Subject: Creating unique household id in Women's data
Posted by priyoma on Mon, 09 Jan 2017 14:59:51 GMT
View Forum Message <> Reply to Message

Hello.

I am a student, I am working on my thesis with data for India-NFHS Round 2. The data file is IAIR42FL, ie the women's daya. In the data there is a unique caseid for each of the 90303 women surveyed.
I am doing an analysis of the effect of female education on the fertility of women in India. Now I need to create a unique Household ID for this data, can you please give me a STATA code which does that.
When I tab v002 (ie household number) I get frequency as 588 and even higher in some,so it is not that there are 588 individuals in a household right? So I need a unique HHID ie tells me which of the 90303 women belong to which household.

Awaiting a quick reply.

Subject: Re: Creating unique household id in Women's data
Posted by Bridgette-DHS on Tue, 10 Jan 2017 21:49:10 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Quote:In IR files, v001 is the cluster number; v002 is the household number (within v001) and v003 is the woman's line number (within v001 and v002). Together, v001 and v002 identify the household; v001 and v002 and v003 identify the individual woman. In IAIR42FL.dta, caseid is a character string that includes v001, v002, and v003. To get a household id, you can remove the last three columns (which give v003) from caseid. That is, use this line in Stata:

gen hhid=substr(caseid,1,12)

Subject: Re: Creating unique household id in Women's data
Posted by priyoma on Thu, 12 Jan 2017 08:08:12 GMT
View Forum Message <> Reply to Message

Thank you. Just to clarify.

So with this code, if someone's caseid is 2 1 66 4 and another's caseid is 2 1 66 7, then they belong to the same household with hhid 2 1 66? Thus there are two individuals belonging to the same household?

With this hhid, can i generate household fixed effects in a regression either by i.hhid as a

regressor or by manually generating household dummies?

Regards,
Priyoma

Subject: Re: Creating unique household id in Women's data
Posted by Bridgette-DHS on Fri, 13 Jan 2017 12:55:34 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Quote:In that example, the household has two women, who are on lines 4 and 7. The hhid is a combination of cluster id (hv001 or v001) and household id (hv002 or v002). (The line number is hvidx or v003.)

Most households will have one eligible respondent--if they have any at all. Only a few will have more than one. You will use up all your degrees of freedom if you have fixed effects for household. I would not even try random effects for households. The density of women per household is too low. Fixed or random effects for cluster would be as far down into the data as I would go.

You would definitely use "xi." for fixed effects--no need to construct dummy variables (fortunately!). For random effects, use an "me" model.

However, we really cannot advise on your choice of model, just on data-related issues and to a limited degree on Stata syntax.