

---

Subject: Dataset issues

Posted by [cgranberg](#) on Mon, 09 Feb 2015 09:44:33 GMT

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

---

Hi!

I am a Master student in Sweden, Uppsala and I am doing my master thesis right now in International Health and I am doing secondary analysis on DHS data. I have some questions about this because it is my first time working with DHS. I am having difficulties trying to figure out how to understand the variables that I have in my dataset. My dataset is from Kenya 2008-09 and I want to look at the dataset for the women. I have looked at the recode manual (DHS-V) and have managed to figure out some of the variables but not all of them. For example in the recode manual there are some variables about MTCT and it says "Ways of transmission from mother to child " and there are three variables for that; V774A: During pregnancy , V774B: During delivery , V774C: By breastfeeding. In the report I can only see that they ask about knowledge about MTCT and transmission through breastfeeding, so first of all, are these variables knowledge (and not actual transmission of HIV to her child?) and second, it seems like all three variables are in the dataset although not brought up clear in the report, is that correct? Same thing with this variable "V824: Drugs to avoid AIDS transmission to baby during pregnancy ". Is that knowledge or if the mother actually took drugs? In the report it seems like they ask about knowledge anyways.

In the map with the dataset for the women I have a word document with some codes, though to me this is not very clear. How to interpret this? What is "SECTIONS AND VARIABLES UNUSED"? and "SECTIONS AND VARIABLES ADDED"? How is this connected to the big general recode manual?

Finally I want to divide the women by HIV status, as I have understood it, it is possible to do that? I have the dataset with HIV test results, but does that dataset include both men and women who gave blood for HIV testing? If I only want the women, how to do this in the best way?

I would really appreciate some answers if anyone has any clue how to do this and how to best interpret the variables!

Thank you!

Caroline

---

---

Subject: Re: Dataset issues

Posted by [Trevor-DHS](#) on Sat, 14 Feb 2015 22:00:29 GMT

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

---

1) v774a, v774b, and v774c are created from the three parts of question 911 (page 389 of the KDHS 2008-09 final report <http://www.dhsprogram.com/pubs/pdf/FR229/FR229.pdf>), and these three variables are about the respondents knowledge of mother to child transmission. v824 is created from question 912A ("Are there any special drugs that a doctor or a nurse can give to a woman infected with the AIDS virus to reduce the risk of transmission to the baby?"), and is also a knowledge-related question.

2) All datasets from different surveys are recoded to common formats, known as the recode files, described in the Recode Manual. The word document you are referring to contains information

related to the Kenya DHS 2008-09 survey concerning the sections and variables not used for Kenya, and the sections and variables added that were specific to the Kenya DHS 2008-09 survey and are not in the standard format.

3) The HIV dataset does include the HIV test results for both men and women. If you merge the HIV data to the Women's Recode (IR) data according to the cluster number, household number and line number, this will merge the data just for the women (as the IR file has no men's data). See this thread for information about merging the HIV data to the IR data.

---

---

Subject: Re: Dataset issues

Posted by [Hassen](#) on Sun, 06 May 2018 09:43:41 GMT

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

---

Thank you all!!

---