

---

Subject: DHS 2022

Posted by [Anonymous](#) on Tue, 14 Jan 2025 13:19:00 GMT

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

---

I would be grateful if someone could advise me on how to get an estimate for the question described below (it is just a fictitious example):

Country: Tanzania

DHS 2022

Region: Dar es Salaam

Clusters: 141+ 142

Desired estimates: (1) Total number and (2) the proportion of children under 5 years old, in the clusters 141 and 142, over the total population of Dar es Salaam.

I know that in Dar es Salaam this proportion is around 11.6 (<https://sensa.nbs.go.tz/publication/report7.pdf>), but I need to know an estimate for particular clusters in relation to the respective region. In the GPS file TZGC81FL.csv, I found the cells (1) UN\_Population\_Count\_2020, (2) U5\_Population\_2020, (3) All\_Population\_Count\_2020, for each cluster, but I noticed that the cells (1) and (3) seem to have very different numbers.

---

Subject: Re: DHS 2022

Posted by [Janet-DHS](#) on Wed, 15 Jan 2025 13:55:48 GMT

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

---

Following is a response from DHS staff member, Tom Pullum:

The Stata lines below illustrate how to get the kinds of numbers I think you are looking for. I am using the PR file for the 2022 survey.  
use "...TZPR82FL.DTA" , clear

\* hv024 is region  
describe hv024

\* the label for hv024 is HV024; list it  
label list HV024

\* which clusters are in hv024=7  
tab hv001 if hv024==7

\* Look at children in clusters 141 and 142

```
gen agelt5=0
replace agelt5=1 if hv105<5
label define agelt5 0 "Age 5+" 1 "Age0-4"
label values agelt5 agelt5
```

\* Unweighted

```
tab agelt5 if hv024==7
```

```
tab agelt5 if hv001==141 | hv001==142
```

\* Weighted

```
tab agelt5 if hv024==7 [iweight=hv005/1000000]
```

```
tab agelt5 if hv001==141 | hv001==142 [iweight=hv005/1000000]
```

---