Setting up DHS Account and Data for HW2

Garrett-Allen

2025-01-22

Pick country

Please pick a country from the following list by putting your name on this Google Sheet. Please only pick one person per country- there are more than enough, so there should be no doubling up.

Request Data

In ordr to do HW2, you'll need access to the required DHS survey data. DHS maintains strict data access rules, particularly for sensitive GPS-tagged data needed for SAE analysis.

Follow these steps to obtain access to DHS datasets (instruction videos also available here)

Register an account; Start by creating an account on the DHS website.

Then, provide project information and specify you're using the data for a class project. Request survey access and geographic access.

Load Data

After you get approval from DHS, you will need to load in the data into R. To do this, please follow the SurveyPrev vignette here.. Please read this in order to do HW2, it will be very helpful. Below I have summarized the steps if you are loading the data from the website rather than downloading the data as a ZIP file.

Make sure to install surveyPrev using devtools::install_github to get the most recent version, as the package is still in development and you'll need the most up to date version.

IMPORTANT: If you copy the below code into your RMD, make sure to remove "eval = FALSE" from the headers, or the code will not run.

```
library(devtools)
devtools::install_github("richardli/surveyPrev")
library(surveyPrev)
library(dplyr)
library(kableExtra)
library(labelled)
library(naniar)
library(sjlabelled)
library(haven)
library(dplyr)
library(geodata)
```

```
#if getting data using rdhs, configure rdhs:
```

```
rdhs::set_rdhs_config(email = "your-email-here",
                      project = "your=project-here")

# example settings
indicator <- "your-indicator-variable-here"
year <- 2023 # for example
country <- "your-country-here"

# get data from rdhs()
dhsData<-getDHSdata(country = country, indicator = indicator, year = year)
geo <- getDHSgeo(country = country, year = year)
```

If you wish to download the data as a ZIP instead, download the dataset you want to analyze from your project page on DHS, as well as the geographic data (the .shp file at the bottom after you click on a specific set of datasets). Then, load the data in as follows:

```
# load data locally
file_path <- "../ZMIR71DT/ZMIR71FL.DTA" # put your filepath here for data
dhsData <- as.data.frame(read_dta(file_path))
file_path_geo<- "../ZMGE71FL/ZMGE71FL.shp" # put your shapefile filepath here
geo <- st_read(file_path_geo)
```

Regardless of which method you choose, to get your data in a more manageable format with the indicator you desire as well as get the info for Admin 1 and Admin 2 levels, run the following code:

```
#get dataframe with your indicator
data <- getDHSindicator(dhsData, indicator = indicator)
head(data)

#get your geographic data
geo <- getDHSgeo(country = country, year = year)

#get your admin1, admin2 geographic level data
poly.adm1 <- geodata::gadm(country="ZMB", level=1, path=tempdir())
poly.adm1 <- sf::st_as_sf(poly.adm1)
poly.adm2 <- geodata::gadm(country="ZMB", level=2, path=tempdir())
poly.adm2 <- sf::st_as_sf(poly.adm2)
```

Continue following the surveyPrev vignette in order to do further analysis, and please reach out or make a discussion post if anything doesn't work!