

---

Subject: Appending files

Posted by [Ugonna](#) on Thu, 04 Mar 2021 22:09:04 GMT

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

---

Hello ,

Please I need help here,

I'm looking at trends using 2008, 2013, and 2018 NDHS, of which I have appended appended the files using directions gotten from this forum

For instance I applied these steps to the files before and after appending:

```
*In the 2008 survey
egen stratum_ID_2008=group(sstate v025)
gen tempvar=stratum_ID_2008
```

```
*In the 2008 survey
gen stratum_id_2013=v023
gen tempvar=stratum_id_2013
order tempvar, after(v025)
```

```
* In the 2018 survey
gen stratum_ID_2018=v023
gen tempvar=stratum_ID_2018
```

```
* append and construct "survey" using v007
```

```
* In the combined file
egen stratum_ID=group(tempvar survey)
drop tempvar
```

```
egen cluster_ID=group(v001 survey)
```

```
svyset cluster_ID [pweight=v005], strata(stratum_ID) singleunit(centered).
```

Please I need to understand if I handled 2013 file correctly, because I noticed it has 73 strata (Lagos state only has urban) unlike 2008 and 2018 with 74 strata. I observed that when I run these commands on the appended file

```
tab v023 v024 if year==0
tab v023 v024 if year==1
tab v023 v024 if year==2
(0=2008, 1=2013, 2=2018) 2008 and 2018 comes out properly ordered while 2013 output places states in wrong geopolitical zones.
```

I don't know what I could be doing wrong.  
Will appreciate some help here.

Thank you.

---

Subject: Re: Appending files

Posted by [Bridgette-DHS](#) on Sat, 06 Mar 2021 00:12:41 GMT

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

---

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

You are right--the 2013 survey has only 73 strata, and the two other surveys have 74. In the 2013 survey, Lagos does not have rural cases, whereas in the 2008 and 2018 surveys it DOES have rural cases. The stratification variable, hv023, consists of all combinations of shstate and hv025. (hv024 is the six major regions.) The labels for shstate are consistent across the surveys, but the numbering for hv023 is not consistent.

I don't know whether the absence of rural cases for Lagos in the 2013 survey was due to a temporary re-definition of Lagos as all urban or was due to something in the sample design. However, what you have done is fine and nothing in your construction of svyset needs to change..

---

Subject: Re: Appending files

Posted by [Ugonna](#) on Sun, 07 Mar 2021 21:28:30 GMT

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

---

Thank you for your response.

Please what does the following weighting commands on an appended file of 3 datasets do, and is it necessary?

```
summarize wt
scalar N=r(N)
scalar Nshare=N/3
generate wt_new=wt
```

```
summarize wt if year==0
scalar sumwt_new =r(sum)
replace wt_new =wt *Nshare/sumwt_new if year==0
```

```
summarize wt if year==1
scalar sumwt_new =r(sum)
```

replace wt\_new =wt \*Nshare/sumwt\_new if year==1

summarize wt if year==2  
scalar sumwt\_new =r(sum)  
replace wt\_new =wt \*Nshare/sumwt\_new if year==2

Thank you

---

---

Subject: Re: Appending files  
Posted by [Bridgette-DHS](#) on Wed, 10 Mar 2021 12:56:34 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

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

These commands will give equal weight to each survey. This may or may not be something you want to do, but if you do not do it, each survey will count in proportion to its sample size. What do you plan to do with the pooled surveys?

---

---

Subject: Re: Appending files  
Posted by [Ugonna](#) on Wed, 10 Mar 2021 22:46:40 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hello Tom,

Thanks for your response.

I plan to do trend analysis and check temporal effects on under-five mortality in Nigeria from 2008 to 2018 using the pooled surveys.

---

---

Subject: Re: Appending files  
Posted by [Bridgette-DHS](#) on Thu, 11 Mar 2021 20:48:34 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

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

For a trend analysis, I recommend staying with the original weights and NOT doing that equalization step.

Sometimes, for a relatively rare event, users will pool surveys to get more statistical power. (This has to be done with caution, because the pooled population is not usually well defined.) That's a situation in which you might want to give each survey equal weight or change the weights to take

account of population size. But when studying trends, I don't think the results would be affected by such changes to the weights, so there's no reason to make those changes.

---

---

Subject: Re: Appending files  
Posted by [Ugonna](#) on Thu, 11 Mar 2021 23:47:52 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hello Tom and Bridgette,

Thank you for your kind response.  
much appreciated.

---