
Subject: duplicated almost blank variables
Posted by [fionakee](#) on Thu, 23 Jan 2020 17:26:18 GMT
[View Forum Message](#) <> [Reply to Message](#)

For eg. I downloaded Cambodia dta.file for DHIS 2014. I went to look for some antenatal care indicators

e.g. m70 "Baby postnatal check within 6 weeks".

1) While I am browsing the data for m70_1, I noticed there were a lot of m70_2, m70_3 etc... However, most of them are blank. So what's the relevancy of these blank variables? This happens to other indicators as well.

But then, I noticed that for H3_1 "tetanus vaccination", the H3_2, H3_3 ..some of them are not blank.

So how do I know which is exactly relevant?

I am only looking for one clean response, per indicator, per individual.

Also, to summarize regression code, can you confirm that the following code will work? What will be the code difference, if i would like to run multi regression?

```
syvset[pw=weight],psu(v021/v001)strata(v022/v023)
svy: reg H3_1 m70_1
```

Thanks!

Subject: Re: duplicated almost blank variables
Posted by [Bridgette-DHS](#) on Fri, 24 Jan 2020 15:40:13 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

You are probably using the IR file, which has women as units. Each woman has m70_1 through m70_6 for up to 6 children born in the past five years. If a woman had, for example, two children in the past few years, then only m70_1 and m70_2 will be filled in, and m70_3 through m70_6 will be blank. The blank just means "not applicable". For some of these subscripted variables, only the first one (for the most recent birth) will have any data in it, because the question was only asked about the most recent birth. I strongly recommend that you switch to the KR file, which includes almost all of the mother's data but has a separate record for each child. The variable m70_1, for example, becomes m70, with a new variable on the record called bidx; for the most birth, bidx=1. For the next most recent birth m70 will be the same as m70_2 and will have bidx=2.

Using the KR file, you could have, for example,

syvset [pw=weight], psu(v021) strata(v023) singleunit(centered)
svy: regress h3 m70

Subject: Re: duplicated almost blank variables
Posted by [fionakee](#) on Mon, 27 Jan 2020 08:32:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks DHS team.

Another question for your advice:

1) What does a KR file refers to? In all the 3 countries below, I dont see any KR file download. There were only Household, Individual, HIV.

2) I downloaded IR dataset for Cambodia, PNG, and Phillippines (STATA). However, within the zip file of each country, the variable code DOC file is not usually complete with each label definition. Some codes do not have a definition in the DOC. Is it possible for you to share the complete variable code DOC file for these 3 countries?

3) I could not find the original questionnaires for KHM, PNG, and PHL. I was not able to compare the variables with the model questionnaires on your website (I assume each country's questionnaire have been modified slightly). Could you possibly share the original questionnaires for these 3 countries?

4) For LOW BIRTHWEIGHT of the mother's child, will I be able to find it in KR or IR file? If yes, could you refer which variable code it's?

Subject: Re: duplicated almost blank variables
Posted by [Bridgette-DHS](#) on Fri, 31 Jan 2020 14:11:43 GMT
[View Forum Message](#) <> [Reply to Message](#)

1) What does a KR file refers to? In all the 3 countries below, I dont see any KR file download. There were only Household, Individual, HIV.

The KR file exists for the countries below - please check the respective dataset pages. The KR file is the Children's Data. This dataset has one record for every child of eligible women, born in the last five years. It contains the information related to the child's pregnancy and postnatal care and immunization and health. The data for the mother of each of these children is included. This file is used to look at child health indicators such as immunization coverage, vitamin A supplementation, and recent occurrences of diarrhea, fever, and cough for young children and treatment of childhood diseases. The unit of analysis (case) in this file is the children of women born in the last 5 years (0-59 months).

2) I downloaded IR dataset for Cambodia, PNG, and Philippines (STATA). However, within the

zip file of each country, the variable code DOC file is not usually complete with each label definition. Some codes do not have a definition in the DOC. Is it possible for you to share the complete variable code DOC file for these 3 countries?

The .doc file contains useful information on country specific variables and differences in the Standard Recode definition. You will also need the DHS Recode Manual. The Recode Manual contains a general description of the recode data file, including the rationale for recoding; a description of coding standards and recode variables, and a listing of the standard dictionary, with basic information relating to each variable.

3) I could not find the original questionnaires for KHM, PNG, and PHL. I was not able to compare the variables with the model questionnaires on your website (I assume each country's questionnaire have been modified slightly). Could you possibly share the original questionnaires for these 3 countries?

The questionnaires are in the appendix of the final reports. Please search for the final reports at: <http://www.dhsprogram.com/publications>

4) For LOW BIRTHWEIGHT of the mother's child, will I be able to find it in KR or IR file? If yes, could you refer which variable code it's?

The child's birth weight (M19) is in both the KR and IR files.
