

---

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!