
Subject: NGHR7BFL

Posted by [Mayo](#) on Mon, 27 Jun 2022 16:33:27 GMT

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

Hello,

I am using NGHR7BFL to do an analysis and in the data file, I noticed that there are multiple variables for "Highest educational level attained" (HV106\$01 onward). Which one should I use?

Subject: Re: NGHR7BFL

Posted by [Janet-DHS](#) on Tue, 28 Jun 2022 13:40:37 GMT

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

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

You are using the HR file, which has one very wide record per household. The subscripts 01, 02....refer to the line number of the person in the household. Life will be easier if you use the PR file, which has one record per person. In the PR file, the line number is given as hvidx and hv106 does not have a subscript.

Subject: Re: NGHR7BFL

Posted by [Mayo](#) on Tue, 28 Jun 2022 15:02:51 GMT

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

Thanks so much!

Subject: Re: NGHR7BFL

Posted by [Mayo](#) on Wed, 10 Aug 2022 20:51:21 GMT

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

Hi everyone,

how to properly weight this survey data in SPSS.

Thanks,

Subject: Re: NGHR7BFL

Posted by [Shireen-DHS](#) on Fri, 12 Aug 2022 12:30:35 GMT

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

Hello,

When using household data (HR file) or population data (PR file) you use the hv005 weight but you need to divide this by 1 million.

This YouTube video explains about the different weights in DHS data and how to apply them in SPSS: https://www.youtube.com/watch?v=NNg8HD_IKow&t=88s

I also wanted to share another resource. We have standardized code to produce all DHS indicators in SPSS which is organized into chapters based on topic. When you go to the chapter of interest you can see "tables" syntax files that tabulates these indicators while applying weights. If you use this code, please be sure to read the notes in the readme file and the main file. <https://github.com/DHSProgram/DHS-Indicators-SPSS>

Thank you.

Best,

Shireen Assaf
The DHS Program

Subject: Re: NGHR7BFL
Posted by [Mayo](#) on Tue, 16 Aug 2022 13:32:20 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks again.

Another question is: how do I know if there is data missing at random (MAR) or MCAR?

Subject: Re: NGHR7BFL
Posted by [Janet-DHS](#) on Thu, 01 Sep 2022 20:34:37 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

You are asking questions that go beyond DHS data and therefore beyond the scope of the forum.

DHS has very low levels of "missing" data. A blank or dot in a DHS data file should be interpreted as Not Applicable (NA). If you are thinking of "missing" as "don't know" or "refused" or something like that, we use special codes such as 8, 9, 9994, etc., depending on the variable. The frequencies of those codes are usually very low.

In general, to test whether "missing" is random with respect to some potential covariate, you construct a binary variable that is 1 if "missing" and 0 if "not missing" and do a logit regression of that variable on the covariate, to see whether there is a statistically significant relationship.

Subject: Re: NGHR7BFL

Posted by [Mayo](#) on Mon, 10 Oct 2022 18:59:48 GMT

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

When using the PR file, what is the difference between the HV270 and HV270a variables? Which is best to use for a binary logistic regression?

Subject: Re: NGHR7BFL

Posted by [Janet-DHS](#) on Wed, 12 Oct 2022 13:37:44 GMT

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

Following is a response from DHS staff member Tom Pullum:

Beginning with DHS-7, most surveys include hv270a in the PR file, v190a in the IR, KR, and BR files, and mv190a in the MR file. The "a" indicates that the wealth quintiles are residence-adjusted, i.e. calculated separately for urban and rural areas. A problem with the original, unadjusted wealth quintiles is that, in most surveys, there are very few households in the top quintile in rural areas and very few households in the bottom quintile in urban areas. If you use the unadjusted wealth quintiles in a model, much of the information is actually an urban/rural distinction. If you use the unadjusted wealth quintiles in a model, AND include urban/rural (hv025, etc.) then you have a better separation of wealth and residence, but the model may run into estimation issues because there are (typically) so few cases in the two combinations I mentioned.

Bottom line: if your model includes urban/rural, which it probably should, then you may want to use the adjusted wealth quintiles rather than the unadjusted. But there's no law saying you have to do that. It would be good to tell the reader which version you are using.

Subject: Re: NGHR7BFL

Posted by [Mayo](#) on Sun, 13 Nov 2022 15:43:00 GMT

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

Hi Janet,

When using the NGPR7BSV file, I would like to determine the unweighted sample size. Would I use HV002 for this?

Subject: Re: NGHR7BFL
Posted by [Janet-DHS](#) on Wed, 16 Nov 2022 14:10:38 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member Tom Pullum:

The best variable for this purpose is hv000 (in Stata, HV000 in SPSS). That variable only takes one value, the string "NG7". If you tab that variable, without weights, you get 188,010 cases in the PR file, i.e. individuals in the household survey. If you do the same thing in the HR file, which has households as units, you get 40,427 households.

Subject: Re: NGHR7BFL
Posted by [Mayo](#) on Tue, 07 Mar 2023 17:46:06 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Janet,

I'm curious to know how often data is collected for the NDHS? I was looking at the website and saw that the first one occurred in 1990. It seems that data is not collected annually. Am I right?

Subject: Re: NGHR7BFL
Posted by [Janet-DHS](#) on Tue, 07 Mar 2023 21:14:49 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member, Tom Pullum:

Surveys in Nigeria, as in most countries that participate in the DHS survey program, are done approximately once every five years. Peru and Senegal are the only countries that have annual surveys.

Subject: Re: NGHR7BFL
Posted by [Mayo](#) on Tue, 07 Mar 2023 22:07:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks!

Subject: Re: NGHR7BFL
Posted by [Mayo](#) on Mon, 13 Mar 2023 18:00:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Janet,

Regarding is wealth quintile calculated after data collection or beforehand?

Thanks in advance,

Subject: Re: NGHR7BFL

Posted by [Janet-DHS](#) on Wed, 15 Mar 2023 14:31:23 GMT

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

Following is a response from DHS staff member, Tom Pullum:

The wealth quintiles are constructed from the raw data files after the data have been collected. Using household level characteristics (or "assets"), a principal components analysis is done with households as units. The first principal component is interpreted as the continuous wealth index. The quintiles (hv270 in the PR file) are constructed from the continuous index (hv271). The DHS website provides more details.
