Subject: R DHS aggregation by Indigenous Identity in Guatemala
Posted by hmwoods02 on Tue, 15 Feb 2022 19:52:50 GMT
View Forum Message <> Reply to Message

Hello,

I am trying to use DHS data in R to create health data that is broken down by tribal group in Guatemala. I have identified variable SETID (self identification), s114 (language learned to speak) or s117(languages spoken at home) as relevant for identifying indigenous identity, however I am not sure how to construct such a conplex variable in R. There are 25 indigenous groups identified in the data.

Thanks

Subject: Re: R DHS aggregation by Indigenous Identity in Guatemala
Posted by Bridgette-DHS on Mon, 21 Feb 2022 13:37:19 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Sampling Specialist, Mahmoud Elkasabi:

You can try the following function to group the variables into one variable (tribalg).

GUBR71 <- GUBR71 %>%
        mutate(tribalg = group_indices(., setid, s114))