
Subject: Re: Egypt DHS District Data
Posted by [user-rhs](#) on Sun, 16 Mar 2014 18:59:26 GMT
[View Forum Message](#) <> [Reply to Message](#)

Amira,

I'm not a QGIS user, so my explanation is how to do the join in ArcGIS. I imagine the steps are translatable to QGIS. If you do not know how to do them in QGIS, consult the user manual or help file for QGIS (<http://www.qgis.org/en/docs/>).

The way I would do it is to spatially join GIS data first in QGIS, and THEN once I have the district info (from the spatial join), I would join with the rest of the DHS woman's recode file in my statistical software.

To do this:

1. Join the GIS data spatially in QGIS. When you download the GIS file from the DHS site, unzip the entire folder into its own folder in the working directory. It is crucial that your working folder has all of the files in that zip file in there (file extensions DBF, PRJ, SBN, SBX, SHP, etc.), because those files are what will allow you to plot the cluster centroids on your map. Here's a screenshot of where I unzipped some Ethiopia DHS GIS files so I can spatially merge it with a layer that contains Woreda boundaries:

2. Open QGIS, and add the Egypt administrative boundary layers (province, district, etc.) that you want to join to the DHS data.

3. Then import the data points into QGIS.

4. Spatially join the data points with the district layer. You do a spatial join because there is no identifier that is common to the district layer and your DHSGIS file that would allow you to join by variable.

5. If you have the option of exporting the joined layer as a new layer or new shapefile, select that option. Make note of where the file will be saved, as the software will create all of the files necessary to map that layer the next time you open it in QGIS.

The next steps assume that QGIS creates a DBF file when you export a joined layer. DBF is basically where the data are stored. This DBF file will contain the original variables in the DHSGIS file in addition to the data from the boundary layer, namely, the district names you are interested in.

6. Take the DBF (or however QGIS stores the data from the spatial join) and import it into your statistical package (SAS, Stata, SPSS, etc.). Rename DHSCLUST into v001 and sort by v001. Join with the DHS woman's recode file using the v001 variable, taking care to sort the woman's recode file by v001 before joining w/ the DHS file.

Opening the joined file in different packages: If you are a SAS user, you can import the DBF directly using

```
proc import out= dhs_geoc
  datafile= "&loc.\JOINED FILE LOCATION\joined_file_name.dbf"
  dbms=dbf replace;
run;
```

If you are a Stata user, you should open the DBF in Excel, select the entire sheet, copy, and then manually paste into a Stata data editor window.

Since I don't work in SPSS unless I really have to, I can't tell you the exact steps, but I assume it's fairly straightforward to open DBFs (file->open->change file type to "all", etc.). Stata is the stubborn one when it comes to opening files that are not "stata data file" format.

hth,
RHS

File Attachments

- 1) [gis_example.jpg](#), downloaded 1526 times
-