Subject: Residing in same household in KR file
Posted by kmdshoyaib on Tue, 05 Mar 2024 02:14:20 GMT
View Forum Message <> Reply to Message

I am working on a analysis related to vaccine dropouts and vaccine delays using KR file for most recent survey in India. I have considered multiple factors at the household level to assess the risk factors for delay in vaccination (such as Distance for Health facility: Big problem, Type of Residency etc).

In this regard I would like to generate a binomial variable "Residing in same household: Yes or No" for all the entries in KR file to account for the clustering effect on the other variables during the regression analysis. I am NOT looking to generate any variable for siblings.

I am using STATA MP 18 for analysis.

One way I see how it can be done is ... for all observations, if cluster number is same, then within that cluster if household number is same then for that observation, Residing in same household can be marked as Yes or else No. But for this statement i am unable to write a command in STATA.

Kindly help me with a STATA command or provide a stata command to accomplish the task using some other methodology.

Any help is really appreciated.

Thanks in advance

Subject: Re: Residing in same household in KR file
Posted by Bridgette-DHS on Tue, 05 Mar 2024 14:54:45 GMT
View Forum Message <> Reply to Message

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

In the KR file for the NFHS's, children in the same household will have the same values of v024 v001 v002. Children with the same mother (in the household) will have the same values of v024 v001 v002 v003. You cannot do what you want to do by just adding a binary variable for "same household" or "same mother" because you want to describe a relationship between cases.

One way to approach this would be with a multi-level model. You can find some literature on household clustering or maternal clustering. Frankly, however, I don't think that will be very helpful.

I suggest another approach, which is to construct a data file of PAIRS of children in the KR file who have the same mother. You might expect a huge number of such pairs but I calculate that there would only be 62,170 pairs in the KR file for the NFHS5, compared with 232,920 children in that file.

I suggest that you use the following Stata program to construct a file that pairs the child with bidx=1 with the child with bidx=2, for women who have 2+ children. There are 50,253 such pairs. If you bring in other pairs, such as pairs of bidx=1 and bidx=3, you will be repeating the children with bidx=1. With this file you can look at similarities and differences between the two children. You will, however, have some difficulties. For example, there will not be many pairs in which both children are age 12-23 months. Note that some of these pairs are twins. Hope this will help.

```
* Construct a file of pairs of the two most recent births in the same household

cd e:\DHS\DHS_data\scratch

use "...IAKR7EFL.DTA", clear
tab bidx
rename v024 state
rename v001 cluster
rename v002 HH
rename v003 MO_line

save temp0.dta, replace
keep if bidx==1
rename b* b*_1
rename m* m*_1
rename h* h*_1
rename HH hh
rename MO_line mo_line
save temp1.dta, replace

use temp0.dta, clear
keep if bidx==2
keep state cluster HH MO_line b* m* h*
rename b* b*_2
rename m* m*_2
rename h* h*_2
rename HH hh
rename MO_line mo_line
merge 1:1 state cluster hh mo_line using temp1.dta
tab _merge
keep if _merge==3
drop _merge

* File with the data for the two most recent births, if 2+
save temp12.dta, replace
tab bidx*
```

Subject: Re: Residing in same household in KR file
Posted by RashadUpton on Mon, 21 Jul 2025 06:31:03 GMT
View Forum Message <> Reply to Message

Step 1: Create a unique household identifier
gen hh_id = v001*1000 + v002
Step 2: Calculate the number of people in each household
bysort hh_id: gen hh_count = _N
Step 3: Create binary variable "Resides with household"
gen same_household = hh_count > 1

Where geometry dash:
v001 is the cluster code
v002 is the household number
hh_id is the unique code for each household
_N in bysort will count the number of observations in each hh_id
same_household will be:
1: if the household has >1 person (has someone living with the household)
0: if the household has only 1 person (no one living with the household)