Subject: Link Service availability data in 2011 to Individual/HH?
Posted by vega25 on Thu, 13 Nov 2014 19:59:09 GMT
View Forum Message <> Reply to Message

Hello,

I see that the community questionnaire fielded in the Bangladesh DHS of 2011 is available in the "Service Availability" dataset. Is the way to link this file with the women's individual or household datasets simply to merge the two datasets on the "cluster" variable? Is there anything else I should be aware of while doing this?

Secondly, would anyone have a recommendation about whether cluster-level variables should be calculated (obviously?) from the Service Availability dataset where the relevant variables of interest are available in it, and/or calculated by aggregating all the individual responses within a cluster?

Thank you very much.

Subject: Re: Link Service availability data in 2011 to Individual/HH?
Posted by Bridgette-DHS on Tue, 16 Dec 2014 14:40:52 GMT
View Forum Message <> Reply to Message

Your post is being reviewed by a DHS moderator, and someone will respond soonest.

Thanks for your patience.

Subject: Re: Link Service availability data in 2011 to Individual/HH?
Posted by Trevor-DHS on Tue, 16 Dec 2014 19:07:48 GMT
View Forum Message <> Reply to Message

1) Yes, use the cluster number for matching.
2) Typically you would use the service availability data to add community attribute data to the individual data and then analyze the individuals. However you can use both calculate variables from the service availability data, or aggregate individual responses, but you have to be careful with what you are doing. If you are aggregating within cluster, realize that the aggregate you produce is based on very few cases and is unlikely to be very representative of that cluster. Also you will need to pay attention to weighting of data when aggregating individual data. Note that as the weights are calculated by cluster, you don't need to weight when calculating the cluster level aggregates, but you will need to weight when aggregating up further, and for that reason you may need to keep aggregates of the numerators and denominators at the cluster level too to facilitate that aggregation. Mostly, though, users use the service availability data as community attributes with the individual level data.