Subject: Understanding variables in PR & IR files
Posted by Mercysh on Mon, 05 Jan 2015 11:31:50 GMT
View Forum Message <> Reply to Message

I am doing data quality checks before analysis and checking consistency between hv112, child
reported mother's line number (hv112) and b16 children's line number and hc60 (line number of
mother from women's questioonaire).
I merged the Lesotho PR & IR files for 2009 using the following code:
use LSIR60FL
sort v001 v002 v003
save LIR_merge.dta
use LSPR60FL.DTA
gen v001 = hv001
gen v002 = hv002
gen v003 = hv112
save LPR_merge.dta
merge m:1 v001 v002 v003 using LIR_merge.dta
*drop if IR & PR matched
drop if _merge==3
tab b16_* if b16_* !=0
*only for complete women questioonaire
tab hv112 hc60 if hc60 !=994

Please help me understand how the codes 993 & 995 are derived when hv112 is valid line
number (1 to 14).
I also don't know if I am interpreting correctly that there are women who reported individuals as
their children but the children did not report them as their mothers.

Lastly:
using Lesotho PR file for 2009
I weight the data:
generate weight= hv005/1000000
tab hv105 [iweight= weight]
I get total of 43 162, which is 10 000 more than the 33 719 total reported in the report.

Please help

Subject: Re: Understanding variables in PR & IR files
Posted by Bridgette-DHS on Mon, 26 Jan 2015 15:04:46 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata specialist, Tom PullumFirst, hc60, like all of the hc
variables, is only coded for young children. In LSPR60FL.dta, hc60 is missing (".") if hc1 (months
of age) is missing, i.e. if hc1>59. If you do "label list hc60" you get this:

. label list hc60

hc60:
        993 mother not de facto
        994 incomplete woman's interview
        995 not in household

If you then do "tab hc60" in LSPR60FL.dta, there are 2,584 non-missing values, one for every child with hc1<=59. Of these, 29 have code 993; 233 have code 994, and 591 have code 995.

Code 993 would typically mean that the mother was in the household but she was a "visitor" who did not sleep there the previous night, so she was not eligible for the woman's interview; code 994 means "incomplete woman's interview", i.e. she is a de facto member in the household listing but is not in the IR file, even though she should have been; code 995 means that the mother was not in the household at all, in which case the child should have hv112=0. However, I see that of the 591 children with hc1<=59 and hc60=995, there are 125 who do not have hv112=0. There are 470 children with hc1<=59 and hv112=0, and 4 of them do NOT have hc60=995.

I doubt that these 125+4=129 apparent discrepancies between hc60=995 and hv112==0 are data quality problems, because codes such as these are usually forced to be consistent. If I had more time, I would look into that.

However, I hope this answers your question about the coding of hc60.