

---

Subject: Re: Sample weights and stratification - Nigeria 2008 and 2018

Posted by [Goethe2014](#) on Thu, 21 May 2020 15:39:00 GMT

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

---

Dear Tom,

Thank you for the updated code. I applied it to the datasets in the following way (Note: The state variable goes by sstate not shstate in Nigerian DHS data):

1. Open DHS 2008 data file and command:

```
egen stratum_ID_2008=group(sstate v025)
gen tempvar=stratum_ID_2008
```

2. Open DHS 2018 data file, command and save:

```
gen stratum_ID_2018=v023
gen tempvar=stratum_ID_2018
save MYSTORAGEPATH, clear
```

3. In DHS 2008 file command:

```
append using MYDHS2018STORAGEPATH
gen survey=.
replace survey=1 if v007==2008
replace survey=2 if v007==2018
egen stratum_ID=group(tempvar survey)
drop tempvar
egen cluster_ID=group(v001 survey)
```

This results in the following (see screenshots of data - example of DHS2008 section, example of DHS2018 section, last entries of DHS 2008 and first entries of DHS 2018).

I have to admit I can not judge whether my result is now correct and looks how it should look like. The survey variable was correctly coded for sure but maybe you could tell me whether the other values also look like they are supposed to

(e.g. stratum\_ID starting going up in odd numbers for 2008 data 1,3,5....147; stratum\_ID going up in even number for 2018 data 2,4,6....148; v001 going from 1-888 for DHS 2008 and 1-1400 for DHS2018; cluster\_ID ranging from 1-1763 and 2-2275 for DHS2018)

I would be very grateful if you could tell me whether my result looks how it should and I have done everything right or whether there is still an issue with the code applied.

Thank you in advance!

## File Attachments

---

- 1) [DHS2008DHS2018Border.JPG](#), downloaded 709 times
  - 2) [DHS2008Example.JPG](#), downloaded 698 times
  - 3) [DHS2018 Example.JPG](#), downloaded 680 times
-