
Subject: Bangladesh 2014

Posted by [loudermilke](#) on Fri, 13 Oct 2017 18:40:57 GMT

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

Hello,

I am interested in looking at oral rehydration therapy, wealth index, blood in stools, and having eaten or drank while having diarrhea in the past two weeks. I am using SAS. I am missing a lot of data. Below is my code. Can you help me determine where I am making a mistake and why I have over 7,000 observations missing? I started out with 7,886 and have not yet deleted any missing. Also, when I compare my statistics to the Bangladesh statistics from 2014, I do not get the same numbers.

Thank you in advance for your help!

```
libname bds 'E:\Quinn_GA\BD_2014\bdkr72sd';
data bds.new;
set bds.bdkr72fl;
run;

proc contents data=bds.new;
run;
/*
H38: Had diarrhea in last two weeks: amount offered to drink
H39: had diarrhea in last two weeks: amount offered to eat
H11B: Blood in the stools
V113: Source of drinking water
V190: Wealth index
H11: had diarrhea recently-Whether the child had diarrhea in the last 24 hours
or within the last two weeks
B8: Current age of child
H13: received oral rehydration therapy
*/

data bds.new1;
set bds.new;
keep h38 h39 h11b v113 v190 h11 b8 h13 S510JG;
run;
proc contents data=bds.new1;
run;
proc freq data=bds.new1;
tables h38 h39 h11b v113 v190 h11 b8 h13 S510JG;
run;

data bds.new2;
set bds.new1;
```

```

if b8 <=5 then child=1;
Else delete;

if h11 = 2 then diarrhea=1;
if h11 = 0 then diarrhea=0;

if h11b = 0 then blood=0;
if h11b = 1 then blood=1;

run;

proc freq data=bds.new2;
tables (child blood v190)*diarrhea h13 h13*diarrhea;
run;
/*Total of 370 children had diarrhea; 15.4% of them had blood in their stools*/
proc freq data=bds.new2;
tables diarrhea child child*diarrhea blood blood*diarrhea; run;
/*Sample: 7,541, 4.92% of children under the age of 5 had diarrhea in the last 2 weeks*/

proc contents data=bds.new2;
run;

proc freq data=bds.new2;
tables h38 h39 h11b v113 v190;
run;

```

Subject: Re: Bangladesh 2014
 Posted by [kingx025](#) on Sun, 22 Apr 2018 06:57:01 GMT
[View Forum Message](#) <> [Reply to Message](#)

I believe that you are ending up with a large number of missing observations because some of the variables you are interested in have only a small number of cases in the universe (i.e., they apply to only a small subset of the data). Cases that are excluded from a variable are coded as blank in the original DHS files and will appear as missing cases. For example, number of children with diarrhea in the past 2 weeks is only a small fraction of all cases of children under 5, and the question about blood in stools and eating and drinking during an episode of diarrhea excludes any children who didn't recently have diarrhea (so those cases appear as missing).

Bangladesh 2014 is now included in the IPUMS-DHS database, and you can check the variable-specific documentation to confirm which cases are in the universe for a given variable (and thus don't appear as blanks in the DHS files). For example:
 For H11 (recently had diarrheal disease), the universe is: Bangladesh 2014: Surviving children under age 5, born to ever-married women age 15-49 -- so you are dropping the dead children.
 For H11B (had blood in stools during recent diarrheal disease), the universe is: Bangladesh 2014:

Surviving children under age 5 with diarrhea in the past 2 weeks, born to ever-married women age 15-49-- so your sample is reduced to the small number of children who survived and had recent diarrheal disease.

You should also be sure that you are using weights when you compare your numbers to the published numbers in the DHS final reports.

Miriam King
