Understanding Society User Support - Support #2278

Effect of self-completion weight on weighted count for sample of people who died

09/01/2025 04:03 PM - Emma Maun

Status:	In Progress	**Start date:**	09/01/2025
Priority:	Normal	**% Done:**	10%
Assignee:	Olena Kaminska		
Category:	Weights		

Description

Dear Understanding Society team,

I would like to check the effect of the self-completion weights for my particular analysis (estimating prevalence and correlates of loneliness in the last year of life), as when applied they increase rather than decrease the weighted count from my original sample (this query is separate to, but related to query number 2154). For background I have pooled data from waves 9-13 for people who were identified as having died and had a valid interview in the wave before. I have pooled the weights without scaling as previously advised. To identify my sample, I exclude people who had missing loneliness questions and covariates. The weights work as expected in the pooled data until I applied them to my final selected sample. After removing anyone with missing loneliness and covariates, the sample size is 743 people but the weighted count is 879. People who died but were missing loneliness and covariates had a lower mean weight:
240 people with missing covariates, mean weight 0.4042998 (sd 0.605)
743 people with full data, mean weigh 1.182772 (sd 0.737).

One reason I can think for this unusual effect is that my sample of people who died may have a high proportion of characteristics common to non-responders, so they are upweighted but I am nervous to apply the weight without advice.

A further difficulty is that the questions on loneliness weren't included in the first 6 months of wave 12 due to COVID, so for wave 12 the sample excludes people interviewed in those months. In addition to the weight reducing the sample for NI (because NI sample only collected in the first half of each wave), I'm not sure if this affects the appropriateness of using weights in other ways e.g. for pooled analysis, the FAQs indicate an even number of 24 month samples should be pooled.

I'd be very grateful for advice on all the above,
Thanks,
Emma

History

#1 - 09/03/2025 12:59 PM - Understanding Society User Support Team

- Status changed from New to In Progress

- Assignee changed from Emma Maun to Olena Kaminska

- % Done changed from 0 to 10

- Private changed from Yes to No

Hello Chloe

Many thanks for your enquiry. The Understanding Society team is looking into it and we will get back to you as soon as we can. We aim to respond to simple queries within 48 hours and more complex issues within 7 working days.

Best wishes,
Understanding Society User Support Team

#2 - 09/03/2025 01:54 PM - Olena Kaminska

Emma,

Thank you for your question. You mention two separate issues.
On your sample selection based on valid answers: this is item nonresponse that weights do not deal with. Weight only deal with unit (person) nonresponse. If your loose many people due to missing item nonresponse (where applicable, meaning in the sample where they were asked), you should try to correct for this missingness in an additional step (through imputation of weighting), especially if you think the item nonresponse is related to your analysis.

On the question being included after 6 months of the fieldwork start: note, if the above missingness comes largely from this issue, you don't need a separate nonresponse correction (as missingness is by design and not person's decision). But you need an additional correction for country distribution and ethnic minority distribution. My suggestion is to post-stratify (match) the distribution of country by ethnic minority in the full sample,

and the subsample in your analysis such that your analysis subsample looks identical to the full sample.

This is easily done by generating a correction = P/p, where P is a proportion in the cell of your subsample, and P is the proportion in the same cell of full sample;
final_wgt=weight*correction.
Use final weight to double check the distribution in the weighted subsample - after the correction it should match the full sample distribution.

Hope this helps,
Olena