
Subject: Setting the Data for Survey Analysis-Stata
Posted by [dnameispaone](#) on Tue, 13 Jun 2023 11:09:26 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi There,

I am trying to replicate the NFHS-5 tables related to Family Planning Indicators.

I using the following command to set my data for survey analysis

```
gen weight = v005/1000000  
gen psu = v021  
gen strata = v023
```

```
svyset psu [pw = weight], singleunit(certainty) strata(strata) vce(linearized)
```

```
svy, subpop (curr_married): tab ster, per missing ci format (%10.2f)  
svy, subpop (ever_marry): tab ster, per missing ci format (%10.2f)
```

Are the above commands correct? Particularly the svyset command?

Because my results are not matching those given in the NFHS-5 report.

Subject: Re: Setting the Data for Survey Analysis-Stata
Posted by [Bridgette-DHS](#) on Wed, 14 Jun 2023 11:50:09 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS staff member, Tom Pullum:

To construct tables, you do not need a full svyset command. For frequencies, percentages, means, and rates you only need to use weights. The adjustments for clustering and stratification only affect standard errors, confidence intervals, and test statistics.

To construct a table of A vs B, where A and B are categorical variables, you can choose between two options, either "tab A B [iweight=v005/1000000]" or "svyset [pweight=v005]" and "svy: tab A B". The first option is easier but they will give the same results.

If you use the first option, with iweight, you have to divide v005 by 1000000. If you use the second option, with pweight, you do NOT need to divide by 1000000. You can if you want, but it will make no difference for your results.

You would use the full svyset command with estimation commands, such as "regress Y X". You have specified it correctly. I would not usually include a design option ("vce()") in svyset but you

can certainly do that.

Subject: Re: Setting the Data for Survey Analysis-Stata
Posted by [dnameispaone](#) on Wed, 14 Jun 2023 14:33:41 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you very much for taking out the time to clear my doubts.

Subject: Re: Setting the Data for Survey Analysis-Stata
Posted by [fdsouza](#) on Tue, 20 Jun 2023 12:41:36 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi,

I am using the DHS India data to evaluate the effect of rainfall on birthweights. The variables I am using include birthweight, weight of mother, religion, caste, education and wealth. Just to confirm, do I need to survey set my data before running my regressions or not? I am not using any weighted variables.

Subject: Re: Setting the Data for Survey Analysis-Stata
Posted by [Bridgette-DHS](#) on Fri, 30 Jun 2023 15:03:38 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member, Tom Pullum:

Sorry for the delay in this response. All the variables you mentioned are in the standard recode files and are affected by the survey design--weights, clusters, and strata. You should use svyset and svy for any estimation commands.

The adjustments for weights, clusters, and strata apply to all cases, not to specific variables.
