Subject: Re: Creating a household level variable
Posted by Trevor-DHS on Thu, 20 Nov 2014 05:02:45 GMT
View Forum Message <> Reply to Message

It looks to me that you have < and > reversed in a couple of cases:
1) In your keep condition, if hv105>=97 & hv102==1 probably should be if hv105<=97 & hv102==1
2) replace adolescents =1 if hv105 <=15 & hv105>=17
should be
replace adolescents =1 if hv105>=15 & hv105<=17

I'm not quite following your code, but if I understand it correctly then here is how I would create the household variable:
use "LSPR60FL.DTA"

* Recode age into the relevant age groups
recode hv105 (20/59=1)(60/97=2)(0/19=3) if hv105<=97 & hv102==1,
gen(household_composition)

* Collapse and use the minimum value for each household
collapse (min) household_composition, by(hhid)
lab def hh_comp 1 "adult presence (ages 20-59)" 2 "older adult (ages 60 and over)" 3
"children-adolescents only (0-19 years)"
lab val household_composition hh_comp

* Save this variable for merging
keep hhid household_composition
sort hhid
save "hh_comp.dta"

* Open women's data file
use "LSIR60FL.DTA", clear
* generate hhid from caseid by dropping the last three characters for the line number
gen hhid = substr(caseid,1,length(caseid)-3)
sort hhid
* merge the data - many women to one household
merge m:1 hhid using "hh_comp.dta"
* drop households without interviewed women
drop if _merge==2

* tabulate to check.  Note some cases are missing due to households with no de jure members or
ages all don't know or missing
tab household_composition,m