
Subject: Beyond District Infon in DHS data 2018 Pakistan

Posted by [khan88](#) on Wed, 27 Dec 2023 16:05:13 GMT

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

Hi, Does anyone know code book to indentify household cluster or cluster numbers or any information beyond district numbers in DHS 2018-19 ? I am conducting research to see vulnerbality within districts i.e. require info for tehsil/villages within each districts, if that can be provided? Thank you

PS: I dont need any indentity information, just data like

District___Village or Tehsil codes/names so i can perform analysis within each district in Pakistan.

Regards

Dr. Khan

Subject: Re: Beyond District Infon in DHS data 2018 Pakistan

Posted by [Janet-DHS](#) on Tue, 02 Jan 2024 17:10:02 GMT

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

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

You can merge the HR file with the GE file using the cluster id in the HR file, hv001, and the cluster id in the GE file (the last three characters of DHSID; I will paste below the Stata lines to do this merge). If you have a shape file for Bangladesh that includes the district boundaries at the time of the 2017-18 survey, then you can identify which district each clusters is in. The latitude and longitude of the clusters (LATNUM and LONGNUM) have been displaced slightly, but they should have been kept within the same admin 2 units (districts).

* Merge the HR and GE files from the BD 2017-18 DHS survey

* Specify workspace

```
cd e:\DHS\DHS_data\scratch
```

```
use "...BDGE7SFL.DTA"  
gen hv001=substr(DHSID,12,3)  
destring hv001, replace  
sort hv001  
save temp.dta, replace
```

```
use "...BDHR7RFL.DTA"  
sort hv001  
merge hv001 using temp.dta  
tab _merge  
drop _merge
```

* Save the merged file with a new name
