

---

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 complex variable in R. There are 25 indigenous groups identified in the data.

Thanks

---