
Subject: Aggregating data from facility to Province level
Posted by [spet2094](#) on Tue, 23 Feb 2021 19:52:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello,

My team is looking at several indicators for SPA and we have a question on how to aggregate data on facilities of different sizes up to province level. For example could you provide guidance on how to relatively weight the refrigeration capabilities of a clinic versus a hospital and then get to a province level estimate?

Looking forward to your response, -- also if I need to provide more detail do let me know.

Subject: Re: Aggregating data from facility to Province level
Posted by [Bridgette-DHS](#) on Wed, 24 Feb 2021 13:31:09 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

The SPA files include a weight variable. For example, the facilities file has "facwt". You would use that weight for any outcome variable. For example, you could cross-tabulate region and "has refrigerator" , using `iweight=facwt/1000000`.

Subject: Re: Aggregating data from facility to Province level
Posted by [spet2094](#) on Tue, 02 Mar 2021 11:58:07 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Bridgette,

Thanks for the response. My question is: do the weights take into account the size AND the type of the facility? For example, if I look at the variable in Facility (FC) dataset, and I see the variable q1008 'refrigeration observed' in the facility (1/2 for 0/1 respectively), will the existence of refrigeration in a provincial hospital count for the same as the existence of refrigeration in a health clinic? If not, what will be the difference, and how is the weight therefore calculated? Is it weighted by the number of beds in the facility for example? Or the capacity of staff? If neither, how would DHS recommend to deal with this question when thinking about vaccine refrigeration capacity?

Looking forward to your response.

Subject: Re: Aggregating data from facility to Province level
Posted by [Bridgette-DHS](#) on Tue, 02 Mar 2021 20:37:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

The weight for a specific facility in the data file is based on just the sample design. It is inversely proportional to the probability of selection. When the weights are used, you will get (for example) an unbiased estimate of the proportion of ALL facilities of that type that have a refrigerator. It is not (for example) an estimate of the proportion of the population that uses a facility that has a refrigerator. The latter would be good to know but I don't think we can estimate it. If you want to re-weight the data by some measure of size, you can certainly do that.

There is a growing literature on effective coverage, which combines SPA data with DHS survey data. DHS has produced several reports on that topic (e.g. Analytical Study 67, <https://www.dhsprogram.com/pubs/pdf/AS67/AS67.pdf>). It may provide a strategy for what you are thinking of.

Subject: Re: Aggregating data from facility to Province level

Posted by [spet2094](#) on Tue, 16 Mar 2021 23:39:33 GMT

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

Dear Tom

Thank you very much for your response on refrigeration and for sharing the report.

I have a follow up question regarding the number of health care workers per 10,000 population.

I found two points of guidance on this:

1. (The number of health workers at a given time in a given country or region / Total population for the same geographical area)

The ratio can be adjusted to per 10,000 population by multiplying the numerator and denominator by the same factor required for the denominator to equal 10,000.

and

2. For the SARA indicator: $N \text{ per } 10,000 \text{ population} / 23 * 100 \text{ (max.100)}$

So I think this means that the first indicator can only be calculated at the admin level seeing as in most cases we do not know the population that the health facility itself serves. Therefore if we were to map health facility provision, for individual facilities, it would make most sense to map the raw numbers of total health care workers per facility. But for aggregation to health zone level, we would sum the health care workers across facilities and divide by the population of the health zone. So e.g. if there were a total of 20 healthcare workers over a health zone of 20,000 we would do $20 * .5 / 20,000 * .5$ ie. 0.001 HCW for 10,000 population. Is that correct?

What if we were to want to know the average number of health care workers (HCW) per health zone. For that would we add any facility weighting?

Then finally, for the SARA indicator, on the understanding that 23 is the metric for the minimum number of HCW per 10,000 population according to WHO standards, is it correct to say that the indicator is telling us the percentage of the target number of HCW per 10,000 population that the health zone/admin area is achieving?

Thanks again for your response
Regards

Subject: Re: Aggregating data from facility to Province level
Posted by [Bridgette-DHS](#) on Wed, 17 Mar 2021 20:09:18 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

We do not normally produce estimates of healthcare workers per 100,000 population. To weight the numbers of health workers in the sampled facilities up to the population of all facilities, you would need to know the non-normalized sampling fraction for each facility. This may not be available. You would have to look at the sampling information in the SPA report. If the sampling fractions are not given explicitly in the report or in the data files, then they are not available. For SPA censuses, as in Malawi and Haiti, there is no problem. The total number of health workers would be divided by the total population and multiplied by 100,000. But in any case there may be health workers in some settings, such as private doctors' offices, who are not captured by a SPA.

We do not work with SARA data and we cannot answer questions about it.
