Subject: Multiple datasets
Posted by annette.cassy on Wed, 01 May 2019 16:42:53 GMT
View Forum Message <> Reply to Message

Hi
I am trying to do some analysis using more than one dataset.
I have appendend but I don't know how to use the weights with more than one dataset.
Can you help me, please?

I need to do:
1. Propottion of HH with universal covarage (I need to understand which variables to use) / year (survey) / province
2. % of people who sought care in Community health workers / year (survey) / province
3. % of under 5 that slept under ITN / year (survey) / province
4. % of population that slept under ITN by gender / year (survey) / province
5. % of HH with IRS / year (survey) / province

Subject: Re: Multiple datasets
Posted by Bridgette-DHS on Wed, 01 May 2019 18:38:27 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Lead Malaria Research Analyst, Cameron Taylor:

Thanks for your question. For appending datasets you do need to take into account weighting. You need to construct unique ID codes for the clusters (PSUs) in the separate surveys. I suggest the following way of creating a new cluster and stratum id that is unique taking into account both survey years. Construct a unique PSU ID with "egen clusterid=group(hv007 hv021)". Similarly, if hv024 is the stratum variable, you can construct a unique stratum ID with "egen stratumid=group(hv007 hv024)". Then you put clusterid and stratumid in the appropriate places in svyset.

For your questions about how to calculate different indicators I have included the code below. In each of the example code I limited the tabulation to the Mozambique 2015 data (hv007==2015), since I want to match the report to make sure my tabulation is correct before proceeding. I wasn't sure what you were wanting with request #2 (% of people who sought care in Community health workers). However, if you are wanting to look at malaria care seeking data you need to use the KR file since your unit of analysis is children with fever.

1. Proportion of HH with universal coverage (I need to understand which variables to use) / year (survey) / province
5. % of HH with IRS / year (survey) / province

append using MZHR62FL.DTA MZHR71FL.DTA

egen clusterid=group(hv007 hv021)
egen stratumid=group(hv007 hv024)

```
gen wgt=hv005/1000000

svyset [pw=wgt], psu(clusterid) strata(stratumid)

//Household ITN ownership
      gen mal_ITNinHH=0
          forvalues x=1/7 {
                  replace mal_ITNinHH=1 if hml10_`x'==1
          }
      lab var mal_ITNinHH "Household owns at least one ITN"

//Number of ITNs per household
      gen mal_numitnhh=0
          forvalues x=1/7 {
                  gen itnhh_0`x'=(hml10_`x'==1)
          }
      replace mal_numitnhh=itnhh_01 + itnhh_02 + itnhh_03 + itnhh_04 + itnhh_05 + itnhh_06 +
itnhh_07
      lab var mal_numitnhh "Number of ITNs per household"

//Potential ITN users in Household
      gen mal_potuse = mal_numitnhh*2
      lab var mal_potuse "Potential ITN users in household"


//Households with > 1 ITN per 2 members
//Potential users divided by defacto household members is greater or equal to one
      gen mal_hhaccess = ((mal_potuse/hv013)>=1)

      lab var mal_hhaccess "Households with >1 ITN per 2 household members"

//IRS Household Sprayed
//Was the household sprayed in the past 12 months by someone other than a household member
      g irs=0
      replace irs=1 if hv253==1 & (hv253a==1|hv253b==1|hv253c==1)
      lab var irs "Household sprayed in the past 12 months by someone other than a household
member"

//TABULATIONS
svy: tab hv025 mal_hhaccess if hv013>=1 & hv007==2015, row ci obs

svy: tab hv025 irs if hv007==2015, row ci obs



3. % of under 5 that slept under ITN / year (survey) / province
4. % of population that slept under ITN by gender / year (survey) / province
```

```
append using MZPR62FL.DTA MZPR71FL.DTA

egen clusterid=group(hv007 hv021)
egen stratumid=group(hv007 hv024)

gen wgt=hv005/1000000

svyset [pw=wgt], psu(clusterid) strata(stratumid)


//Categorizing nets
        gen mal_netcat=0 if hml12==0
        replace mal_netcat=1 if hml12==1|hml12==2
        replace mal_netcat=2 if hml12==3
        lab var mal_netcat "Mosquito net categorization"

//ITN net variable
        gen mal_itn=(mal_netcat==1)
        lab var mal_itn "ITN"

//TABULATIONS
*Request #3
svy: tab hv024 mal_itn if hv103==1 & hml16<5 & hv007==2015 [iw=wgt] , row ci obs

*Request #4
svy: tab hv024 mal_itn if hv103==1 & hv007==2015, row ci obs

svy: tab hv104 mal_itn if hv103==1 & hv007==2015, row ci obs
```

Subject: Re: Multiple datasets
Posted by annette.cassy on Wed, 01 May 2019 19:04:22 GMT
View Forum Message <> Reply to Message

Thanks a lot.
I am going to try this right now.