
Subject: Re: Confine my Dataset

Posted by [kingx025](#) on Sat, 21 Apr 2018 22:42:41 GMT

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

I am not sure of your motivation for limiting your dataset, but an easy solution is to create a customized dataset with just the samples and variables you need from the IPUMS-DHS database (at idhsdata.org). This will limit the number of variables and size of the file to what fits your research project. Downloading data from IPUMS-DHS also saves you the work of merging across file types, because the IPUMS-DHS staff have attached household characteristics to the files for women and attached the household and mother's characteristics to files for young children and births. Once you have made your datafile, check the universe (who is included with answers in variable) in the documentation for variables of interest. For example, you will find that for many variables relating to children, dead children are coded as "Not in universe" and don't have meaningful responses for issues like recent illness or nutrition. If you want to further restrict your population for analysis, I think it would be better to do it with programming statements after you have created a dataset for analysis.

IPUMS-DHS currently includes DHS data for 28 countries; you don't mention which countries interest you.

Miriam King
