

---

Subject: Mismatch of number of clusters

Posted by [MrZ](#) on Wed, 09 Jun 2021 16:37:02 GMT

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

---

Dear DHS - Team,

I'm trying to merge GPS data, household and children data files for a number of countries. I've noticed that sometimes there is a mismatch between the number of clusters (as given in v001 and hv001) across different survey files. For example, in the Namibia 2013 data, there are 550 clusters in the household file and 537 clusters in the children file. However, it may well be the case that some clusters do not contain any children, which would explain a slightly smaller number of clusters in the children file.

In addition, the number of locations in the GPS files do match the number of clusters in the survey files relatively seldom, even though the discrepancies are usually rather small. However, in some cases the difference is large, e.g., the Democratic Republic of Congo GPS file from 2013-14 holds 492 locations of clusters while the household file lists 536 clusters.

I am a bit worried that these "missing" clusters could cause a mismatch between GPS coordinates and survey information. Can we be reasonably sure that, e.g., "cluster no. 52" will be the same cluster in all survey and GPS files, even when the number of clusters are not perfectly aligned?

Thank you!

---