

---

Subject: Mapping Z-scores

Posted by [hm\\_DHS](#) on Wed, 02 Jul 2014 23:29:42 GMT

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

---

Hello -

I'm using the Malawi 2010, 2004, and 2000 DHS data.

We're using the GPS coordinates of the clusters to represent mean cluster height-for-age and weight-for-height Z scores spatially.

We are interpolating from the DHS cluster data using Inverse Distance Weighting.

This gives us an estimated mean Z-score for each pixel.

We've then extracted Z-score statistics (min, max, range, mean, std, sum) for each traditional authority (sub-district level). I've attached an example of one of the maps.

I have two questions

- (1) do we need to use the sampling weights when we plot and interpolate the GPS cluster data?
- (2) if we're using inverse distance weighting to make the maps, are we making faulty inferences if we extract data at the sub-provincial level, given that the DHS Malawi 2010 is only meant to be representative at the level of the region or province?

Thank you for your help.

HC

---

File Attachments

1) [IDW\\_test.jpeg](#), downloaded 1014 times

---

---

Subject: Re: Mapping Z-scores

Posted by [Reduced-For\(u\)m](#) on Sun, 06 Jul 2014 21:16:54 GMT

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

---

I just want to bump this question up, in case someone at the DHS has a good answer. I've been thinking about it, but this is a problem we haven't really dealt with here specifically.

My (totally unofficial, based solely on experience and answers to similar problems) two cents - I don't think you can reliably estimate cluster-level mean HAZ, at least not in the sense that the estimate would be representative of the population of the cluster (and I'm guessing that some cluster samples are really small, so there is a small sample problem too for inference).

One thing that might make sense is to only use the smallest nationally representative unit, and then interpolate from there - basically assuming a smooth change in mean HAZ from one regional center to the neighboring regional center (which may or may not make any sense). That could maybe be refined to adjust for urban/rural differences. All this assumes of course that all three survey years are representative at the same geographic levels, that you've re-normalized your weights across survey rounds, that you are happy to assume away time-dependent changes in HAZ from 2000-2010, etc.

I think the trade-off you'll have to make is either getting really spatially rough maps (interpolated from just a few big geographic units) that are population-representative, versus getting finer grained maps that are just representative of the DHS sampling frame - which you could do without using the weights at all. The consensus among the DHS people seems to be to always use weights, but this is not the consensus among academic economists who regularly seem to say "oh, I don't use the weights."

But like I said - I'm hoping someone from DHS weighs in on this, because its a hard question, but one that I think is important given current interests in mapping things like this.

---

---

**Subject: Re: Mapping Z-scores**

Posted by [Thea-DHS](#) on Mon, 07 Jul 2014 19:48:19 GMT

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

---

The DHS Program is currently writing a new report in the Spatial Analysis Report series that will outline some of the key considerations when interpolating Demographic and Health Survey data. There will be an announcement on The DHS Program website once the final version is released. In the meantime, we'd like to make you aware of some guidelines, and direct you to the two most recent Spatial Analysis Report publications, which will help you to understand the displacement of The DHS Program GPS datasets, and how to account for the displacement in your analysis: Geographic Displacement Procedure and Georeferenced Data Release Policy for the Demographic and Health Surveys

<http://dhsprogram.com/publications/publication-SAR7-Spatial-Analysis-Reports.cfm> Guidelines on the Use of DHS GPS Data <http://dhsprogram.com/publications/publication-SAR8-Spatial-Analysis-Reports.cfm>

The DHS Program strongly recommends using sampling weights whenever you use DHS data in analysis, including when interpolating the spatial data.

It is possible to create interpolated layers of DHS indicators using cluster-level estimates and there are numerous published studies that have done this. However, it is important when interpreting the results of such analyses, or conducting your own analysis, to take into account the fact that most DHS surveys are not designed to be representative below the level of the DHS regions (to view or download the DHS regions for Malawi surveys, please visit <http://spatialdata.dhsprogram.com/boundaries.html?country=MW>). Therefore, interpolating the cluster level data or using it to conduct small area estimation will introduce error into your study. Inverse Distance Weighting and other univariate methods of interpolation do not yield any error or uncertainty estimates. Alternatively, multi-variate methods are capable of producing measurements of error alongside the interpolated indicator estimates.

---

If this has not answered your question adequately, please reply.

---