Subject: Child development module in Honduras 2011-2012
Posted by helgaurke on Tue, 20 Mar 2018 20:51:48 GMT

Hi,

I am using the early child development variables in the Honduras 2011-2012 survey for my work, but cannot seem to figure out what the true sample size is. From what I understand, these variables are collected on children 36-59 months. In the report, the N is 2710, but in the data the sample is 2918, with an additional 913 children as system missing on all items pertaining to the development index. Can anyone explain to me why this is, and whether I need to select on a specific variable?

Thanks in advance!

Helga Urke

Subject: Re: Child development module in Honduras 2011-2012
Posted by Bridgette-DHS on Fri, 06 Apr 2018 17:46:05 GMT

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

Sorry for the delay with this response. Questions like this should identify the specific table in which the problematic number (in this case 2710) appears. I scanned the report and found the number in table 17.1.

The relevant variables appear to be s562* through s566* in the KR file (HNKR61FL.dta). If a variable is coded with a blank or dot then the variable is not applicable, i.e. is NA. For most of the s56* variables, there are 3614 unweighted cases that are not NA. If I restrict to hw1>=36 & hw1<=59, then 2923 cases are not NA. (The questions were asked about a large number of children, about 700, whose age turned out to be below 36 months or above 59 month.) Of the 2923, 5 cases consistently have code 9, which is not a legal code. If you just removed those cases you would get 2923-5=2918. I am guessing that that is how you got 2918. However, this is an unweighted frequency and the n's in table 17.1 are weighted. If I do

tab s564a if hw1>=36 & hw1<=59 [iweight=v005/1000000]

then the weighted number of cases that are not NA on the s56* variables and are in the range 36-48 months is 2778.2934, or 2778. The way I would read the table, that's the number that should be 2710. My best guess, 2778, is too high by 68 cases.l

This was a special module and table 17.1 is a special table, i.e. non-standard. If I were you I would simply take the cases that are not NA on s564a and are in the age range 36-59 months as the eligible cases and proceed to analyze them. I cannot match the table.