

---

Subject: Re: Calculating a Representative Wealth Index for Clusters Using DHS Sample Weights

Posted by [Rean](#) on Fri, 17 Jan 2025 10:30:08 GMT

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

---

Thank you so much for your insightful response!

Regarding the use of weighted and unweighted means for representing the cluster-level Wealth Index (WI), I would like to confirm my understanding. Since all households within a cluster share the same weight, I am using two approaches to derive a representative value for the cluster's wealth level for machine learning training labels. The two methods are:

1. Unweighted mean WI: The average WI for all households within the cluster.
2. Weighted mean WI: The average WI of all households within the cluster, weighted by the cluster's associated weight (hv005).

In my machine learning experiments, I observed that the model's performance was slightly better when using the weighted WI as training labels rather than the unweighted mean WI. Could the analysis you mentioned, which involves weighting by the cluster's weight, explain this result? In other words, using the weighted mean WI as the label for the clusters' overall wealth status might provide a more representative and useful feature for training the model.

I greatly appreciate your time and valuable insights, which have truly helped me a lot regarding this issue!

---