

---

Subject: Re: Query on Cluster-Level Modeling with DHS Data and Sampling Weights

Posted by [sayianka](#) on Sun, 22 Sep 2024 04:33:48 GMT

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

---

Thank you for your detailed explanation.

I have successfully recreated the R version of the Stata code.

I needed clarification on this line:

```
collapse (first) v005 v023 (sum) nch4plus cases (mean) v190_*, by(v001)
```

I've noted that we use simple means for the variables in this line; Is this meant to be so, or we should weigh the means ?

I've also taken note of your comment: "Substantively, this would not be a good analysis of the data.", as this sentiment has been echoed elsewhere in this forum as well.

The only reason I am attempting this cluster-level analysis is to perform an analysis similar to what the DHS team does for geospatial covariates and modeled map surfaces. According to their report, this analysis must be conducted at the cluster level.

Thank you for your assistance.

---