Subject: Filtering children's data for nutrition indicators
Posted by Mark_H_22 on Fri, 08 Nov 2024 10:10:36 GMT
View Forum Message <> Reply to Message

Hello,

I have a question about filtering children's datasets to calculate dietary diversity. In the Guide to DHS statistics (DHS8), the following description for calculating the denominator is given (page 11.61):
To select the cases for the denominator, first filter or keep only living children living with the mother born in the preceding 24 months (keep/select if b19 < 24 & b9 = 0), then keep only the youngest child (keep/select if the first entry in the dataset (_n = 1) or the first for this respondent

file), then finally restrict to children age 6-23 (note that you cannot select for children age 6-23 in the first step as this would include some cases that were not asked the questions.

I understand this as the 3 following steps:
1) Filter to keep only living children, living with their mother, under 24 months old.
2) Keep only the youngest child per mother.
3) Remove children younger than 6 months.

I am unsure why filtering the data to only contain children aged 6-23 months must be the last step, and not the first. In step 2, only the youngest child is kept, and all older siblings are removed. Therefore, if the youngest child is younger than 6 months, this child is also removed in step 3 and I end up with no children for this particular respondent (mother).
Surely the data should be filtered to keep children in the range 6-23 months first, and then keep only the youngest child?

Any clarification on this would be really helpful. Thank you.

Subject: Re: Filtering children's data for nutrition indicators
Posted by Janet-DHS on Thu, 14 Nov 2024 21:13:36 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:
Either sequence of steps will work, but I agree with you that it would be more efficient to restrict the age range to 6-23 months in the first step. I will paste below the Stata lines to identify the reference child with "egen seq()". Some people add a step to require that b5=1, but that is not necessary because if b9=0 then b5 must be 1.
* Steps to find the youngest child age 6-23 months living with the mother

* Illustrate with Nepal 2022 survey, KR file

use "...NPKR82FL.DTA", clear

```
gen select=0

egen sequence=seq() if b19>=6 & b19<=23 & b9==0, by(v001 v002 v003 bidx)

replace select=1 if sequence==1
```