

---

Subject: Problem with dates in the Ethiopia datasets  
Posted by [DHS user](#) on Wed, 20 Feb 2013 16:19:12 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

The variables in the Ethiopia datasets, related to the year of the study do not match the actual timing of the study. Is this an error?

---

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Bridgette-DHS](#) on Wed, 20 Feb 2013 16:28:58 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

All standard date variables in the Ethiopia data file are in the Ethiopian calendar. In general the Ethiopian calendar is 92 months behind the Gregorian (western) calendar. The Ethiopian year runs from September 11 through September 10 and has 12 months of 30 days and 1 months of 5 days. Thus it is only possible to transfer the Ethiopian calendar into the Gregorian calendar when the exact date (day, month and year) is available. Therefore, in the file you will find additional variables indicating the Gregorian dates if they could be established.

Note that all standard variables based on calendar dates and century month codes are given in the Ethiopian Calendar. In general, the Ethiopian year consists of 365 days, divided into 12 months of 30 days and one month of 5 days (6 days in a leap year). Ethiopia's new year falls on September 11 and ends the following September 10 according to the Gregorian calendar. From September 11 to December 31, the Ethiopian year runs seven years behind the Gregorian year, thereafter, the difference is eight years. Since the exact day is not available for most dates, it is not possible to convert the dates exactly, but only approximate it to the month. There is a difference of 92 months between the two date systems. The 13th month of the Ethiopian calendar falls in September. Thus, to keep the higher precision available in the Ethiopian calendar, these were used for all standard recode variables where applicable. In general, dates in the Gregorian calendar are provided as country specific variables. However, the calendar is transferred to the Gregorian calendar.

To explain somewhat further, the number of months in the CMC variables (even for Ethiopian dates) are computed as follows: each year has 12 months, but when the event was in the 13th month then 13 is added for the number of months. Thus e.g. a person was born in the 13th month of 1960 in the Ethiopian calendar, then the CMC is  $60 * 12 + 13$ . Thus the 5 days in the 13th month really go to the following year when you recomputed the CMC back to years (with 12 months). But this will cause little bias. So just use the CMC as you would for other surveys (intervals should not be divided by 13, but by 12). However, know that on average the Ethiopian calendar lags 92 months behind our Gregorian system.

The Ethiopian calendar was kept as the standard variables, since not all dates (e.g. vaccination dates) could be transferred.

Here is a link that you can use to convert Ethiopian dates:

<http://www.funaba.org/en/calendar-conversion.cgi>

Attached is a little Stata script: "ETcalconvert", that you may find useful (written by DHS Data User: Keith Kranker).

I hope this helps.

Bridgette-DHS

---

### File Attachments

1) [ETcalconvert.ado](#), downloaded 2424 times

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [lillo?S](#) on Fri, 22 Jan 2016 17:02:38 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hello,

From what I can see in the DHS 2005 there are country-specific variables that define the dates in the Gregorian calendar, the same does not hold for DHS 2011.

In the 'Individual Recode Documentation' (p.8) I find this:

"Before the production of any indicators with these data the Ethiopian calendar was converted to the Gregorian calendar but conserving the Ethiopian year; however, the Ethiopian first month is considered in the logic as January, the 2nd as February, etc. For dates including year, month and day the conversion is precise since both calendars have 365 or 366 days; for dates including only year and month, the 13th month was included in December."

I understand that the months have already been converted, whereas the years have not. When I go to the 'CSPRO Process Summary' file and look at v006, the months of interview range from 4 through 9. According what's written in the IRD word file, I should interpret this as 4=April and 9=September because the conversion is supposed to have been already done. However, in the website at <http://dhsprogram.com/what-we-do/survey/survey-display-359.cfm> I see that: Fieldwork: December 2010 - May 2011. These months do not match the ones above, whereas they could coincide if the conversion had not been done.

My questions are:

- 1- has the conversion been done or not?
- 2- how are the CMC variables given?

Thanking you in advance.

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [lillo?S](#) on Fri, 22 Jan 2016 17:04:21 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hello,

From what I can see in the DHS 2005 there are country-specific variables that define the dates in the Gregorian calendar, the same does not hold for DHS 2011.

In the 'Individual Recode Documentation' (p.8) I find this:

"Before the production of any indicators with these data the Ethiopian calendar was converted to the Gregorian calendar but conserving the Ethiopian year; however, the Ethiopian first month is considered in the logic as January, the 2nd as February, etc. For dates including year, month and day the conversion is precise since both calendars have 365 or 366 days; for dates including only year and month, the 13th month was included in December."

I understand that the months have already been converted, whereas the years have not.

When I go to the 'CSPRO Process Summary' file and look at v006, the months of interview range from 4 through 9. According what's written in the IRD word file, I should interpret this as 4=April and 9=September because the conversion is supposed to have already been done. However, in the website at <http://dhsprogram.com/what-we-do/survey/survey-display-359.cfm> I see that: Fieldwork: December 2010 - May 2011. These months do not match the ones above, whereas they could coincide if the conversion had not been done.

My questions are:

- 1- has the conversion been done or not?
- 2- how are the CMC variables given?

Thanking you in advance.

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Bridgette-DHS](#) on Fri, 29 Jan 2016 17:56:35 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

In the 2011 survey, all dates are in the Ethiopian calendar rather than the Gregorian calendar. The cmc is also calculated consistently with the Ethiopian calendar. For example, you can confirm that  $v008 = 12(v007 - 1900) + v006$ . This consistency is found in all surveys, whether using the Gregorian, Ethiopian, or Nepalese calendars.

To convert to the Gregorian calendar, add 92 to the Ethiopian cmc. Then get the Ethiopian year and month from the Ethiopian cmc as follows, illustrated again for v006, v007, and v008:

```
gen v007=int((v008-1)/12)
gen v006=v008-12*v007
```

replace v007=v007+1900

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [st89](#) on Fri, 19 Feb 2016 15:12:52 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

A follow-up question about the dates in the 2011 survey (specifically the children's recode): variable h9d (day of month of measles vaccination, if known) takes on values from 1 to 31. However, my understanding is that all of the dates are in the Ethiopian calendar, which should have 12 months of 30 days + one month of 5 or 6 days. I'm unclear on how to use h9d values of 31 since it is not in the Ethiopian calendar, and am thus not able to convert the three variables h9y-h9m-h9d to Gregorian dates.

Thanks for your help!

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Bridgette-DHS](#) on Fri, 19 Feb 2016 17:02:36 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Another response from Tom Pullum:

I will start by repeating an answer I gave on the forum on January 29, but this time with a correction of a typo in the January 29, version, where in the second paragraph I said "Ethiopian" but meant "Gregorian"--sorry about that.

In the 2011 survey, all dates are in the Ethiopian calendar rather than the Gregorian calendar. The cmc is also calculated consistently with the Ethiopian calendar. For example, you can confirm that  $v008 = 12(v007 - 1900) + v006$ . This consistency is found in all surveys, whether using the Gregorian, Ethiopian, or Nepalese calendars.

To convert to the Gregorian calendar, add 92 to the Ethiopian cmc. Then get the Gregorian year and month from the Gregorian cmc as follows, illustrated again for v006, v007, and v008. I will use vg006, etc, to indicate that these are Gregorian versions of the original Ethiopian codes.

```
gen vg008=v008+92
gen vg007=int((vg008-1)/12)
gen vg006=vg008-12*v007
replace vg007=vg007+1900
```

Now to get to your question. Although you may be right that there is a 5-day month in the Ethiopian calendar, I think it must be absorbed in one or both of the adjacent months. If you go to the BR file and enter "tab h9m" or "tab b2", etc., you will see 12 numbered months, all with about the same number of cases. I recommend that you simply do the month and year conversion described above, but keep the day as h9d. I would look at how the days 29, 30, and 31 are

converted to a day with "day=mdy(h9m,h9d,h9y)". You will find that certain combinations of year and month and day (even year, because of Leap Year) are rejected, and will produce "day=" because the specified Gregorian month has only 30 days, or in the case of February has only 28 days, except in a Leap Year, when it has 29! (Ten days from now we will have a Feb. 29!) Then I personally would recode h9d for all of the rejected values to 28, but you could go to more trouble and successively convert to 30, then to 29, then to 28, until all combinations of year, month, and day are accepted. The kind of error that will be incurred with this sort of adjustment will be trivial, especially compared with other reporting errors.

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [kcaglaya@tulane.edu](mailto:kcaglaya@tulane.edu) on Fri, 17 Nov 2017 17:39:50 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hi,

I use b1 (month of birth) and hw16 (day of birth of child) in my analysis and initially assumed that both of these information are provided according to the Ethiopian calendar. However, as mentioned before in the forum; hw16 takes the values from 1 to 31 in 2011 and 2016 Surveys. (In 2000 and 2005 max value for hw16 is 30) I looked at the months in which hw16 is equal to 31 and this is how it looks like:

. tab b1 if hw16==31

month of   birth	Freq.	Percent	Cum.
1	37	24.03	24.03
3	21	13.64	37.66
5	25	16.23	53.90
7	19	12.34	66.23
8	30	19.48	85.71
10	16	10.39	96.10
12	6	3.90	100.00
Total	154	100.00	

These are the Gregorian months with 31 days. Does this mean that b1 (month of birth) is also in Gregorian calendar? Also see the months with 30 and 31 days:

. tab b1 if hw16==31 | hw16==30

month of   birth	Freq.	Percent	Cum.
1	73	14.07	14.07

3	52	10.02	24.08
4	32	6.17	30.25
5	67	12.91	43.16
6	38	7.32	50.48
7	72	13.87	64.35
8	59	11.37	75.72
9	27	5.20	80.92
10	52	10.02	90.94
11	26	5.01	95.95
12	21	4.05	100.00

-----+-----  
Total |      519      100.00

It lacks February as we expect from the Gregorian Calendar.

My conversion from Ethiopian to Gregorian calendar depends on b1. I would greatly appreciate a clarification regarding the variable b1.

Regards,

Koray Caglayan

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Bridgette-DHS](#) on Mon, 18 Dec 2017 17:26:42 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior Data Processing Specialist, Ladys Ortiz:

#### Birth Month and Day Variables in DHS Ethiopian Household Survey

Your Comments: We use B1 (month of birth) and HW16 (day of birth of child) in our analysis and initially assumed that both of these variables are reported according to the Ethiopian calendar. However, HW16 takes values from 1 to 31 only in 2011 and 2016 Surveys. (In 2000 and 2005 the max value for HW16 is 30) We looked at the months in which HW16 is equal to 31 and this is what it looks like:

Response:

A) FYI: For the date of birth of the child you should use DD=B17/MM=B1/YY=B2; variable HW16 is for children with Height/weight.

B) One of the differences between EDHS 2016 and previous DHS's is that the DAY of birth of the child was included in the birth history (DD/MM/YYYY) instead of MM/YYYY, as in prior surveys. Because of that we could convert the day and the month of birth to Gregorian calendar but the year remains as Ethiopian calendar.

C) You will find more information about how Ethiopian's dates were processed for previous surveys by checking the recode documentation included in our website.

Hope my response helps in answering your questions.

Ladys

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [kcaglaya@tulane.edu](mailto:kcaglaya@tulane.edu) on Mon, 18 Dec 2017 18:21:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Dear Ladys,

Thank you so much for your response. It helps a lot.

There is one more thing that I would like to be sure of, since the results of our analysis heavily depend on the day of birth.

1) b17 is identical with hw16 in 2016 survey. We are using hw16 since it is present in the data across all survey years. We pool the data from all survey years in our estimations. Height and weight are among our variables of interest, so we are using observations with height and weight variables available anyway.

2) The information you provided about 2016 survey in terms of birth day and birth month is very helpful and answers our question. However, we suspect that this might also be the case for 2011 survey; i.e. hw16 and b1 are also reported in Gregorian Calendar in 2011 survey. I base this on the fact that in 2011 survey data; hw16 varies between 1-31, the birth months with hw16=31 are the Gregorian months with 31 days and also birth months with hw16=30 or hw16=31 lack February as we would expect from Gregorian Calendar. So, I would greatly appreciate if you can confirm whether the following table regarding the way these dates are reported is correct or not. I also attach the table as a separate file.

	B17 (Day of Birth)	HW16 (Day of Birth)	B1 (Month of Birth)	B2 (Year of Birth)
2000 Survey	-	Ethiopian	Ethiopian	Ethiopian
2005 Survey	-	Ethiopian	Ethiopian	Ethiopian
2011 Survey	-	Gregorian	Gregorian	Ethiopian
2016 Survey	Gregorian	Gregorian	Gregorian	Ethiopian

Once again, thank you for your time and help.

Regards,

Koray

---

#### File Attachments

1) [dhs\\_ethiopia\\_dates.docx](#), downloaded 1032 times

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Bridgette-DHS](#) on Mon, 18 Dec 2017 22:15:40 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Another response from Ladys Ortiz:

Dear Koray,

I forwarded your email to the Data Processing Specialist who processed EDHS-2005, so he can address your questions. Hopefully, it will be sooner than later (there could be delays as many of us are already on vacation for the holidays).

Thanks,

Ladys Ortiz

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Bridgette-DHS](#) on Tue, 19 Dec 2017 19:31:59 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS Data Processing Manager, Albert Themme:

I checked the imputation and recode programs for 2000 and 2005. Both handled the construction of HW16, B1 and B2 in the same way.

As the recode documentation states all dates are in the Ethiopian calendar.

Thus in 2000 HW16 was derived from HC16, which is a copy of QHC29D as recorded on the questionnaire.

Similarly, in 2005 HW16 was derived from HC16, which is a copy of QC53D again as recorded on the questionnaire.

B1 and B2 are derived from B3 (CMC data of birth of child) with is a copy of Q215C. The computation of Q215C had one difference from our standard procedure. The Ethiopian month and year of birth were used, but when the month of birth was recorded as 13, then 13 months were added to the last year.

E.g. the child was born 13 1960, then the CMC is computed as  $60 * 12 + 13$ .

When the year and month of birth are then returned from this CMC, then the 13th month is really added to the 1st month of the new year.

Let me know if this is not clear.

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

---

Posted by [kcaglaya@tulane.edu](mailto:kcaglaya@tulane.edu) on Tue, 19 Dec 2017 19:50:28 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Dear Bridgette, Ladys and Albert,

Thank you so much for your time and help. We greatly appreciate your kind effort. We are very close to clarify this issue completely. I just need your help and expertise on one last thing. (at least for this project :))

Yesterday, Ladys confirmed that a child's day of birth (b17 or hw16) and month of birth (b1) are reported in Gregorian Calendar while the year of birth (b2) is reported in Ethiopian Calendar in 2016 survey data.

Today, Albert confirmed that a child's day (hw16), month (b1) and year (b2) of birth are all reported in the same way and in Ethiopian Calendar in 2000 and 2005 survey data.

If somebody can confirm the way these variables (i.e. hw16, b1, b2) are reported in 2011 survey data, we would be able to solve the problem completely.

I think, in 2011 these variables are reported as they are reported in 2016 (i.e. hw16 and b1 in Gregorian Calendar and b2 in Ethiopian Calendar) because of the reasons that I mentioned in previous posts. However, I would really appreciate a confirmation from you on this matter.

Once again, thank you very much.

Regards,

Koray

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Bridgette-DHS](#) on Tue, 19 Dec 2017 21:31:20 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Another response from Ladys Ortiz:

As recommended in my previous email, you should read the recode documentation that we usually post within the zipped dataset. The following is excerpted from the documentation mentioned earlier, and hopefully, this answers your question:

Quote:All standard variables are based on calendar dates and century month codes are given in the Ethiopian Calendar. The Ethiopian year consists of 365 days, divided into 12 months of 30

days and one month of 5 days (6 days in a leap year). Ethiopia's new year falls on September 11 and ends the following September 10 according to the Gregorian calendar. Before the production of any indicators with these data the Ethiopian calendar was converted to the Gregorian calendar but conserving the Ethiopian year; however, the Ethiopian first month is considered in the logic as January, the 2nd as February, etc. For dates including year, month and day the conversion is precise since both calendars have 365 or 366 days; for dates including only year and month, the 13th month was included in December.

Thanks,

Ladys Ortiz

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [kcaglaya@tulane.edu](mailto:kcaglaya@tulane.edu) on Tue, 19 Dec 2017 22:03:34 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi Ladys,

Found it!

I thought you were referring to general Recode Manuals which don't have specific information about this issue. I found the word document for the 2011 survey in the "etir61" folder and it confirms that the reporting of the dates are the same as in 2016.

I would like to thank you, Albert and Bridgette for your help and support.

Happy Holidays.

Regards,

Koray

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Ogriv](#) on Tue, 05 Feb 2019 15:36:03 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hello

I am now working with the 2016 EDHS survey.  
I understand from the thread above that b1 (birth month) is in the Gregorian calendar, but b2 (birth year) is still in the Ethiopian calendar.

I also understand that for this survey, because day of birth has also been collected, a fairly accurate conversion can be done.

How can I get an accurate Gregorian birth year (b2) in the 2016 survey?

(I have tried Tom Pullum's method for 2011, but then I realised it might not work for 2016.)

Thanks  
Sandra

---

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [kcaglaya@tulane.edu](mailto:kcaglaya@tulane.edu) on Tue, 05 Feb 2019 15:48:09 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi Sandra,

I don't know which software you are using, but here is the Stata code I wrote to convert the dates in 2011 and 2016 Surveys;

\*Convert from Ethiopian to Gregorian for 2011 and 2016

```
foreach y of numlist 2011 2016 {
```

```
use "`projectdir'/data/eth_`y'_raw/etkr`y'dt/ETKR`y'FL.dta", clear
```

```
*observations with exact birhtday  
replace hw16=. if hw16==98 | hw16==99  
gen birthday_available=hw16!=.
```

```
gen y_gregorian=.  
replace y_gregorian=b2+7 if (b1>=9 & hw16>=6) | b1>9  
replace y_gregorian=b2+8 if (b1==9 & hw16<6) | b1<9  
gen m_gregorian=b1
```

```
gen child_birthyear=y_gregorian  
gen child_birthmonth=b1  
gen child_birthday=hw16
```

```
gen child_birthdate_g_days=mdy(child_birthmonth, child_birthday, child_birthyear)  
gen child_birthdate_gregorian=child_birthdate_g_days  
format child_birthdate_gregorian %td
```

\*convert interview date from Ethiopian to Gregorian Calendar

```
gen interview_y_gregorian=.  
replace interview_y_gregorian=v007+7 if (v006>=9 & v016>=6) | v006>9  
replace interview_y_gregorian=v007+8 if (v006==9 & v016<6) | v006<9  
gen interview_m_gregorian=v006
```

```
gen interview_year=interview_y_gregorian
```

```
gen interview_month=v006  
gen interview_day=v016
```

```
gen interview_date_g_days=mdy(interview_month, interview_day, interview_year)  
gen interview_date_gregorian=interview_date_g_days  
format interview_date_gregorian %td
```

```
tempfile `y'_temp
```

```
save "`y'_temp"
```

```
}
```

Best,

Koray

---

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Ogriv](#) on Tue, 05 Feb 2019 17:14:07 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Thanks Koray  
I am using Stata.

This looks very helpful.

I only need Gregorian versions of b1 and b2, but I will run this and let you know how I get on.

Many thanks

Sandra

---

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Ogriv](#) on Wed, 06 Feb 2019 13:05:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi there Koray  
I ran the code and it seems to have worked well.

Interesting to see that child birth year looks like this:

```
child_birth |  
   year |   Freq.   Percent   Cum.
```

child_birt	child_birtheayear	2010	2011	2012	Total
hmonth					
1	0	0	265	1,156	
2	0	0	201	953	
3	0	0	231	987	
4	0	0	235	955	
5	0	5	243	1,047	
6	0	28	183	941	
7	0	50	195	897	
8	0	90	199	902	
9	108	130	132	751	
10	148	135	133	740	
11	127	122	145	695	
12	114	116	101	617	
Total	497	676	2,263	10,641	

I can only imagine that in 2010 and 2011 there was not full sampling.  
And I assume that full sampling did happen in 2016.  
Is that what you concluded?

Best Wishes  
Sandra

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Ogriv](#) on Wed, 06 Feb 2019 14:01:31 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi there  
I have now looked at a cross-tab of the converted birth month and birth year.  
It looks like this:

child_birt	child_birtheayear	2010	2011	2012	Total
hmonth					
1	0	0	265	1,156	
2	0	0	201	953	
3	0	0	231	987	
4	0	0	235	955	
5	0	5	243	1,047	
6	0	28	183	941	
7	0	50	195	897	
8	0	90	199	902	
9	108	130	132	751	
10	148	135	133	740	
11	127	122	145	695	
12	114	116	101	617	
Total	497	676	2,263	10,641	

child_birt	child_birtheayear	2013	2014	2015	Total
hmonth					

-----+-----+-----				
1	247	219	209	1,156
2	168	192	184	953
3	200	190	181	987
4	207	172	169	955
5	218	177	209	1,047
6	186	231	179	941
7	167	170	206	897
8	180	184	182	902
9	139	179	56	751
10	145	178	1	740
11	121	180	0	695
12	120	166	0	617
-----+-----+-----				
Total	2,098	2,238	1,576	10,641

child_birt			
child_birt	hyear		
hmonth	2016	Total	
-----+-----+-----			
1	216	1,156	
2	208	953	
3	185	987	
4	172	955	
5	195	1,047	
6	134	941	
7	109	897	
8	67	902	
9	7	751	
10	0	740	
11	0	695	
12	0	617	
-----+-----+-----			
Total	1,293	10,641	

From the EDHS 2016 report, interview data collection took place from January-June 2016. If this is the case, then some of these birth months must be incorrect for 2016.

Also it looks like full data are only available for 2012-2014 from the table above.

Sandra

---

Subject: Re: Problem with dates in the Ethiopia datasets  
 Posted by [kcaglaya@tulane.edu](mailto:kcaglaya@tulane.edu) on Wed, 06 Feb 2019 14:03:07 GMT

Hi Sandra,

Your sample size for 2016 (children with exact birth date) looks fine. However, it looks like there is a problem with your 2011 survey data, since the sample size (children with exact birth date) should be similar to 2016. Are you sure that you are using the right survey data? (You know that there are different version of the survey such as all children, children under five or only male households, etc...) I attach a table showing our final sample derivation for the paper. I hope it helps.

Best,  
Koray

---

#### File Attachments

1) [dhs\\_ethiopia\\_samplesize.docx](#), downloaded 949 times

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Ogriv](#) on Wed, 06 Feb 2019 14:09:15 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Thanks Koray

I'm interested in children under 5 in the 2016 survey.

You have a total of 10,006.

I have not yet weighted the data as I'm first trying to establish birth month and year.

But my unweighted total for children under 5 in the 2016 survey is 10,641.

I think I am looking at the right sample, as it's those children included in the hw16 variable (same as b1 and b2).

But I'm worried about the months and years after conversion - see my previous response.

Sandra

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Ogriv](#) on Wed, 06 Feb 2019 14:15:58 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hi again

I can also see from discussions on this thread that for the 2016 survey, the birth month has already been converted to the Gregorian calendar.

Whereas the birth year has not.

That means perhaps that different conversion code needs using for the 2016 survey compared to the 2011 survey.

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [kcaglaya@tulane.edu](#) on Wed, 06 Feb 2019 14:21:59 GMT

---

Sandra,

The way dates are reported is the same in 2011 and 2016 surveys. You can also see it from the survey documents, see my response on Dec 19th;

"I found the word document for the 2011 survey in the "etir61" folder and it confirms that the reporting of the dates are the same as in 2016."

So, the conversion from Ethiopian to Gregorian (only for years) should be the same for 2011 and 2016 surveys.

Best,

Koray

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Ogriv](#) on Wed, 06 Feb 2019 14:26:10 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hello Koray

Do you mean that in the 2016 survey only the years need to be converted?

I can see that birth month in the 2016 survey has already been converted, as there are 12 months (not 13).

Am I right to think that the conversion to birth year relies on birth month being correct?

If so, then I can't just convert years on their own.

It looks like from my cross-tabbing of birth month and year that it's the months that are wrong in the converted data (some of them occur after data collection is finished in June 2016).

Sandra

---

---

Subject: Re: Problem with dates in the Ethiopia datasets

Posted by [Mark](#) on Wed, 01 May 2019 05:56:13 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Dear Sir/Madam,

I am using the 2016 EDHS IR dataset to calculate contraceptive method used prior to the most recent birth.

I am actually interested to calculate it for women who are eligible for birth interval variable (I dropped non-eligible respondents (keep if b11\_01!=.)).

Even though I used the command (presented bellow) that should be used for calendar data using Stata software (after reading the 'DHS Contraceptive Calendar Tutorial'), I got the highest number of missing data (30.31%) which is not common in DHS data. Is there anything I may miss in my program? May you check my program code, and provide me the right program code for this particular case the 2016 Ethiopian DHS please.

- \* Step 1.1
- \* length of full calendar string including leading blanks (80)
- \* actual length used according to v019 will be less

```
egen vcal_len = max(strlen(vcal_1))
```

- \* most calendars are 80 in length, but those without method use may be short, so use the max

label variable vcal\_len "Length of calendar"

- \* Step 1.2
- \* position of last birth or terminated pregnancy in calendar

```
gen lb = strpos(vcal_1,"B")
gen lp = strpos(vcal_1,"T")
```

- \* update lp with position of last birth if there was no terminated pregnancy,
- \* or if the last birth was more recent than last terminated pregnancy

```
replace lp = lb if lp == 0 | (lb > 0 & lb < lp)
```

- \* e.g. if calendar is as below ("\_" used to replace blank for display here):

```
* _____00000BPPPPPPP00000055555500000TPP00000000000000BPPPPPP PP000000000
*           ^
```

- \* lp would be 20

label variable lp "Position of last birth or terminated pregnancy in calendar"

label def lp 0 "No birth or terminated pregnancy in calendar"

label value lp lp

- \* get the type of birth or terminated pregnancy
- \* lp\_type will be set to 1 if lp refers to a birth,
- \* and 2 if lp refers to a terminated pregnancy using the position in "BT" for the resulting code

```
gen lp_type = strpos("BT",substr(vcal_1,lp,1)) if lp > 0
```

label variable lp\_type "Birth or terminated pregnancy in calendar"

label def lp\_type 1 "Birth" 2 "Terminated pregnancy"

label value lp\_type lp\_type

list vcal\_1 lp lp\_type in 1/5

tab lp lp\_type, m

- \* Step 1.3
- \* if there is a birth or terminated pregnancy in the calendar then calculate CMC
- \* of date of last birth or pregnancy by adding length of calendar to start CMC
- \* less the position of the birth or pregnancy
- \* calendar starts in CMC given in v017
- \* lp > 0 means there was a birth or terminated pregnancy in the calendar

```
gen cmc_lp = v017 + vcal_len - lp if lp > 0
```

label variable cmc\_lp "Century month code of last pregnancy"

- \* e.g. if calendar is as below and cmc of beginning of calendar (V017) = 1321:

```
* _____00000BPPPPPPP00000055555500000TPP00000000000000BPPPPPP PP000000000
```

- \* cmc\_lp would be 1381, calculation as follows:
- \* 1321 + 80 - 20 (80 is the vcal\_len, and 20 is the position of lp)

```
list v017 lp vcal_len cmc_lp in 1/5
```

- \* check the variables created.

tab lp

tab cmc\_lp

- \* list cases where cmc\_lp and b3\_01 don't agree if the last pregnancy was a birth

```
list cmc_lp b3_01 if lp > 0 & lp == lb & cmc_lp != b3_01
```

- \* there shouldn't be any cases listed.

\* Step 1.4

- \* get the duration of pregnancy and the position of the month prior to the pregnancy
- \* start from the position after the birth in the calendar string by creating a substring
- \* indexnot searches the substring for the first position that is not a "P" (pregnancy)
- \* piece is the piece of the calendar before the birth ("B") or termination ("T") code

```
gen piece = substr(vcal_1, lp+1, vcal_len-lp)
```

- \* find the length of the pregnancy

```
gen dur_preg = indexnot(piece, "P") if lp > 0
```

- \* dur\_preg will be 0 if pregnant at the start of the calendar

label variable dur\_preg "Duration of pregnancy"

- \* e.g. if calendar is as below:

```
* ____00000BPPPPPPPP00000055555500000TPP00000000000000BPPPPPP PP000000000
*      |12345678^
```

- \* dur\_preg would be 9 for the last pregnancy (1 B plus 8 Ps)
- \* if we find something other than a "P" then that is the month before the pregnancy
- \* if it returns 0 then the pregnancy is underway in the first month of the calendar
- \* now get the position in the calendar to reflect the full calendar
- \* not just the piece before the birth, by adding lp
- \* \_bp means 'before pregnancy'. pos\_bp means position before pregnancy

```
gen pos_bp = dur_preg + lp if dur_preg > 0
```

label variable pos\_bp "Position before pregnancy"

label def pos\_bp 0 "Pregnant in first month of calendar"

label val pos\_bp pos\_bp

- \* e.g. if calendar is as below:

```
* ____00000BPPPPPPPP00000055555500000TPP00000000000000BPPPPPP PP000000000
*      ^
```

- \* pos\_bp would be 29

```
list vcal_1 lp dur_preg pos_bp in 1/5
```

tab dur\_preg lp\_type, m

\* Step 1.5

- \* find the last code that is not 0 before the pregnancy (using indexnot),
- \* searching in a substring of the calendar from the month before pregnancy and earlier,
- \* but not more than 5 years back
- \* lnz means 'last non-zero before the pregnancy'

```
gen lnz = indexnot(substr(vcal_1, pos_bp, vcal_len - pos_bp + 1), "0") ///
if inrange(pos_bp, 1, vcal_len)
```

- \* get the actual position in the calendar of the last non-zero before the last birth

```
gen pos_lnz = pos_bp + lnz - 1 if inrange(lnz, 1, vcal_len)
```

- \* if last non-zero is more than 5 years before interview, set position to 0

```
replace pos_lnz = 0 if lnz == 0 | (pos_lnz != . & pos_lnz > v018+59)
```

label variable pos\_lnz "Position in calendar of last non-zero before pregnancy"

label def pos\_lnz 0 "No non-zero preceding the pregnancy in the last 5 years"

label val pos\_lnz pos\_lnz

- \* list a few cases to check

```
list vcal_1 lp pos_bp pos_lnz in 1/5
```

\* Step 1.6

- \* check if the respondent is using a method before the pregnancy but in the last 5 years

```
gen code_lnz = substr(vcal_1, pos_lnz, 1) if inrange(pos_lnz, v018, v018+92)
```

```

replace code_lnz = "0" if pos_lnz == 0
* if the code is NOT(!) a zero ("0"), a "B", "P" or "T" then the respondent was using a method
gen used_bp = !inlist(code_lnz, "0", "B", "P", "T") if code_lnz != ""
label variable code_lnz "Last non-zero code before pregnancy"
label variable used_bp "Using a method before the last pregnancy"
label def used_bp 0 "No" 1 "Yes"
label val used_bp used_bp
* list a few cases to check
list vcal_1 lp pos_bp pos_lnz code_lnz used_bp in 1/5
* Step 1.7
* last method used before pregnancy, but may have been followed by a period of non-use
* converting the string variable to numeric, although it isn't really necessary for most analyses
* set up a list of codes used in the calendar, with each position matching the coding in V312
* use a tilde (~) to mark gaps in the coding that are not used for this survey
* e.g. Emergency contraception and Standard days method do not exist in this calendar
* note that some of the codes are survey specific so this list may need adjusting
scalar methodlist = "123456789WNALCF~M~"
gen method_bp = strpos(methodlist,code_lnz) if code_lnz != ""
* convert the missing code to 99
replace method_bp = 99 if code_lnz == "?"
* now check if there are any method codes that were not converted, and change these to -1
replace method_bp = -1 if method_bp == 0 & used_bp == 1
* alternatively,
* use the do file below to set up survey specific coding using scalar methodlist and label method
* and recode the method and/or reasons for discontinuation
* include the path to the do file if needed
*run "Calendar recoding.do" code_lnz method_bp
* and skip the value labeling in step 2.8 as the do file above includes the value labeling
* if no method was used, set method_bp to 0
replace method_bp = 0 if used_bp == 0
* Step 1.8
* label the method variable and codes
label variable method_bp "Method used before the last pregnancy (numeric)"
label def method ///
  0 "No method used" ///
  1 "Pill" ///
  2 "IUD" ///
  3 "Injectable" ///
  4 "Diaphragm" ///
  5 "Condom" ///
  6 "Female sterilization" ///
  7 "Male sterilization" ///
  8 "Periodic abstinence/Rhythm" ///
  9 "Withdrawal" ///
  10 "Other traditional method" ///
  11 "Norplant" ///
  12 "Abstinence" ///
  13 "Lactational amenorrhea method" ///

```

```
14 "Female condom" ///  
15 "Foam and Jelly" ///  
16 "Emergency contraception" ///  
17 "Other modern method" ///  
18 "Standard days method" ///  
99 "Missing" ///  
-1 "***Unknown code not recoded***"  
label val method_bp method  
tab method_bp
```

---

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Bridgette-DHS](#) on Mon, 06 May 2019 11:55:46 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from our Research & Data Analysis Director, Tom Pullum:

It's possible that the problem could be due to the use of a different calendar (months/years, not the DHS calendar!) in Ethiopia. That seems unlikely to me, because this survey is internally consistent; ALL dates in the data are given in the Ethiopian version. Is there anything in the report or on STATcompiler that you can calibrate with?

Sorry--I'd like to help but just don't have time to review your code.

---

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [kbhan89](#) on Tue, 21 Apr 2020 20:19:52 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi,

Sorry to bring up this issue again.

I'm currently working on the 2016 Ethiopia DHS data. I have deliberately read the discussions seen above and the recode documentation.

However, I still have some issues making me confused. I would be grateful if you could help me clarify the following points.

1) I wonder if v009 (respondent's month of birth) has already been converted to the Gregorian calendar or not. If I understood the recode and documentation the discussions above correctly, v009 should be the one based on the Ethiopian calendar since the information on a respondent's day of birth was not collected and then precise conversion cannot be done. If I compute v011

(date of birth - cmc) manually using v009 (respondent's month of birth) and v010 (respondent's year of birth - based on the Ethiopian calendar), I can get the same values of v011 in the dataset, leading me to guess that v009 is also from the Ethiopian calendar.

2) Thanks to the clear explanations done above, I have understood that b1(month of birth) has already been converted to the Gregorian calendar while b2(year of birth) has not. Then, could you let me know how b3 (date of birth - cmc) was computed? If I do the same computation as I did for v011 (i.e.  $b3 = (b2 - 1900) * 12 + b1$ ), I get the same values of b3 in the dataset. Does it mean that these cmc values are coming from the months by the Gregorian calendar and the years by the Ethiopian one? I need to convert b2 to the Gregorian one by using b3, but the values of b3 in the data file are a little confusing to me.

I might be missing something important. Please correct me if I have made wrong points.

Thank you very much for your help in advance.

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Trevor-DHS](#) on Wed, 20 May 2020 14:16:13 GMT  
[View Forum Message](#) <> [Reply to Message](#)

The dates in the Ethiopia surveys have not been converted to the Gregorian calendar but they have been "squeezed" from a 13 month calendar down to fit a 12 month calendar (comparable to other countries). In this "squeezing" process the years are unchanged and only the month and day are affected. For dates that only have month, but not day, the 13th month is recoded to the 12th month. For dates with month and day, the dates are squeezed by moving the dates back into 12 months, with some months having 31 days and some 30 days, similar to Gregorian months.

The century month codes are based on these squeezed Ethiopian dates.

For Gregorian dates, the simplest thing to do is to add 92 months to the CMC. This won't be perfect as it doesn't take into account the day of birth or interview, but you could adjust for that by subtracting 1 month from the CMC if the day was less than 12.

See the Guide to DHS Statistics and search in the PDF for Ethiopia for more information.

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [VictorJanVilla](#) on Tue, 07 Mar 2023 18:36:29 GMT  
[View Forum Message](#) <> [Reply to Message](#)

Hi,

Sorry to bring up this issue again.

I'm currently working on all the available Ethiopia DHS data on Stata. Reading the conversation I

saw a few possible solutions to convert the Ethiopian date to gregorian.

I am mainly interested in the year and month of the interview since I am merging these datasets with external climatic data and thus, I would like to obtain the most precise date possible.

Trying the following commands shared by Bridgette-DHS:

```
gen vg008=v008+92
gen vg007=int((vg008-1)/12)
gen vg006=(vg008-12)*v007  /// I included parenthesis in here
replace vg007=vg007+1900
```

This strategy works for the year but not for the months(since it does not give months' values). I saw that other users have simply assigned the same month to the CMC, which I think is not the right approach.

I also checked with 2000 and 2005 rounds that include gregorian dates and the gen "vg006=(vg008-12)\*v007" do not respect the distribution of the months in gregorian format.

Did anyone solve this problem?

Thank you very much in advance!

Best,

Victor

---

Subject: Re: Problem with dates in the Ethiopia datasets  
Posted by [Bridgette-DHS](#) on Tue, 07 Mar 2023 21:05:46 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS staff member, Tom Pullum:

We have nothing new to add or to suggest. You can get information about the Ethiopian calendar from websites.

---