
Subject: Caseid variable is not available in the height-weight data for BDHS 2004
Posted by [m.hasan1](#) on Wed, 05 Jun 2013 11:37:56 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello,

I would like to have some advice on the following:

The 2004 BDHS height-weight dataset doesn't have the key variable 'caseid' which is used to merge height-weight data with children's data. I fail to find a key variable that can be used to merge the two datasets (BDKR 2004 dataset & BDHW 2004 dataset). Can anyone please suggest how can I create the caseid variable or how can I merge the two datasets? Please note that I am using STATA.

Thanks in advance for help.

T

Subject: Re: Caseid variable is not available in the height-weight data for BDHS 2004

Posted by [Liz-DHS](#) on Tue, 25 Jun 2013 20:34:48 GMT

[View Forum Message](#) <> [Reply to Message](#)

Here is an answer from one of our experts, Alexander Izmukhambetov:

There is an ID variable "HWHHID" in the BDWH 2004 height-weight data set. This variable contains household identification for every case (cluster + household number). There is also another ID variable "HWLINE", which identifies the household line number, relating every height-weight case to a household member.

Subject: Re: Caseid variable is not available in the height-weight data for BDHS 2004

Posted by [vega25](#) on Mon, 06 Oct 2014 20:48:23 GMT

[View Forum Message](#) <> [Reply to Message](#)

In order to merge the height and weight file with the children's file for Bangladesh 2004, the merge.doc file which accompanies the height and weight file mentions:

"Use HWCASEID and HWLINE, from the Height and Weight file, with CASEID and MIDX, from the Children's recode file to merge it with the Children's data."

HWCASEID, I presume, is meant to refer to "hwhhid" which is in the height and weight file.

However, hwhhid in the height and weight file and caseid in the children's recode file do not match! What should one do?

Subject: Re: Caseid variable is not available in the height-weight data for BDHS 2004

Posted by [Trevor-DHS](#) on Mon, 06 Oct 2014 21:44:58 GMT

[View Forum Message](#) <> [Reply to Message](#)

Try the below:

* HHID is just the first 12 characters of CASEID, so you can use the following to construct HHID in the KR file:

```
gen hwhhid = substr(caseid,1,12)
```

* The line number variable to use is then b16 (line number from the household schedule).

* Note that b16 is 0 if the child does not live in the household, but was a visitor.

```
gen hwline = b16
```

* You can then merge using the following command. This uses a many to 1 merge because of the cases above that can have 0 on b16.

```
merge m:1 hwhhid hwline using "bdhw4jfl.dta"
```

* Drop the extra cases from the 'using' file that we don't want.

```
drop if _merge==2
```

Subject: Re: Caseid variable is not available in the height-weight data for BDHS 2004

Posted by [Wolf](#) on Wed, 29 Jun 2022 12:30:29 GMT

[View Forum Message](#) <> [Reply to Message](#)

Thanks for your help above.

Please could you update this in the merge instructions? To reflect that it is not the 'midx' variable in the child recode that is equivalent to 'HWLINE' in the new file, it is instead 'b16'.

Just to confirm, in this case I will need to drop the last digit of the 'caseID' in the child recode, and instead replace it with 'b16' so that the codes will match in both files and can then be merged together?

Subject: Re: Caseid variable is not available in the height-weight data for BDHS 2004

Posted by [Trevor-DHS](#) on Wed, 29 Jun 2022 13:50:57 GMT

[View Forum Message](#) <> [Reply to Message](#)

You are really creating the household ID which is the first 12 characters of the caseid (effectively dropping the woman's line number - the last 3 characters of caseid [not just the last digit]). You are matching on the household ID and the household line number in the two datasets so that the two datasets can be merged.
