Subject: Re: Next IPUMS-DHS Release
Posted by kingx025 on Wed, 01 May 2019 19:52:53 GMT
View Forum Message <> Reply to Message

I suggest that you use IPUMS-DHS to avoid the frustration you experience when working with the original DHS files, with labels varying across samples. We recode the data from the original DHS files so that the same meaning is given the same code and label across samples, and you don't have to deal with the issue you describe.

For example, when you are working with multiple samples for a country, choose the "Integrated geography" variables from the drop down menu in IPUMS-DHS:
https://www.idhsdata.org/idhs-action/variables/group?id=geog _integ

This integrated region variable includes as many samples as possible and has consistent codes identifying areas with the same geographic footprint across all samples included for a country. There may be less detail than in the single sample region variables but there will be comparability in codes, labels, and meanings for the GEO_ integrated variables.

Our data harmonization is designed to insure that the same variable meaning has the same codes and labels across samples; that is why we integrate the data. We also release single-sample region variables for all samples, because they sometimes include more detail than is available in the integrated geographic variable, but most other variables are integrated across multiple sample years (and, apart from a few variables dealing with geography and ethnicty, across countries).

Miriam King
IPUMS-DHS Project Manager