Subject: Identifying geographic location.
Posted by JonKoch on Wed, 08 Nov 2023 14:14:23 GMT
View Forum Message <> Reply to Message

Hello dear users and staff,

I am currently trying to write my thesis on preferred family size and polygyny. I want to replicate a study done by Ezeh on the topic in 98, but with more current data. For my research design, I want to create an index of areas with high/middle/low levels of polygyny and identify average desired family size, but unfortunately, I have found myself unable to ascribe any meaning to the cluster variable. Can anybody tell me how to group the subjects by geographic location within Kenya?

If anybody has experience with a similar region/variable design, some advice would be greatly appreciated. I am working with RStudio.

Subject: Re: Identifying geographic location.
Posted by Bridgette-DHS on Thu, 09 Nov 2023 17:06:37 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS staff member, Tom Pullum:

Clusters are census enumeration areas that are used as primary sampling units in the sample design. Typically about 30 households are sampled from each cluster. Clusters loosely correspond with neighborhoods in urban areas and villages in rural areas, but they may be internally heterogenous. You can construct cluster-level variables such as the proportion of households that are polygamous. There are separate geographic files that give cluster-level characteristics.

The geographic coordinates of clusters are slightly displaced, to maintain anonymity, and it is difficult to combine them into meaningful groups. Many surveys include districts, a generic label for the 2nd administrative level. You could also treat districts as units.

I hope other users will add their thoughts.

Subject: Re: Identifying geographic location.
Posted by JonKoch on Wed, 15 Nov 2023 14:26:54 GMT
View Forum Message <> Reply to Message

Thanks for your reply. Could you perhaps tell me whether there is any sort of codebook or guide specifically for the Kenya data that I might be overlooking?

As for the geographic data: I do not need exact information on where the clusters are situated, but the rough administrative district or region would be great so that I can bundle the clusters

sensibly. The clusters themselves are a little too small for my analysis, so grouping them together seems like the way to go.

Is the geographic data the one that has to be downloaded separately? I haven't done that yet as I'm not sure of what exactly the content of it is. I assumed, it was data on GPS coverage, which didn't seem important for my design.

Can the geographic /GPS data be merged with the survey data sets so that the location of the clusters becomes evident?

Excuse me if those are too many basic questions. If you can direct me to a good place to read up on it, I will do so.

Subject: Re: Identifying geographic location.
Posted by Bridgette-DHS on Wed, 15 Nov 2023 15:46:10 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS staff member, Tom Pullum:

You should be able to find a file "KEGE8AFL.dta", which gives the (displaced) coordinates of the clusters. You can merge with the IR file, for example, matching v001 in the IR file with DHSCLUST in the GE file. In the IR file, v024 is region. "Region" is the generic label for v024, the first administrative level, but in Kenya these are counties. When the clusters are displaced, they are kept in the original county--they are never moved into an adjacent county.

I don't believe any smaller geographic units are coded in the data. Let us know if you have difficulty with the merge.

Subject: Re: Identifying geographic location.
Posted by JonKoch on Tue, 21 Nov 2023 12:14:46 GMT
View Forum Message <> Reply to Message

Thank you very much. V024 worked perfectly and I also managed to classify the countries according to polygyny thresholds.
