

---

Subject: Creating a household level variable  
Posted by [Mercysh](#) on Wed, 19 Nov 2014 17:15:23 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

I am trying to generate a variable household composition using ages of household members (hv105) in households in which at least a member is aged 0-19 years. The categories should be adult presence (ages 20-59), older adult (ages 60 and over) and children-adolescents only (0-19 years). In multigenerational households the order of precedence should be the same as above. I have the following code but it does not seem to work:

I use the Lesotho 2009 (LSPR60FL.DTA) household dataset and keep only if hv105>=97 & hv102==1

\*\*\*\*\*

```
sort hhid
egen hhrestrict=tag(hhid)
```

```
*Create Adolescents
gen adolescents =.
replace adolescents =1 if hv105 <=15 & hv105>=17
```

```
*Households with adolescent member
egen hhadolescents = max(adolescents), by (hhid)
tab hhadolescents if hhrestrict==1
```

```
**Household composition
gen household_composition=.
replace household_composition =1 if(hv105 >=20 & hv105 <=59) & hhadolescents ==1 &
hhrestrict ==1
replace household_composition =2 if hv105 >=60 & hhadolescents ==1 & hhrestrict ==1
replace household_composition =3 if hv105 <=19 & hhadolescents ==1 & hhrestrict ==1
```

I merge with the (LSIR60FL.DTA) individual dataset after restricting to if v502==0, v012 <=17, v135==1, v531 >=90 and is still fine, the first output is

```
. tab household_composition
```

| Household composition  | Freq. | Percent | Cum.   |
|------------------------|-------|---------|--------|
| -----+-----            |       |         |        |
| Adult                  | 683   | 65.74   | 65.74  |
| Older adult            | 263   | 25.31   | 91.05  |
| Child-adolescents only | 93    | 8.95    | 100.00 |
| -----+-----            |       |         |        |
| Total                  | 1,039 | 100.00  |        |

and only a problem when I add the svy command here:

svy: tab household\_composition and gives the following output:

|                  |   |           |                 |    |          |
|------------------|---|-----------|-----------------|----|----------|
| Number of strata | = | 19        | Number of obs   | =  | 63       |
| Number of PSUs   | = | 53        | Population size | =  | 62.20275 |
|                  |   | Design df | =               | 34 |          |

```
-----  
Household |  
compositi |  
on        | proportions  
-----+-----  
Child-ad |      1  
-----
```

Key: proportions = cell proportions

I have already svyset  
generate weight= v005/1000000  
svyset [pw=weight], psu( hv021) strata(hv023)

Please advise

Thank you

---

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

---

---

Subject: Re: Creating a household level variable  
Posted by [Mercysh](#) on Thu, 20 Nov 2014 14:22:31 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Its my first time to use collapse, instead I was using tag which did not work. Thank you for your help.

---