
Subject: Merging GPS and household data
Posted by [imcohen](#) on Mon, 08 Aug 2022 19:07:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello,

I am interested in linking household records to geocoordinates for a number of Uganda DHS surveys, but have found that not all clusters appear in the relevant GPS data. When linking hv001 with dhsclust, I've found in most cases a significant fraction of clusters are missing. If it's relevant, I've been using `st_read` from the `sf` library in R to access the shp files. A specific list of the number of records in each is below.

I know some clusters may not have been geocoded, but just want to ensure that I'm not making any mistakes and that these numbers of mismatches are as expected. Any guidance on: 1) if there's some other place to find this data; 2) any issues that might be causing this; 3) any suggestions on other ways to find relevant ADM2 units for households; and 4) any clarity on why the GPS coordinates are missing would be greatly appreciated!

The specific datasets are:

Uganda 2001 DHS: 298 unique clusters in UGHR41FL, 189 observations in UGGE43FL.shp
Uganda 2006 DHS: 368 unique clusters in UGHR52FL, 280 observations in UGGE53FL.shp
Uganda 2009 MIS: 170 unique clusters in UGHR5AFL, 143 observations in UGGE5AFL.shp
Uganda 2011 DHS: 404 unique clusters in UGHR61FL, 340 observations in UGGE61FL.shp
Uganda 2014 MIS: 210 unique clusters in UGHR72FL, 181 observations in UGGE71FL.shp
Uganda 2016 DHS: 696 unique clusters in UGHR7BFL, 562 observations in UGGE7AFL.shp
Uganda 2018 MIS: 340 unique clusters in UGHR7IFL, 270 observations in UGGE7IFL.shp

Many thanks,
Isabelle

Subject: Re: Merging GPS and household data
Posted by [brad-DHS](#) on Tue, 09 Aug 2022 14:54:55 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello Isabelle,

I took a look at the two most recent surveys and am seeing the correct number of clusters. I see 340 clusters in the 2018 MIS shapefile and 696 in the 2016 DHS shapefile. Some of the clusters in both datasets are set to the coordinates 0,0 because they are missing or withheld. The 2018 MIS has 24 clusters at 0,0 and the 2016 DHS has 11. But, even when subtracting these numbers from the total cluster counts, I'm not getting 270 and 562. Given the information you provided, I think you might need to do some troubleshooting with your code. If you would like to quickly verify that your shapefiles have the correct number of clusters (without using GIS), you can open the DBF file in Excel and count the rows.

As you continue troubleshooting this, please let me know if you have any additional questions in the thread below. I hope this helps!

Cheers,
Brad

Subject: Re: Merging GPS and household data
Posted by [imcohen](#) on Wed, 10 Aug 2022 17:16:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Brad,

Many thanks - your response was really helpful. I took another look at my code, and am now getting 24 missing clusters in the 2018 MIS and 11 in the 2016 DHS.

Best,
Isabelle
