
Subject: Household Respondent's age (Philippines, years 2013,2017 and 2022)

Posted by [josueestrada@ph](#) on Sun, 18 Jun 2023 10:54:57 GMT

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

Hello,

I am doing a research involving data from the Philippine DHS for the years 2013, 2017, and 2022. As my analytic sample, I will be including the respondents to the household survey. I am trying to determine their age but when I prepare the dataset for my analytic sample, I get respondents with ages below 15 and as high as 96+. I also have a few respondents with unknown ages. I've tried creating the dataset both from the household (HR) and individual members (PR) datasets, matching/keeping information based on hv003 (line number of each respondent) for each household number on a cluster and for a particular year. As an example, here are the commands I used in Stata to extract the information in the PR datasets:

```
bysort hv007 hv001 hv002: gen seq=_n  
bysort seq: keep hvidx==hv003
```

Can someone help me figure out if I did something wrong? I would appreciate your help. Thank you very much.

Kind regards,
Josue