
Subject: Denormalizing data

Posted by [cbdolan](#) on Fri, 15 Apr 2016 15:57:47 GMT

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

I am using the 2007 and 2013/14 DRC BR and GE files. I have listened to both the YouTube videos <https://www.youtube.com/playlist?list=PLagqLv-gqpTN8IZQBy7vAYw10NjynAn2Z> as well as the webinar "Analyzing DHS data: Weights and other adjustments for the survey design". I have several follow up questions:

1. How does the use of the subpop command (ie. svy, subpop (rural) logistic....) impact the process for weighting my data? I want to run my specification on a sub-sample of only rural locations. However, when making the stratification adjustment I grouped on (v024 and v025).
2. Since I am combining surveys for a single country, spanning multiple years and then making comparisons across years I think I need to denormalize the data. I have used the attached document provided by Dr.Ren as a guide. Is the following process correct for the BR file? Also, can you please verify that the process of denormalization I've outlined is correct.

$V005^* = V005 \times (\text{total births in the country at the time of the survey}) / (\text{total number of births in the survey})$

PROCESS FOR DENORMALIZING THEN APPLYING WEIGHTS, CLUSTER AND STRATIFICATION ADJUSTMENT

*generate weight

$V005^* = V005 \times (\text{total births in the country at the time of the survey}) / (\text{total number of births in the survey})$

gen wgt=v005*/1000000

*make unique strata values by region/urban-rural)

egen stratum=group(ADM1_CODE v025)

*tell stata the weight (using pweights for robust standard errors, cluster (psu), and strata

svyset [pw=wgt],psu(v021)strata(stratum)

*prefix regrss with "svy:stata will now know how to weight your data and compute the right standard errors

3. After applying the above process the standard errors are larger. Is this because the clusters are independent, but the households within the same clusters are not independent. Accounting for v021 increases the standard errors.

File Attachments

1) [Note+on+de-normalization+of+DHS+standard+weight.pdf](#) ,
downloaded 2242 times

Subject: Re: Denormalizing data

Posted by [Bridgette-DHS](#) on Mon, 18 Apr 2016 16:09:27 GMT

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

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

According to Stata documentation, you should use the "subpop" option rather than reducing the file to the subpopulation of interest. For example, if you only want to estimate a model for the rural cases, you must construct a binary variable ("rural") that is 1 if v025=2 and 0 otherwise, and then include "subpop(rural)" in the command rather than, say, "if rural==1" or "if v025==2". If you try both, and do a comparison, you will see that there is, indeed, a small difference between the two. However, the difference is not in the estimates of coefficients, but in the standard errors, and the difference is usually very small (at least in several comparisons that I did) and can either increase or decrease the standard errors. So--my recommendation is that you do what Stata recommends, but if you don't, your conclusions are very unlikely to be affected.

Regarding denormalization, please see a response I just wrote to message #9538. I will add to that, however, a question about what you mean by the total number of births. You could use the BR file to get at the total number of births in, say, the calendar year before the survey, and then scale up to the UN Population Division estimate of the number of births in the population in that calendar year. However, I don't see how you could use the total number of births in the BR file. That does not correspond with any population estimate that you are likely to find anywhere.
