
Subject: Re: Merging Children with Household
Posted by [user-rhs](#) on Sun, 27 Mar 2016 18:46:51 GMT
[View Forum Message](#) <> [Reply to Message](#)

Befulina,

I stand corrected. You need to manually check the 26 duplicates and determine whether it makes sense to include them in the analysis. I noticed that some of the children don't have a line number in the HH roster, which probably means that they were dead at the time of interview (remember the birth hx questionnaire asks the woman of ALL births regardless of current vital status). I also saw that some children had "not listed in household" as the value for b16, which means that they belong to the woman being interviewed, but they do not live in the household with her. An example would be if the child lived with grandparents or if the woman was divorced/separated from the child's father and the child lives with him instead of her.

Here's a way to manually inspect the duplicates:

*Create a variable that flags duplicates using v001, v002, v003, b16
bysort v001 v002 v003 b16: gen count=_N /*_N is one of two built-in count variables in Stata that indicates the highest count within a group, in this case (v001 v002 v003 b16) */

*See how many are duplicates
tab count

*Visually inspect the duplicate records
sort b5 v001 v002 v003 b16
browse v001 v002 v003 b16 b5 b6 b8 if count>1

After you determine why the children don't have line numbers or are not listed in the HH roster, you can make a decision whether to include or exclude them. Remember that you can merge HH characteristics to the children without having their HH roster line number by merging using the HH unique identifier (v001 v002). If you're interested in all children born to a woman in the past 5 years, then it would be a good idea to include the dead children as well. If a child does not live with the mother, my inclination is that it would not make sense to link the mother's HH characteristics to the child, but again, that is your call to make.

HTH,
rhs
