Subject: Re: Interpolated surfaces
Posted by Janet-DHS on Mon, 18 Jul 2022 14:08:07 GMT
View Forum Message <> Reply to Message

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

The new (collapsed) file will have clusters as units. The two crucial variables are prop_diarrhea and wt_diarrhea. Prop_diarrhea would be the outcome variable and wt_diarrhea would be the weight for each cluster. It is a combination of the sampling weight for the cluster (every child in a cluster has the same value of v005) and the number of children for whom the individual-level outcome could be assessed. You don't need to do anything else with the weights, such as divide by 1000000. If you stay within Stata, and use pweights, Stata will renormalize so the mean weight is 1 for each unit (i.e. for each cluster).

I would expect you to carry along, in the collapse, other cluster-level variables such as place of residence (the same for everyone in a cluster), perhaps something related to wealth quintile, education level of the mothers, etc. You could have additional outcome variables, such as fever or cough, but because they would have different numbers of children in the denominator, they would have their own weights.

You would use the cluster GIS codes to merge with a file of cluster-level characteristics that are external to the DHS data. The clusters as geographic data points would be the basis for the interpolated surface. Good luck.