
Subject: Aggregating up to district level in NFHS-4
Posted by [zoeg](#) on Thu, 22 Oct 2020 13:27:18 GMT

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

Hello!

I am trying to merge the NFHS-4 dataset with another dataset that contains variables at the district-level. Are there district weights somewhere in the NFHS-4 dataset that will allow me to aggregate up the individual-level responses to the district level? Has someone else done this and knows how to do it in Stata?

Thank you!

Subject: Re: Aggregating up to district level in NFHS-4
Posted by [Bridgette-DHS](#) on Mon, 26 Oct 2020 21:58:17 GMT

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

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

I don't know which file you want to use, but let's say it's the PR file (IAPR74FL.dta). In that file, district is given by shdistri. The numerical codes for the names will be given if you enter "label list SHDISTRI". In your district-level file, be sure that the district ID has the same name and the same codes. Sort that file by shdistri and save it. Then open the PR file and sort it by shdistri. Then merge with the saved district-level file. There are many examples of merges on the forum.

This will attach the district-level variables to the individual cases in the PR file. You can then treat those variables as covariates. You can collapse the file (try "help collapse") but that's not necessary. Ideally, you would do the analysis with multi-level models.
