
Subject: A general approach to merge GPS and Individual Data

Posted by [Doon](#) on Wed, 09 Mar 2016 18:02:55 GMT

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

I am doing an analysis that requires geographic information. I downloaded the GPS, Individual Recode, and Men Recode for all sub-Saharan African countries when available. I have a file of all individual recodes for different countries and phases, all appended and a separate shapefile with all the available GPS files merged. I have a question regarding the approach I need to use to merge the geographic information to the individual files.

In the geographic files, the identifier is of the form:

I use the case of Liberia, phase 5 as an example (File LBGE5CFL.shp)

Survey = Two letter country code + Year , ex: LB2009

Cluster = Cluster number, ex: 1

=> I thought a unique identifier would be of the form LB2009_1

However, it is not easy to match such information with the individual recode file.

There are 2 Stata files that can be downloaded: (LBIR51FL and LBIR5AFL) They both have the country code LB5, although they run for years 2007 to 2009. My question is then: are they run in the same villages, meaning that I can merge all the clusters in these files with those in LBGE5CFL.shp, or can I just only merge the subsample of the individual recode that were collected in 2009?

In other words,

What is the best way for me to create a unique identifier in the appended file of *all* the individual recodes in sub-Saharan Africa that I'll manage to merge correctly with the GPS files, provided that phase of the DHS collected GPS information?