Subject: Replicating DHS tables on child health
Posted by lexgw on Tue, 17 Nov 2015 00:00:31 GMT

View Forum Message <> Reply to Message

Hello, I am trying to replicate tables on prevalence of Diarrhea, fever, ARI & weight and size at birth for Uganda DHS 2011 Report. Questions;
1. How do I generate the variables (mother's age at birth-grouped & mother's smoking status) in table 10.1 page 124 to replicate the rows for these two variables in stata. I have managed to replicate all other rows other than for these two variables.
2. Table 10.7 page 132, how do I replicate values in column two (diarrhea with blood),(-in fact instead of 4% of children that had diarrhea with blood, me I get 17.82% and I cannot replicate column two at all) and rows; source of drinking water & toilet facility?
3. How do i retrieve the number of children? is there a command i can add?
4. How in stata can one produce a table like 10.1 or 10.7?

Thanks

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Wed, 18 Nov 2015 06:17:43 GMT

View Forum Message <> Reply to Message

See the below code for examples for question 1 and 2.
cd "C:\Data\DHS_Stata\"
* Use kids recode dataset
use "UGKR60FL.DTA", clear

* set weight
g wt=v005/1000000

* Q1.

* create mother's age at birth variable =
* CMC date of birth of child - CMC date of birth of woman - , divided into 15 year groups (180 months)
g mageatb = b3 - v011
recode mageatb (0/239 = 1 "<20") (240/419 = 2 "20-34") (420/599 = 3 "35-49"), g(mab)
lab var mab "Mother's age at birth"
tab mab [iw=wt]

* creates smokes variable
g smokes = 2
replace smokes = 1 if v463a == 1 | v463b == 1
*replace smokes = 9 if smokes == 2 & (v463a == 9 | v463b == 9)
lab def smokes 1 "Smokes cigarettes/tobacco" 2 "Does not smoke"
lab var smokes "Mother's smoking status"
lab val smokes smokes
tab smokes [iw=wt]Note that there appears to be a typing error in table 10.1 and the number of

women who do not smoke should be 8010, not 8000.

* Q2.

* Diarrhea
g diarrhea = (h11==2) if b5==1
tab diarrhea [iw=wt] if b5==1
* Diarrhea with blood
g diarr_bld = (h11b==1) if b5==1
tab diarr_bld [iw=wt] if b5==1
For question 3, the tab commands above provide the numbers of children as well as the percentages.
For question 4, a simple way of producing the output you need is to try something like the following:
* Q4.

g total = 0
foreach var in v102 v101 v190 total {
  tab `var' diarrhea [iw=wt] if b5==1, row
}
This produces results for a number of background characteristics, but not a single table.
If you want to produce tables that looks like those in the report, you can try using alternative tabulation commands, such as tabout, or use a different software to produce the output desired (DHS uses CSPro with some additional tools to produce the tables as found in the reports).

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Wed, 18 Nov 2015 12:41:42 GMT
View Forum Message <> Reply to Message

Thanks dear Trevor, let me try it out. Will get back.

Gabriel

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Wed, 18 Nov 2015 23:47:45 GMT
View Forum Message <> Reply to Message

Dear Trevor,

Just a quick one. All worked! Thanks a lot. A few questions;

1. How do you gen the children numbers in the middle of the table other than the totals? I could get totals by your guidance.

2. You did tell me how rows; Source of drinking water and Toilet facilities were classified as improved & not improved. I tried to categorize according to table 2.1 & 2.2 but I don't exactly replicate the values for these rows

Thanks

Gabriel

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Thu, 19 Nov 2015 00:05:58 GMT
View Forum Message <> Reply to Message

1. The number in the middle of a table are just produced by selecting the data. For example,
g weighed = (m19a==1 | m19a==2)
tab weighed [iw=wt]
provides the numbers of births weighed at birth given in table 10.1 (5th column). [The percents are the same as those in the 1st column]

2. Source of drinking water and toilet facilities are produced by merging the household recode (HR) dataset with the kid's recode (KR) dataset and recoding the source of drinking water and toilet facility variables according to the rules in table 2.1 and 2.2. These are not based just on v113 and v116 as these variables exclude non de jure (non usual) residents. Instead the variables from the household data are used.

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Thu, 19 Nov 2015 14:32:14 GMT
View Forum Message <> Reply to Message

Thanks dear Trevor.

Kind regards,

Gabriel

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Fri, 20 Nov 2015 01:07:53 GMT
View Forum Message <> Reply to Message

hello,

---

1. Does replicating table 11.1 (nutritional status of children) require merging KR file with any other? I need help here because I have tried to generate variables from Height for weight but i am not replicating at all.
2. I also tried to merge KR to HR by; generating v001=hv001, v002=hv002 in HR file1
 then use KR file, sort v001 v002
 then merge m:1 v001 v002 using HR file1
I am not sure this is right. I was trying to look for hv103 which seem to have been used to generate table 11.1 and is in HR file. thanks

Gabriel

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Fri, 20 Nov 2015 04:06:04 GMT
View Forum Message <> Reply to Message

The anthropometric data are collected in the household questionnaire and the data used for table 11.1 can be found in the PR file. The KR file only includes the data for children whose mothers were interviewed and not for all children in the household. With the PR file you will need to select the de facto children (children who slept in the household the previous night, using hv103.

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Sat, 21 Nov 2015 10:55:37 GMT
View Forum Message <> Reply to Message

Hello, how do you get the mean Z-score(SD) say under height for age?

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Sat, 21 Nov 2015 20:13:18 GMT
View Forum Message <> Reply to Message

svy: mean hc70 if hc70<9996 & hv103==1
then divide the resulting mean by 100 to get the decimal in the right place.

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Sun, 22 Nov 2015 02:12:34 GMT
View Forum Message <> Reply to Message

Okay, but this gives you one value (-1.468763) and the column for "Mean Z-score (SD)" has many values! For example under age in months, the the column has (-0.3, -0.4, -1.0, -1.4, -1.7, -1.8, -1.7 & -1.5). How do i get these? thanks

---

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Sun, 22 Nov 2015 22:59:26 GMT
View Forum Message <> Reply to Message

So you disaggregate by age in months:

recode hc1 (0/5=1 "<6")(6/8=2 "6-8")(9/11=3 "9-11")(12/17=4 "12-17")(18/23=5 "18-23")(24/35=6 "24-35")(36/47=7 "36-47")(48/59=8 "48-59")if hv120==1, gen (age_months)
svy: mean hc70 if hc70<9996 & hv103==1, over(age_months)

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Mon, 23 Nov 2015 14:22:30 GMT
View Forum Message <> Reply to Message

Great! Thanks Dear Trevor,

I can see values are slightly higher (1= -.4374683, 2= -.4934267, 3= -1.127726, 4= -1.456142,5= -1.672893, 6= -1.802239, 7= -1.752084, 8= -1.604335) is this okay or there is something else I have to fix? Thanks

---

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Mon, 23 Nov 2015 15:50:23 GMT
View Forum Message <> Reply to Message

I'm getting different results from you:
. svy: mean hc70 if hc70<9996 & hv103==1, over(age_months)
(running mean on estimation sample)

Survey: Mean estimation

Number of strata =   19   Number of obs   =   2,336
Number of PSUs   =   397   Population size = 2,350.3319
                     Design df   =   378

    _subpop_1: age_months = <6
    _subpop_2: age_months = 6-8
    _subpop_3: age_months = 9-11
    _subpop_4: age_months = 12-17
    _subpop_5: age_months = 18-23
    _subpop_6: age_months = 24-35
    _subpop_7: age_months = 36-47
    _subpop_8: age_months = 48-59

```
          |          Linearized
     Over |    Mean   Std. Err.    [95% Conf. Interval]
-------------+----------------------------------------------
hc70      |
  _subpop_1 |  -34.44011   12.39171     -58.80543   -10.07478
  _subpop_2 |  -43.11069   14.73607     -72.08562   -14.13575
  _subpop_3 |  -101.8908   14.76032     -130.9134   -72.86815
  _subpop_4 |  -140.5865   11.14251     -162.4955   -118.6774
  _subpop_5 |   -166.106   14.58119     -194.7764   -137.4356
  _subpop_6 |  -177.6505   7.966516     -193.3148   -161.9863
  _subpop_7 |  -174.0726   7.855612     -189.5188   -158.6264
  _subpop_8 |  -146.6423   9.608246     -165.5346    -127.75
----------------------------------------------------------------
```

If you divide the results here by 100, you get the same results as in table 11.1 of the Uganda report. I think you must have a small mistake somewhere. Are you weighting with the correct variable?

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Mon, 23 Nov 2015 15:56:59 GMT
View Forum Message <> Reply to Message

I can see these are exact! I am using gen wgt=hv005/1000000 as the weight

---

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Mon, 23 Nov 2015 15:58:34 GMT
View Forum Message <> Reply to Message

That looks correct. And you are using wgt in the svyset command too?

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Mon, 23 Nov 2015 16:00:34 GMT
View Forum Message <> Reply to Message

recode hc1 (0/5=1 "<6")(6/8=2 "6-8")(9/11=3 "9-11")(12/17=4 "12-17")(18/23=5 "18-23")(24/35=6 "24-35")(36/47=7 "36-47")(48/59=8 "48-59")if hv120==1 & hv103==1, gen (age_months2)
svy: tab age_months2 hbel_minus3 if survey==1 & hv103==1 & hc70<9996, row
svy: tab age_months2 hbel_minus2 if survey==1 & hv103==1 & hc70<9996, row
svy: mean hfa if hc70<9996 & hv103==1 & hc13==0, over(age_months2)
svy: mean hfa if hc70<9996 & hv103==1 & hc13==0, over(hc27)
tab age_months2 hbel_minus2 if survey==1 & hv103==1 & hc70<9996 [iw=wgt]

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Mon, 23 Nov 2015 16:05:45 GMT
View Forum Message <> Reply to Message

Sorry I sent you something I was changing but this is what I meant to send you;

```
clear
use "C:\Users\Gabriel\Desktop\DHS\DHS 2011\data\UGPR6052FL.DTA"
log using anthropometric.log, replace

gen wgt=hv005/1000000

egen strata = group(hv024 hv025)
replace hv022 = strata
replace hv023 = strata
svyset[pw=wgt], psu(hv021) strata(hv022) singleunit(centered)

*Height for age
gen hbel_minus3=0
replace hbel_minus3=1 if hc70<-300 & hc13==0
gen hbel_minus2=0
replace hbel_minus2=1 if hc70<-200 & hc13==0


recode hc1 (0/5=1 "<6")(6/8=2 "6-8")(9/11=3 "9-11")(12/17=4 "12-17")(18/23=5 "18-23")(24/35=6
"24-35")(36/47=7 "36-47")(48/59=8 "48-59")if hv120==1 & hv103==1, gen (age_months2)
svy: tab age_months2 hbel_minus3 if survey==1 & hv103==1 & hc70<9996, row
svy: tab age_months2 hbel_minus2 if survey==1 & hv103==1 & hc70<9996, row
svy: mean hc70 if hc70<9996 & hc13==0, over(age_months2)
svy: mean hc70
 if hc70<9996 & hc13==0, over(hc27)
tab age_months2 hbel_minus2 if survey==1 & hv103==1 & hc70<9996 [iw=wgt]

log close
```

Sorry this is what I am running in stata

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Mon, 23 Nov 2015 16:16:04 GMT
View Forum Message <> Reply to Message

Ooohh, I saw the problem dear Trevor,

i have two surveys in a single data base so I should have indicated survey==1 on generating
age-groups
recode hc1 (0/5=1 "<6")(6/8=2 "6-8")(9/11=3 "9-11")(12/17=4 "12-17")(18/23=5 "18-23")(24/35=6
"24-35")(36/47=7 "36-47")(48/59=8 "48-59")if hv120==1 & survey==1, gen (age_months2)

Thanks a lot.

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Mon, 23 Nov 2015 16:19:40 GMT
View Forum Message <> Reply to Message

Great - glad you worked it out

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Thu, 26 Nov 2015 22:46:55 GMT
View Forum Message <> Reply to Message

Hello, I am still trying to replicate table 11.1 Nutritional status of children. I have a few more questions;
1. Which variables in data file UGPR60FL.DTA are used to generate;
 a) Size at birth
 b) Mother's interview status
 c)Mother's nutritional status
2. How do you generate "first birth"-under "birth interval in months" using hc63 (preceding birth intervals(months))

Thanks

Gabriel

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Thu, 03 Dec 2015 16:55:56 GMT
View Forum Message <> Reply to Message

1a) M18 in the KR file contains the size at birth variable. For size at birth you first need to merge UGKR60FL.DTA to UGPR60FL.DTA. First sort the PR file on hv001 hv002 and hvidx, and the KR file on v001 v002 and b16. Drop cases from the KR file that have b16 == 0 or 99 or missing (.) as these are children that do not live in the household, died, or had other missing data. Then perform a 1:1 merge.

1b) For mother's interview status, you need to look at hv111 and hv112, and then for cases where hv112 is not 0, 99, or missing (.) you need to merge the IR file to see if the mother was successfully interviewed. Use hv001 hv002 and hv112 from the PR file and v001 v002 and v003 from the IR file for the merge. This should be a many to one (m:1) merge from the PR file using the IR file as the mother may have had more than one child. Note that you will need to sort on the

above fields first before merging.

1c) For mother's nutritional status, you need to use HA40 for the mother. This get's trickier as you already have an ha40 in the child's record, but it will be blank. You will need to drop this variable, and then merge in ha40 for the mother from another copy of the PR file. To do this, you will need:
 your PR file sort based on hv001 hv002 and hv112, and with ha40 dropped. gen mother=hv112
a copy of the original PR file sorted on hv001 hv002 and hvidx. gen mother=hvidx
You will then merge 1:1 using hv001 hv002 and mother, and just add ha40 from the copy of the PR file to your main PR file.

2) For the preceding birth interval, once you have done the merge of the KR file in the first step above, then just use B11. B11 is set to missing for first births, so just recode the durations into the groups you need and the missing (.) to the category for first birth.

Subject: Re: Replicating DHS tables on child health
Posted by Trevor-DHS on Thu, 03 Dec 2015 16:58:26 GMT
View Forum Message <> Reply to Message

Added note: There will be cases that do not match for 1a) and 2) as the child listed in the household was not a child of an interviewed woman. For those cases, size at birth and preceding birth interval should be set to missing (.)

Subject: Re: Replicating DHS tables on child health
Posted by lexgw on Thu, 10 Dec 2015 13:35:09 GMT
View Forum Message <> Reply to Message

Dear Trevor,

Hope this finds you well. I have been successful on Qn. 1 a) Birth size (replicated all) and partly on Qn.2 Birth interval (only values for "missing" do not match. I have 330 obs for missing instead of 331 shown in the report).

Qn. 1b) which variable is used for mother's interview status? after merging IR, PR & KR it is confusing!(can't see one with interviewed, not interviewed but in house etc..).
Qn. 1c) Why drop ha40 and again merge? from 1a)we have PR & KR files and ha40 exists and not empty. However it seems like this BMI/ha40 is for women not interviewed because when you tab ha40 v015, there are no obs, tab ha40 hbel_minus3 still don't link up!

thanks for the continued support.

Gabriel

Subject: Re: Replicating DHS tables on child health

Posted by lexgw on Thu, 10 Dec 2015 14:23:38 GMT

See do file attached

File Attachments
1) child nutrition.do, downloaded 984 times

Subject: Re: Merging
Posted by lexgw on Thu, 04 Aug 2016 17:36:39 GMT

Hello there,

I would like to analyse the likelihood of a child experiencing a working mother. This in a way implies I only need mothers with child/ren. But in the analysis I also need household characteristics. Does this mean merging the three data sets KR IR and PR or IR and PR are enough?

Thanks

Subject: Re: Merging
Posted by Trevor-DHS on Fri, 05 Aug 2016 22:13:40 GMT

The IR data is already merged into the KR file, so all of the women's characteristics are attached to the children's data. You would need to just match the household characteristics, which you can do either using the HR (Household Recode) or the PR (Person's Recode - matching to the mother's line number).

Subject: Re: Merging
Posted by lexgw on Sat, 06 Aug 2016 00:25:13 GMT

Thanks dear Trevor

Subject: Re: Replicating DHS tables on child health
Posted by sammy nyarko on Sat, 10 Dec 2016 15:59:24 GMT

im trying to follow your steps but got stuck. After the command below.

---

recode hc1 (0/5=1 "<6")(6/8=2 "6-8")(9/11=3 "9-11")(12/17=4 "12-17")(18/23=5 "18-23")(24/35=6 "24-35")(36/47=7 "36-47")(48/59=8 "48-59")if hv120==1 & survey==1, gen (age_months2)

this was my responds

variable survey not

Subject: Re: Replicating DHS tables on child health
Posted by sammy nyarko on Sun, 11 Dec 2016 08:47:40 GMT
View Forum Message <> Reply to Message

okay i get it now. you were working with two survey data.