
Subject: Weighting Data for Pooled Indonesian DHS Dataset

Posted by [adelia](#) on Fri, 17 Jun 2016 04:07:21 GMT

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

Hi!

I have several questions regarding the weighting and survey set of Indonesia DHS dataset 2007 and 2012

So currently I want to pool the BR dataset for both periods and would like to generate a pooled wealth index using PCA in order to examine the change in wealth between the years, in which I need to weight the data. I'm using STATA for the analysis, and from what I understand I must do the following before I merge the datasets

1. De-normalizing the sample weight

```
gen wgt=v005/1000000  
gen wgt_denorm=wgt*total number of female population age 15-49 in the particular period /  
number of female age 15-49 interviewed in the survey
```

My question about this part is whether it is necessary to use the total number of female population age 15-49 in the particular period or it is sufficient to do the following command?

```
scalar TOTWT=1000000  
quietly summarize v005  
scalar T=r(sum)  
gen wgt_denorm=v005*TOTWT/T
```

If it is indeed necessary to use the total number of female population age 15-49 in the particular period, will it be okay for me to estimate this number by multiplying the total country population projected for that particular period with the percentage of female age 15-49 in the household population as reported in the DHS final report on that year?

2. Correcting the PSU and strata to make it specific for the particular year e.g. by adding 10000 for PSU and strata in 2007 and 20000 for PSU and strata in 2012

Question for this part, I generate the strata manually by using the following command

```
egen strata=group(v024 v025), label
```

Is this the correct strata to use for Indonesia DHS 2007 and 2012? Or should I use v022 instead?

3. Setting the survey design (after merging the dataset) with the following

```
svyset [pweight=wgt], psu(v021) strata(strata)
```

Are these steps correct and do I need additional steps to correct the weighting?

Thank you very much beforehand
