Subject: Daughter age at circumcision Mali 2006
Posted by aligazan on Thu, 13 Nov 2014 10:31:40 GMT
View Forum Message <> Reply to Message

I am working with the Mali 2006 Dataset.

The variable S915 gives the daughters age at the time of circumcision. This as I understand it from the questionnaire is given in 3 digits, the first being the year, and the second two being the month.

The range of data is from 100 - 995. I understood 100 to mean 1 year and no months and that this means that no girls were circumcised younger than 1 year.

However, in the DHS report for Mali 2006 table 18.6 'Âge des filles à l'excision' the data and range of ages shown does not seem to match to the information within variable S915 in the dataset.

Would you be able to explain to me how the data in table 18.6 is calculated?

Many thanks

Janet

Subject: Re: Daughter age at circumcision Mali 2006
Posted by Trevor-DHS on Wed, 19 Nov 2014 19:11:51 GMT
View Forum Message <> Reply to Message

Variable S915 is a composite variable with the first digit being a unit field and the last two digits being a number field. The easiest way to handle this is to split it into two parts. You can use the following code in Stata to do that:

```
gen s915u = int(s915/100)
gen s915n = mod(s915,100)
lab var s915u "Age of daughter at circumcision (units)"
lab var s915n "Age of daughter at circumcision (number)"
lab def s915u 1 "Age in years" 2 "In infancy" 3 "Age in months" 9 "Special answers"
lab def s915n 95 "In infancy" 98 "Don't know" 99 "Missing"
lab val s915u s915u
lab val s915n s915n
tab s915n s915u [iw=v005/1000000]
```

However when I looked at the data more closely and compared to the Mali 2006 report, it is apparent that there are errors in the coding and an error in the tabulation.

First the error in the coding. The unit code 2 (In infancy) in s915u should match up with code 95 (In infancy) in s915n, but there are cases with code 2 on s915u that have values 1-40 in s915n. I believe these should be code 3 on s915u for age in months, not code 2 (In infancy). Also, code 9 (Special answers) on s915u has numeric values 1-18 on s915n. These are probably keyer errors and could be either age in years or months, but I believe are most likely to be age in years (it is

possible they are some of each - I don't think we can tell).

Second the error in the tabulation. When I tried to reproduce table 18.6 it was clear that the table just used s915n from above, assuming all ages were age in years, and ignoring that some were age in months.

Below is code that will first allow you to reproduce the table and then a correction for the age in months:

```
* Create age groups used in table 18.6 in Mali 2006 report
recode s915n (95=0)(0/1=1)(2/4=2)(5/9=3)(10/14=4)(15/94=5)(98/99=9), gen(age_exc)

* Add labels
lab var age_exc "Age of daughter at circumcision"
label def age_exc 0 "In infancy" 1 "0-1" 2 "2-4" 3 "5-9" 4 "10-14" 5 "15+" 9 "Don't know/missing"
label val age_exc age_exc

* Tabulate as in table 18.6 in Mali 2006 report
tab age_exc [iw=v005/1000000]

* Correction for reporting as age in months
replace age_exc = 1 if  s915n< 24 & (s915u==2|s915u==3)
replace age_exc = 2 if (s915n>=24 & s915n<48) & (s915u==2|s915u==3)

* Re-tabulate withh correction
tab age_exc [iw=v005/1000000]
```

Subject: Re: Daughter age at circumcision Mali 2006
Posted by aligazan on Thu, 15 Jan 2015 16:17:37 GMT
View Forum Message <> Reply to Message

Hi Trevor,

Many thanks for this thorough response.

I noticed today that a new version of the Mali dataset has been released with this variable added/corrected which is great - especially as I had not quite yet got on to following your instructions myself.

I am working on it now and will be following your first steps of splitting the unit and number field. As I am using SPSS not Stata I will try to figure out how to do this using your Stata instructions - I may be in touch again if my SPSS syntax skills are not up to the job.

For future reference is there any way of knowing that a variable is a composite variable in the way that you describe? I have noticed that ages are recorded in different ways.

Once again, many thanks for your help.

Janet

Subject: Re: Daughter age at circumcision Mali 2006
Posted by Trevor-DHS on Fri, 16 Jan 2015 13:13:51 GMT
View Forum Message <> Reply to Message

I recommend looking at the DHS recode manual which describes the standard variables and explains the composite variables where they are used. For country specific variables, look at the questionnaire and you should be able to see if a variable is formed from a unit and a number in that unit. Producing a frequency distribution of the variable will also give you an idea too.

Sorry that the code i provide is in Stata - so many users use Stata that I generally respond with Stata code. Hopefully it is clear enough for you to translate it into SPSS easily enough.

Subject: Re: Daughter age at circumcision Mali 2006
Posted by aligazan on Fri, 16 Jan 2015 18:44:58 GMT
View Forum Message <> Reply to Message

Thanks again, and yes I managed to work out the SPSS syntax from your stata example.

One last question (for the time being) concerns variable G109 in the Mali 2006 women's dataset which is "line number of daughter most recently circumcised".

Please can you tell me which variable this relates to and in which dataset. I think it is BORD$01 etc. but would be grateful if you could confirm, as my analysis depends upon this.

For example CASEID 1 9 2, G109 tells us the line number of daughter most recently circumcised is 4, and BORD$04 tells us this was the first born child to this woman, and then B4$01 tells us this child is female. Is that correct? (If so, then the last born girl is not always the last daughter circumcised which is unexpected).

Many thanks

Janet

Subject: Re: Daughter age at circumcision Mali 2006
Posted by Trevor-DHS on Sat, 17 Jan 2015 11:54:40 GMT
View Forum Message <> Reply to Message

You are right to look at BORD$01 ..., but you need to look at the values of these to determine which child it refers to. If you look at BORD$01 it has the value 04, matching the line number in G109. The birth history is in reverse order in the recode file, so the $01 variables are for the last birth, the $02 variables for the next to last, etc. In the vast majority of cases you should find that the value matches BORD$01.

Subject: Re: Daughter age at circumcision Mali 2006
Posted by aligazan on Tue, 20 Jan 2015 03:42:06 GMT
View Forum Message <> Reply to Message

Hello Trevor

I have been looking at variable G109 within this dataset with an interest in whether the last daughter who was circumcised was alive or dead at the time of the survey.

Using the logic: "Look up the BORD$column indicated by G109 to get the column number of the last daughter (=CLLD)" and then look up the B5$column indicated by CLLD - I get 100% of the last daughters to be circumcised as alive. I also get 100% as female (which makes sense and suggests the column logic is working).

This is unexpected as about 10-15% of children are reported as dead in columns B5.

Is there something within the data collection that means all "last daughters circumcised" are by definition alive at the time of the survey? I can't see anything in the wording of the questionnaire which suggests this (Q911).

Excerpts from the SPSS syntax I've used is below (it's very clunky - I'm sure there are more efficient ways of writing this but I haven't figured it out yet!):

```
DO IF  (G109=1).
RECODE BORD$01 (ELSE=Copy) INTO GavCOLNOLAST.
END IF.
VARIABLE LABELS  GavCOLNOLAST 'Column no of last circumcised'.
EXECUTE.

DO IF  (G109=2).
RECODE BORD$02 (ELSE=Copy) INTO GavCOLNOLAST.
END IF.
VARIABLE LABELS  GavCOLNOLAST 'Column no of last circumcised'.
EXECUTE.
```

etc to 15

```
DO IF  (GAVCOLNOLAST=1).
RECODE B5$01 (ELSE=Copy) INTO Mortalitylast.
END IF.
```

```
VARIABLE LABELS  Mortalitylast 'Mortality of last circumcised'.
EXECUTE.

DO IF  (GAVCOLNOLAST=2).
RECODE B5$02 (ELSE=Copy) INTO Mortalitylast.
END IF.
VARIABLE LABELS  Mortalitylast 'Mortality of last circumcised'.
EXECUTE.
```

etc. to 15

The result is that all Mortalitylast cells are either blank (where G109 is 99 or blank) or 1 (indicating daughter is alive).

Your help would be very much appreciated.

Best wishes,

Janet

Subject: Re: Daughter age at circumcision Mali 2006
Posted by Trevor-DHS on Tue, 20 Jan 2015 07:03:17 GMT
View Forum Message <> Reply to Message

Yes, that seems to be the case in the data. All of the daughters are alive. It is not very explicit in the questionnaire, but if you look at the filter in 909 in the questionnaire, there is a filter so that the questions are only asked if there is at least one living daughter. Additionally, the question asks about daughters, implying they are alive. The interviewers instructions were probably explicit in regard to this, but I think it is clear that these questions were only asked about living daughters.

I also tested to check with the following code, which is a bit shorter than your and may be useful to you:

```
compute bindex = $sysmis.
compute bstatus = $sysmis.
do repeat bord=bord$01 to bord$15 /status=b5$01 to b5$15 /idx=1 to 15 .
+  do if (bord = g109).
+   compute bindex=idx.
+   compute bstatus = status.
+  end if.
end repeat.
freq bstatus bindex.
```

Subject: Re: Daughter age at circumcision Mali 2006
Posted by aligazan on Tue, 20 Jan 2015 16:35:40 GMT

View Forum Message <> Reply to Message

That's a shame the data was collected in that way. However, good to know I'm not missing something and the syntax is very useful. Once again, many thanks for your help.

Best wishes, Janet

Subject: Re: Daughter age at circumcision Mali 2006
Posted by aligazan on Fri, 06 Feb 2015 17:18:00 GMT
View Forum Message <> Reply to Message

Dear Trevor

I have a new challenge that I am struggling with. I am still working with the Mali 5 dataset and using SPSS.

I would like to bring over some of the data from the Womans IR file into the Male MR file. In particular I would like to bring over the variables concerning the circumcision status of each man's wife/wives/partner(s) e.g. G101, G102, G103, G104

I don't know if I need to merge the IR and MR datasets (and can't even imagine what that would end up looking like!) or if I can just somehow copy the relevant data from the IR into the MR using the line number of wife variable MV034$1 etc.

I would be so grateful if you could explain how to do this.

Many thanks

Janet

Subject: Re: Daughter age at circumcision Mali 2006
Posted by Trevor-DHS on Mon, 09 Feb 2015 16:18:01 GMT
View Forum Message <> Reply to Message

Have you tried working with the Couple's Recode (CR) file. In this file the men's and women's recode data has already been merged together.

Alternatively you can use the MATCH command. See this thread for an example of the MATCH FILES command. You would match using something like:

```
* Open the women's data.
GET FILE='ccIRvvFL.SAV'.
* Match the men's data to the women's data.
MATCH FILES /FILE=*
  /table='C:\data\ccMRvvFL.SAV'
```

```
/RENAME MV001=V001 MV002=V002 MV003=V034
/BY V001 V002 V034.
```