
Subject: Missing value in Birth file-NFHS4
Posted by [Jayanta](#) on Fri, 19 Jan 2018 12:55:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear DHS team,

I have run the birth file of NFHS-4 for India and found missing cases in v701 "Husband/partner's education level". Can you please let me know what is the meaning of this missing cases.

Thanks,
Jayanta

Subject: Re: Missing value in Birth file-NFHS4
Posted by [Jayanta](#) on Fri, 02 Feb 2018 11:09:39 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello DHS team,

Can you please suggest what is the meaning of missing cases for husband education, women occupation in BR file of NFHS-4 data?

Your help is highly appreciated.

Regards,
Jayanta

Subject: Re: Missing value in Birth file-NFHS4
Posted by [Liz-DHS](#) on Mon, 26 Mar 2018 13:46:54 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear User, From the Standard Recode Manual:

Quote:

Rationale for Recoding

The individual data are transformed into a standardized recode dataset for several reasons:

- First, dates for several key events are imputed as much analysis of the data is based on these events

and their dates are often incomplete or missing. The imputed dates are included in the data file to allow analysts to produce results consistent with those published by DHS and to save analysts the time and trouble of creating their own imputation schemes.

Quote:

Coding Standards

Special codes are used throughout the data file for certain responses. The general coding scheme is presented

below. The codes given apply to 4 digit, 3 digit, 2 digit and 1 digit variables, respectively. If there are other special responses to questions, these are coded in decreasing order from these special codes, i.e., 9996, 996, 96, 6; 9995, 995, 95, 5; etc.

BLANK Variable is not applicable for this respondent either because the question was not asked in a particular country or because the question was not asked of this respondent due to the flow or skip pattern of the questionnaire.

9999, 999, 99, 9 This question should have been answered by the respondent, but the questionnaire contained no information for this variable (missing data).

9998, 998, 98, 8 The respondent replied "Don't know" to this question.

9997, 997, 97, 7 The answer to this question was inconsistent with other responses in the questionnaire and it was thought that this response was probably in error. The response was changed to this code to avoid further problems due to inconsistency of information. This usually takes place during the secondary editing stage of data processing.

Please refer to The Guide to DHS Statistics <https://dhsprogram.com/publications/publication-dhsg1-dhs-questionnaires-and-manuals.cfm>, The DHS6 Standard Recode Manual <https://dhsprogram.com/publications/publication-DHSG4-DHS-Questionnaires-and-Manuals.cfm> and various other resources for working with and using the datasets <https://dhsprogram.com/data/Using-Datasets-for-Analysis.cfm>

Subject: Re: Missing value in Birth file-NFHS4
Posted by [Jayanta](#) on Mon, 26 Mar 2018 14:29:50 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you for your reply.
I have already read this information from the guideline/manual
It would be great if you can provide information on how to calculate husband educational status and women working status that will be helpful.

Please see my specific queries regarding on missing data-

- 1) While using weight (v005/1000000) on caste variable s116, ST 's sample size almost reduced to 50%- why?
- 2) Women's education(v106) information is available but currently working status(v714,v716) shows missing cases-why? I am facing difficulties to generate working women status- kindly suggest how to proceed
- 3) Currently married women (v501) and education(v106) information is available but, for the same respondent, the highest education level for partners(v701) is missing-why?

Regards,
Jayanta

Subject: Re: Missing value in Birth file-NFHS4
Posted by [Trevor-DHS](#) on Mon, 26 Mar 2018 16:14:54 GMT
[View Forum Message](#) <> [Reply to Message](#)

In answer to your questions:

1) Please use s116 in conjunction with v131. s116 is missing for any case on which no caste/tribe is reported in v131, so you may want to recode the missing on s116 to a category labeled "No caste/tribe". For your comment that the scheduled tribes group sample size is reduced by 50% when weighted, that would be because the scheduled tribes were effectively oversampled in the survey and the weights correct that back to the correct proportion of the population.

2) and 3) The questions on working status and education of the partner were only asked in a subset of households using what was known as the state specific modules. The variable ssmod identifies cases in which this module was used. Thus you need to select on ssmod == 1 before analyzing these variables. These variables are missing for all cases not selected for the state specific modules.

Subject: Re: Missing value in Birth file-NFHS4
Posted by [Jayanta](#) on Tue, 27 Mar 2018 07:57:32 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks Trevor for your reply.

I am still worried why 50% of ST population is oversampled.

For the responses of 2&3, if we see the frequency of the currently working status(v714) and highest education level for partners(v701)with or without ssmod == 1, results are same and I think it's already considered the value of ssmod == 1 in the dataset.

The reason may be different.

Regards,
Jayanta

Subject: Re: Missing value in Birth file-NFHS4
Posted by [Trevor-DHS](#) on Tue, 27 Mar 2018 16:50:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

1) It isn't 50% are oversampled. It is that the enumeration areas that people from scheduled tribes

lived in were, roughly on average, sampled at twice the level that they represent in the population, and to correct for this in national data, they receive half the weight (roughly).

2) for points 2 &3, you asked me why so many were missing. The reason they are missing is because of the state specific modules selection, and you can confirm this by crosstabbing v701 or v714 with ssmod with the missing data option on. All cases for which ssmod is 1 will be missing for v701 or v714, so yes, obviously the frequency of the non-missing values are going to be exactly the same whether you filter on ssmod == 1 or not.
