Subject: Merge HIV-Positive Data with Individuals
Posted by Yawo on Tue, 26 Feb 2019 02:24:24 GMT
View Forum Message <> Reply to Message

Hello,

I am starting a project linking HIV seropositivity with a host of covariates from women and men data files.
I am pooling data across countries and across years..

1. Can I link the HIV dataset to the IPUMS data?
If so, do I create my set of samples/data from IPUMS first, and then link them to the HIV data?

2. Also, does the IPUMS-DHS include men-recode data? If not, how do I include men in my data?

Thanks - Yy

Subject: Re: Merge HIV-Positive Data with Individuals
Posted by kingx025 on Tue, 26 Feb 2019 20:41:41 GMT
View Forum Message <> Reply to Message

I also posted answers on the IPUMS User Forum, and I repeat those answers here.

To link IPUMS DHS data to the HIV test result data:
I found the following post on the DHS User forum, about merging HIV test result files with other DHS data files:
"I believe there are person identifiers in the HIV dataset called hivclust, hivnumb, and hivline, the combination of which should be a person's unique identifier. These represent the cluster, household, and line number variables, respectively. The easiest way is to rename these variables to v001, v002, and v003 (as these variables are called in the IR dataset), respectively. If merging with male recode dataset, then you would rename these variables to mv001, mv002, and mv003. After renaming hivclust, hivnumb, and hivline, and sorting by the appropriate variables (v001, v002, and v003 for women and mv001, mv002, and mv003 for men's dataset), you should be able to merge the datasets on these variables."

If you chose women as your unit of analysis in IPUMS-DHS and opt for showing the original DHS variable names, you can select V001 (IPUMS name: CLUSTERNO), V002 (IPUMS name: HHNUM), and V003 (IPUMS name: LINENO) as elements of your customized data extract. Then you would rename the HIV dataset variable names to match and merge on these variables. I would do this for one dataset at a time, or add some other linking variables to distinguish between samples, such as COUNTRY and YEAR.

When we add men as a unit of analysis (very soon), the MV001, MV002, and MV003 identifier variables will be available for IPUMS-DHS samples, as well.

We are currently working on harmonizing standard variables for men for 115 DHS samples and expect those data will be available through IPUMS-DHS by May, 2019, at the latest.

I hope this helps.