
Subject: Education variable inconsistent

Posted by [Kerstin](#) on Tue, 13 Aug 2019 09:07:20 GMT

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

I am using the DHS data of African countries. By cross-checking the educational variable (hv108 education completed in single years), I found inconsistencies with respect to the age of the respondent (hv105). For instance in the Angola PR dataset 2015-16, there are quite some observations where it is reported that the individual went longer to school than (s)he is old. Is this a calculation error of hv108 or a misreport? How should I deal with it?

Best regards

Kerstin

Subject: Re: Education variable inconsistent

Posted by [Bridgette-DHS](#) on Sat, 17 Aug 2019 18:19:30 GMT

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

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

I see that there are 372 cases in which hv108 is coded "98", for "don't know". These are cases for which $hv108 > hv105$, numerically, but they are not errors. However, I see 16 cases with $hv108 > hv105$ and hv108 is in the range 5 to 13 and hv105 is in the range of 3 to 10. These are definitely coding errors. They should have been discovered during editing and given a value of $hv108=97$, "inconsistent".

Of course, years of schooling should not just be greater than hv105, it should be greater than $hv105+5$ (approximately). In your analysis, I suggest recoding hv108 to 97 if $hv108 > hv105+5$.
