months

Question: Which children could be in source01=1?
Answer: a child who is in the PR file but not the KR file means that their mother wasn't interviewed.

As always please carefully review the questionnaires in the back of the report to fully understand who is being asked which questions. This will help you better understand who is in which data file and whether your merge is necessary

Let us know if you have additional questions!

Subject: Re: Merging datasets in SPSS
Posted by Hassen on Sat, 28 Apr 2018 13:45:26 GMT
View Forum Message <> Reply to Message

Dear Bridgette, Thank you for your Attractive Response and stepwise reply to our Questions!! I have done my Master of Public Health Thesis on factors associated with nutritional status of under five children in Ethiopia using 2016 EDHS Data set.I want to get your attractive reply for the following Questions.
1. Can all children in the KR file alive and living with their mothers? What is the difference between children in KR file and children in PR file? It is not clear for me.
2.What is the difference between age of child in variable B19 and Hw1? Which could be inline with my title?
3.Which variables are useful to create Household food security status from EDHS 2016 Data set.
Thank you in advance!!

Subject: Re: Merging datasets in SPSS
Posted by Bridgette-DHS on Tue, 01 May 2018 11:34:50 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

As has been stated in other postings, the PR file includes all children living in the household. It gives the most representative information about children in the household population. The KR file is limited to children whose mother is living and in the household. It does not include children whose mother has died or is living in a different household than the child. The children who are in the PR file but not the KR file tend to be relatively deprived and are more likely to be stunted, etc. However, if you want to relate stunting (etc.) to other characteristics of the child, or characteristics of the mother, it is necessary to use the KR file.

The variable hw1 (in the KR file) or hc1 (in the PR file) is age in months, calculated as month of interview minus month of birth (v008-b3). The new variable b19 is similar but is calculated down to the level of the day. The day of the interview is given as v008a (in century day codes, or cdc). The day of birth is the new b18. The length of a month is approximated as a constant number of days, 365.25/12=30.4375. Then b19 is the integer part of v008a-b18 when divided by 30.4375. Previously, children would be included in the KR file if hw1 was between 0 and 59, inclusive. Currently, the requirement is that b19 is between 0 and 59, inclusive.

If you have a question about an indicator such as food security, please refer to a specific table in which the indicator appears.

Subject: Re: Merging datasets in SPSS
Posted by Hassen on Wed, 02 May 2018 07:21:19 GMT
View Forum Message <> Reply to Message

Thank you very much for your Delicious Response!! I will come up with other challenges After I have seen all the Issues.
Respectfully,Hassen
