

---

Subject: Imputation of missing data

Posted by [Ashlesha Pal](#) on Wed, 21 Aug 2024 20:05:58 GMT

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

---

I have been doing an analysis for which I need to generate women empowerment scores, based on "Swapper scale for women empowerment" (Ewerling F, Raj A, Victora CG, Hellwig F, Coll CV, Barros AJ. SWPER Global: A survey-based women's empowerment index expanded from Africa to all low- and middle-income countries. J Glob Health. 2020 Dec;10(2):020343. doi: 10.7189/jogh.10.020434. PMID: 33274055; PMCID: PMC7699005.) using IR file of India, Pakistan and Bangladesh. I am unable to generate these scores for a subpopulation as the husband's education in about 41 observations in that subpopulation are missing or don't know. IS there a way I can some how impute those values without missing those observations?

---

---

Subject: Re: Imputation of missing data

Posted by [Bridgette-DHS](#) on Thu, 22 Aug 2024 16:28:52 GMT

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

---

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

In this kind of a situation the researcher (you) must make a "judgment call" about what to do. I can suggest potential strategies. One possibility is to drop those cases. Naturally, we don't like to do that. Another would be to assign these men the mean or median or modal value for all men or for all men in some subpopulation that includes these men, such as their district. I would not recommend anything more complicated. Elaborate methods, such as multiple imputation, do exist, but with only 41 cases that would be a waste of effort.

Whatever you do, it would be good to include a comment or footnote describing it, so someone could potentially match your results. If you look at the tables in DHS final reports, you will sometimes find a footnote that says what was done with missing (distinct from Not Applicable) values.

You can also look at whether any estimates appear to change or differ, depending on how you handled such cases. With only 41 questionable cases, you will probably find that the results are not sensitive to whatever option you choose.

---