
Subject: Duplicate variables?

Posted by [WoodyWoodruff](#) on Mon, 16 Sep 2013 20:05:03 GMT

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

Dear All:

I am new to DHS data analysis, so please excuse what may be an overly simple question. I am using the child data from the three surveys in Ethiopia. So far, I have looked extensively only at the DHS 2010 dataset in which there seems to be duplicate variables, most of which contain the same data. However, there is one pair of variable (V415 and M38) which both seem to contain the responses to question 460 on the woman's questionnaire (Did (NAME) drink anything from a bottle with a nipple yesterday or last night?). Although the Recode Manual for DHS6 contains the same description of these two variables, the data in them are different, as demonstrated by a cross-tabulation (see attached file). Does anyone know what is the difference between the data in these two variables?

Thanks in advance for any help.

File Attachments

1) [Cross-tab V415 and M38.jpg](#), downloaded 1180 times

Subject: Re: Duplicate variables?

Posted by [Bridgette-DHS](#) on Fri, 11 Oct 2013 12:39:22 GMT

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

Here is a response from one of our Senior Data Processing Specialists, Lady Ortiz, that should answer your questions.

Question M38 "drank from a bottle with nipple" was asked for all living children born in the last 5 years preceding the survey. Question V415 has the same information as M38 but contains information only for the LAST child born in the last 5 years before the survey.

Subject: Re: Duplicate variables?

Posted by [WoodyWoodruff](#) on Sat, 12 Oct 2013 01:35:58 GMT

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

Dear Ms. Ortiz:

Thank you for your prompt answer. You know what they say: Don't demonstrate your competence or people will put you to work. So here's a few more questions I have about DHS datasets:

1) Two variables (V024 and V101) seem to contain the region of residence. What is the difference between them?

2) Two variables record time after birth breastfeeding was initiated: V426 and M34. This question asked ONLY about most recently born child. In the DHS 2011 Ethiopia child dataset, V426 has only 250 (3.0%) missing values. M34 has 4,114 (35.3%) missing values, which seems more likely given that children born less recently should have missing values. What is the difference between these two variables?

Thank you in advance for your help.

Subject: Re: Duplicate variables?

Posted by [Bridgette-DHS](#) on Tue, 29 Oct 2013 16:25:07 GMT

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

Here is another answer from Ms. Ladys Ortiz:

Dear Woody

Variables V024 and V101 has the same information but different denominators. Unfortunately you can't see the difference because when we export our data to SPSS/Stat/SAS, we only include completed cases. V024 has the region information for all the individuals regardless of the result of the interview; V101 has the region information for individuals with completed interviews only.

In reference to your second question, you should follow the same approach in my previous answer, related to questions M38/V415.

Thanks

Ladys
