
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
