I am currently completing a project that calculates household and women's phone ownership estimates by demographic characteristics (wealth index, gender, rural/urban etc.) across several countries' DHS surveys.

While running the standard code in R for calculating women's phone ownership (I provide the code/function below), I noticed that there is a high rate of NAs in the India data, higher than any other country's (in my dataset spanning 43 country-year surveys). If I run the standard code, I get a low phone ownership rate of approximately 8%. If I omit all NAs (84% of the data!), I still get phone ownership = 53.9% among women, which is higher than the rate stated in the official DHS report and in StatCompiler (50.9%)

Is there a reason for this?

Function for calculating women's phone ownership:

```
analyze_dhs_wm <- function(df_name) {
  df <- get(df_name, envir = .GlobalEnv)
  country_code <- substr(df_name, 1, 3)

  if (!"v169a" %in% names(df)) {
    # Handle the case where v169a does not exist in the dataframe
    phone_ownership_weighted <- NA
    phone_ownership_percent <- NA
    na_percentage <- NA
  } else {
    # Proceed if v169a exists
    phone_ownership_weighted <- sum(df$v169a * (df$v005 / 1e6), na.rm = TRUE) / sum(df$v005 / 1e6, na.rm = TRUE)
    phone_ownership_percent <- phone_ownership_weighted * 100
    na_percentage <- mean(is.na(df$v169a)) * 100
  }

  return(data.frame(country_code = country_code,
            type = "DHS_wm",
            phone_ownership_weighted = phone_ownership_weighted,
            phone_ownership_percent = phone_ownership_percent,
            na_percentage = na_percentage))
}

results_dhs_wm <- do.call(rbind, lapply(dhs_wm_data_frames, analyze_dhs_wm))
```