
Subject: Grouping of ethnic groups Nigeria 2003, 2008 and 2013

Posted by [aligazan](#) on Fri, 03 Feb 2017 19:25:04 GMT

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

Hello

I am trying to compare some trends over the years among ethnic groups, however I am having difficulty because of differences in the ethnic groups captured in each of these surveys.

Nigeria 2003 NGIR4BFL

S118 Ethnicity. This has 100 ethnic groups. Variable V131 is not used. In the published report ethnic groups are categorised into 6 main ethnic groups (Fulani, Hausa, Igbo, Kanuri, Tiv, Yoruba) and others. I would like to know how to group the ethnic groups in S118 into these 6 main ethnic groups.

Nigeria 2008 NGIR53FL

S118 and V131 are used for ethnicity, with S118 being a comprehensive list, and V131 grouping these into 10 main ethnic groups. Unfortunately this grouping cannot be applied to S118 from the 2003 file above as the list of ethnic groups in S118 from each of the years are different.

Nigeria 2013 NGIR6AFL

V131 is used for ethnicity with 270 ethnic groups. The report groups these into 10 main ethnic groups (Tiv, Kanuri, Igala, Fulani, Ibibio, Ijaw, Hausa, Ekoi, Igbo, Yoruba). I would like to know how V131 variables are grouped to get into these ethnic groups. Unfortunately as above I cannot use the group from 2008 as the ethnic groups listed do not compare.

So ideally I would like to have the same major ethnic groups for all three surveys, which I think will need some grouping code to show me how the smaller ethnic groups have been grouped each year into the major groups listed in the reports.

Please can you help me?

Thanks in advance,

Janet

Subject: Re: Grouping of ethnic groups Nigeria 2003, 2008 and 2013

Posted by [Liz-DHS](#) on Fri, 24 Feb 2017 15:41:05 GMT

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

Dear User, A response from Survey Manager (Nigeria 2008 DHS) Adrienne Cox:

Quote:

The most accurate method would be to use the ungrouped ethnicity so that the researcher can apply the same approach to all surveys, or at least for the 2003 NDHS. I am not sure if the ten

ethnic groups from 2008 and 2013 are the same. If they are, the easiest approach to standardize the ethnicity across all three surveys is to use the six main groups from 2003 and place combine Igala, Ibibio, Ijaw, and Ekoi into the "other" group. I think the latter is the best approach because the percentages for the other ethnic groups are so small.

I believe that NPopC developed a more extensive list of ethnicities captured in their other surveys at the time of the 2008 NDHS. I believe the same list was used for the 2013 NDHS.

A response from Data Processing expert for the Nigeria DHS 2013, Claudia Marchena:

Quote:

Variable V131 in the recode data file has all the individual ethnic categories as they were collected in the field. In the final report, Mianmian (who produced most chapters for NDHS 2013) have specific individual categories displayed in the ethnic variable without grouping (in green below), missing displayed as missing (in orange below), and the rest of the categories grouped as "Others (in red below)".

Best,
Claudia

Please let us know if you need additional assistance.

File Attachments

1) [code_for_forum_post_Nigeria_Ethnic_groups.jpg](#), downloaded 1405 times

Subject: Re: Grouping of ethnic groups Nigeria 2003, 2008 and 2013

Posted by [AnvitaDixit](#) on Sun, 11 Jun 2017 21:41:05 GMT

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

Hi! I tried to do this in SAS, but am not able to replicate the DHS report. Please check and advise. Thanks.

```
data new3;  
set new2;  
if v131=91 then v131w=91;  
if v131=109 then v131w=109;  
if v131=130 then v131w=130;  
if v131=133 then v131w=133;  
if v131=137 then v131w=137;  
if v131=140 then v131w=140;  
if v131=179 then v131w=179;  
if v131=276 then v131w=276;  
if v131=298 then v131w=298;  
if v131=998 then v131w=998;
```

```
else if v131=. then v131w=999;  
else v131w=0;  
run;  
proc freq data=new3;  
tables v131w;  
run;
```

Subject: Re: Grouping of ethnic groups Nigeria 2003, 2008 and 2013

Posted by [boyle014](#) on Mon, 12 Jun 2017 17:26:22 GMT

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

To avoid harmonization problems like this, you might try using IPUMS-DHS at www.idhsdata.org to create your dataset. At IPUMS-DHS, the data is already fully harmonized and you can get complete information on the details of each variable (such as how it was asked, and of whom). It's very easy and fast to use, and it includes the three Nigerian samples that interest you.

To create a dataset in IPUMS-DHS (www.idhsdata.org), first login at the top of the page using your DHS username (email) and password.

Then click GET DATA -> WOMEN (assuming you want the women data).

Click SELECT SAMPLES, choose the three Nigerian samples you want, and click SELECT SAMPLE SELECTIONS.

You can choose all the variables you want by browsing the Topics drop-down menu on the far left. Ethnicity is under the CORE DEMOGRAPHIC variables. To add it to your data file, click the plus sign next to it.

To learn more about a variable, such as how it was asked, click on the variable name. That will show you the variable categories, and you will also see a series of tabs across the top of the page, such as "Comparability" and "Survey Text." I can see that ethnicity in Nigeria is unusual because it is based on an open-ended question; fortunately it was asked the same way in all three of the years.

When you have added all the variables you want, click VIEW DATA CART in the upper right of the screen. (Despite the metaphor of a data cart, the data are completely free). You'll see a list of the variables you've selected. You also get some variables whether you ask for them or not, such as sample weights.

Click CREATE DATA EXTRACT. The third line on the next page will read CHANGE DATA FORMAT. You probably want to click "Change" and select "SAS." (I'm not sure what the default is; I've got the data format set to STATA for my account.) You can describe the dataset if you wish, and then click SUBMIT EXTRACT.

Refresh the next page a couple of times, and your dataset will appear under the column FORMATTED DATA. Click on it to download the zipped file. You'll have to unzip it to use it.

I just created a dataset of the three Nigerian surveys with the ethnicity variable (which happens to be called "ethnicityng" in IPUMS-DHS). I determined the percentage of respondents in each ethnic category for each year using the STATA command:

```
tabulate ethnicityng sample [aw = perweight], column
```

This creates a weighted crosstab of ethnicity by sample year, showing column percentages. "sample" is the country-year; "perweight" is the standard person weight variable. BTW, perweight is ready to apply when you download your IPUMS-DHS file. You don't have to divide it by anything, as you do if you're using The DHS Program datasets.

In my crosstab, the numbers match the final reports exactly. It took me less than 5 minutes to check this out.

So, IPUMS solves your problem if your issue is harmonization. It's also possible your problem is not weighting the data. Either way, you may want to check out IPUMS-DHS. It's seems perfect for your research project.

I hope this is helpful.

Liz Boyle
Professor, University of Minnesota
PI, IPUMS-DHS (project supported by NICHD)