
Subject: Survey weighting with HIV and Domestic violence weights
Posted by sureshkumar.19@buckeyemail on Thu, 14 Dec 2023 18:58:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi, I am trying to perform multinomial logistic regression with Mozambique 2015 data using both HIV biomarker data and the domestic violence variables from the women's recode data. The HIV test results and the domestic violence module have the same n=3006. I am not sure how to proceed with the survey weights here. Please advice.

Subject: Re: Survey weighting with HIV and Domestic violence weights
Posted by Bridgette-DHS on Mon, 18 Dec 2023 13:22:32 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

If you have merged the IR and AR files, then you have three possible weights: v005, d005, and hiv05. If your analysis includes both DV variables and hiv03, the weight variable would be either d005 or hiv05. The general rule DHS recommends is to use the weight that has the most adjustment for nonresponse. That is, if there is more nonresponse for DV than for HIV then use d005; if there is more nonresponse for HIV than for DV, then use hiv05. The latter is probably true, but that's an empirical question. You could calculate the nonresponse rates or you could just look at the report, which should give both nonresponse rates, and compare.

It would even be possible to calculate new weights, adjusting for nonresponse on BOTH DV and HIV, but I don't think it would be worth the extra effort.

To summarize: use hiv05 if there is more nonresponse for HIV (by women) than for DV, which I would expect, or use d005 if the reverse is true.

Subject: Re: Survey weighting with HIV and Domestic violence weights
Posted by sureshkumar.19@buckeyemail on Fri, 22 Dec 2023 15:06:42 GMT
[View Forum Message](#) <> [Reply to Message](#)

I am trying to navigate the survey weighting in DHS. Am I right in thinking that it is a 2-stage cluster sampling? If so, is it right to think that fpc = v022?

Is this right?

generate wgt = D005/1000000

svyset V021 [pweight=wgt], psu(V021) fpc(V022) strata(V023) singleunit(centered)

Subject: Re: Survey weighting with HIV and Domestic violence weights

Posted by [Bridgette-DHS](#) on Tue, 26 Dec 2023 22:36:18 GMT

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

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

Yes, DHS surveys are two-stage cluster samples. Your svyset statement has a duplicate reference to v021 and you do not need the finite population correction (fpc). Use the following:

```
svyset v021 [pweight=d005], strata(v023) singleunit(centered)
```

As has been mentioned in other postings, pweight does not require that you divide d005 (or the other weight variables) by 1000000. If you omitted that step you would get the same results (try it!).

There are older surveys in which v023 is not the stratum variable. On GitHub we have a list of the stratum variable in each survey. Sometimes you have to construct it as all combinations of v024 and v025.

Subject: Re: Survey weighting with HIV and Domestic violence weights

Posted by [sureshkumar.19@buckeyemail](#) on Thu, 25 Jan 2024 18:48:42 GMT

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

Thank you so much. I really appreciate it.

I had another query:

I am using the subpopulation of those who have responded to the domestic violence questions and consented to the HIV biomarker testing. Does this mean I would have to generate a subpopulation for HIV status?

So far, svyset V021 [pweight=D005], strata(V023) singleunit(centered) is the stata code I am using. Please let me know and thank you in advance.

Subject: Re: Survey weighting with HIV and Domestic violence weights

Posted by [Bridgette-DHS](#) on Fri, 26 Jan 2024 16:13:19 GMT

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

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

We normally advise people to use d005 when analyzing DV and hiv05 when analyzing HIV. I don't recall someone looking at the subpopulation of BOTH DV and HIV. I would expect that nonresponse is higher for HIV than for DV, and therefore hiv05 would be preferable to d005. However, you could look at the final report and see whether that's correct.

Your choice of d005 or hiv05 should have a negligible impact on the results, but you could try to

check that by comparing some key estimates using d005 or hiv05.
