

---

Subject: Weighting Data Question

Posted by [nholla](#) on Mon, 23 Nov 2015 14:23:43 GMT

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

---

Hello DHS,

I'm currently trying to define my weighting and strata procedures in STATA and had a couple of questions. I'm appending 18 African countries over 3 decades (1980-2010) for a total of 54 total surveys for this analysis (I'm looking at nutrition and child mortality):

I'm using the BR file for each country and year, and am running the following code (before pooling) to denormalize the weights:

```
summarize v005
scalar T = r(sum)
gen weight = v005*fempop/T
```

Where fempop is the female 15-49 population for a particular country. My question is, when summing up the weights, should I be summing up the weights over the BR file or should I be summing them up over the observations in the IR file (since v005 is duplicated for each child of a women in the BR file I wasn't sure)?

I'm also a bit confused on how to define my strata variable. I'm using the following code:

```
gen strata = v101 * v102
egen strata_pooled=group(survey strata)
```

where "survey" is an identifier of the country and year of the survey. Assuming v101 always represents region and v102 always urban/rural, can I safely do this for the 18 countries over three decades that I am working with? It seems like I should be looking at Appendix A in the final report for the sampling methodology, but several of the reports are in French so it's hard for me to discern what's happening in these particular countries.

Thank you for all your help!!

---