## Subject: Linking GPS data to DHS data files
Posted by davidk35 on Sat, 05 Aug 2017 17:33:04 GMT
View Forum Message <> Reply to Message

Hi everyone,

I am currently creating a panel dataset for African countries, which uses years and administrative regions at the first level (as given by gadmn) as the units of analysis.
As I would like to compute for example the share of young men (15-24) to the adult population (e.g. until 54) in a specific region and year, I need to assign each cluster mapped by the GPS file to the corresponding cluster in the DHS file (e.g. GNGE52FL and GNMR52DT).
I thought I could do this by the DHSID but it turns out that the DHSID is not equivalent to any combination of variables in the DHS file (e.g. I thought of creating a string in STATA with the code egen = concat (mv000 mv007 mv001) to create the DHSID, but the result is not eqivalent to the DHSID in the GPS-file wich is GN20050000001 for the first cluster for example).

therefore my question would be how to match/merge the DHS-file with the GPS-file such that I can do regional analysis using the clusters in this region for a given (survey) year?

Thank you very much for your help in advance.

## Subject: Re: Linking GPS data to DHS data files
Posted by Bridgette-DHS on Mon, 07 Aug 2017 20:37:50 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Specialist, Trinadh Dontamsetti:

The DHSID column you are referring to does not have a link to any of the variables you mentioned. The DHSID GN20050000001 is simply the country code (GN), survey year (2005), and the cluster number (001), spaced with placeholder zeroes. To link the GPS data file to the survey dataset, you will need to reference the DHSCLUST column in the GPS file, which should match v001 or hv001 from the recode. Doing so should allow you to match the two datasets.