```
## FEBRUARY 2024
#############################################################################
#==========================
#Set working directory
> setwd("D:/DHS2016/ACIPH")
### DHS 2016 Analysis 2024
# LOAD AND PREPARE DATA ###################################
# INSTALL AND LOAD PACKAGES ###############################

# Install the pacman ("package manager")
if (!require("pacman")) install.packages("pacman")

# Load contributed packages with pacman
pacman::p_load(pacman, party, psych, rio, tidyverse)
# pacman: for loading/unloading packages
# party: for decision trees
# psych: for many statistical procedures
# rio: for importing data
# tidyverse: for so many additional reasons

# Load base packages manually
library(datasets)  # For example datasets

# LOAD AND PREPARE DATA ###################################

# Save data to "df" (for "data frame")
# Import CSV files with readr::read_csv() from tidyverse
##(df <- read_csv("data/StateData.csv"))

# Import other formats with rio::import() from rio
##(df <- import("data/StateData.xlsx") %>% as_tibble())

## Read 2016 DHS Stata data into R
## Data source https://www.dhsprogram.com/data/dataset_admin/index.cfm
## Lognin info required
## Use the foreign package to read Stata data from the DHS website

library(foreign)

#Bring in Births recode using read.data
DHS2016_Births <- read.dta("D:/DHS2016/ETBR71DT/ETBR71FL.DTA")
```

```
#Know the dimensions - number of rows and columns
dim(DHS2016_Births)
# [1] 41392  1287

#Bring in couples' recode using read.data
DHS2016_Couple <- read.dta("D:/DHS2016/ETCR71DT/ETCR71FL.DTA")
#Know the dimensions - number of rows and columns
dim(DHS2016_Couple)
#[1] 6141 2528

#Bring in household recode using read.data
DHS2016_HHold <- read.dta("D:/DHS2016/ETHR71DT/ETHR71FL.DTA")
#Know the dimensions - number of rows and columns
dim(DHS2016_HHold)
#[1] 16650  1966

#Bringing in individual woman recode using read.data
DHS2016_Wom <- read.dta("D:/DHS2016/ETIR71DT/ETIR71FL.DTA")
#Know the dimensions - number of rows and columns
dim(DHS2016_Wom)
#[1] 15683       5902
##############################################################################
### AN ALTERNATIVE WAY OF BRINGING IN EXTERNAL DATA

# Import data with rio::import() from rio
##(df <- import("data/StateData.xlsx") %>% as_tibble())
# ############################################################################
##############################################################################

# DATA REDUCTION: Selecting only the data columns you need
# The DHS_2016Wom file has 15683  rows and 5902 columns/fields/variables
#We want to select only FIVE variables to demonstrate steps followed for analyzing data used in a recent paper titled: The Social Geography of Women's Attitudes Towards Wife-bearing in Ethiopia: A Contribution Toward Proper Application of Spatial Statistics
DHS2016_WomReduced <- DHS2016_Wom %>% select(v001, v101, v005, v106, v744c) %>%
  print()
dim(DHS2016_WomReduced)

##############################################################################
##############################################################################

#Unweighted CLUSTER frequency and proportions of women responding "yes" or "no" to whether or
```

```
# not wife-beating was okay

ClasterFreqWB <- table(DHS2016_WomReduced$v001,DHS2016_Wom$v744c) # A will be rows, B will be columns
ClasterFreqWB # print table
## Calculate the proportions of women responding "yes" or "no"
prop.table (ClasterFreqWB, 1)
## Save UNWEIGHTED DHS2016_Wom wife-beating proportions to the DHS2016 Analysis2024 folder as CSV
write.csv(prop.table (ClasterFreqWB, 1), file = "UnweightedCLWBProp.csv")

##############################################################################

#Unweighted REGION frequency and proportions of women responding "yes" or "no" to whether or
# not wif-beating was okay

RegionFreqWB <- table(DHS2016_WomReduced$v101,DHS2016_Wom$v744c) # A will be rows, B will be columns
RegionFreqWB # print table

## Calculate the proportions of women responding "yes" or "no"
prop.table (RegionFreqWB, 1)

## Save UNWEIGHTED DHS2016_Wom wife-beating proportions to the DHS2016 Analysis2024 folder as CSV
write.csv(prop.table (RegionFreqWB, 1), file = "UnweightedRegionWBProp.csv")

##############################################################################

# WEIGHTED DATA: How to use weights
#weights example
#use questionr package

library(questionr)

#Eth DHS, weights example, SAMPLING CLUSTER-level
#WtdCLWB_DHS16_Wom2 : weighted sampling cluster-level wife beating frequency
# women's record file from Rio import ...DHS16_Wom2
#normwt = FALSE means do not normalize the weights
#na.rm = TRUE means delete records labeled NA (not applicable)
#na.show = FALSE means do not show removed NAs
```

```
#Eth DHS, weights example, SAMPLING CLUSTER-level (wife-beating)

WtdCLWifeBeating <- wtd.table(DHS2016_WomReduced$v001, y=DHS2016_WomReduced$v744c,
            weights = DHS2016_WomReduced$v005/1000000, ##DHS advises dividing weights by a million
                normwt = FALSE, # do not normalize the weight   na.rm = TRUE,   #rewmove NAs
                na.show = FALSE)
WtdCLWifeBeating

## Save WtdCLWB_DHS16_Wom to a folder as CSV
write.csv(prop.table (WtdCLWifeBeating, 1), file = "WeightedClusterWBProp.csv")

#############################################################################
WtdRegionWifeBeating <- wtd.table(DHS2016_WomReduced$v101, y=DHS2016_WomReduced$v744c,
                 weights = DHS2016_WomReduced$v005/1000000,
                 normwt = FALSE, na.rm = TRUE,
                 na.show = FALSE)
WtdRegionWifeBeating
## Save WtdCLWB_DHS16_Wom to a folder as CSV
write.csv(prop.table (WtdRegionWifeBeating, 1), file = "WeightedRegionWBProp.csv")

#############################################################################
#HOUSEHOLD RECODE, February 2014
#############################################################################

# Data reduction applied to DHS Household recode file: select only the fields or columns needed for
analysis
#CLUSTER-level analysis

#Household recode
DHS2016_HHold <- read.dta("D:/DHS2016/ETHR71DT/ETHR71FL.DTA")
dim(DHS2016_HHold)
#[1] 16650  1966

## Household file data reduction
# The DHS_2016HHold file has 16650  rows and 1966 columns/fields/variables
# We would like to choose just six variables to conduct background analysis 1) HV001: cluster number, 2)
HV201:  source of drinking water, 3) HV205:  type of toilet facility, 4) HV206: household has electricity
(yes, no), 5) HV207: radio ownership, 6) HV005 sample weight

DHS2016_HHReduced <- DHS2016_HHold %>% select(hv001, hv005, hv201, hv205, hv206,
hv207,hv270) %>%
  print()
```

```
dim(DHS2016_HHReduced)

#WtdWatrDHS2016_HHold: weighted data SOURCE OF WATER
WtdWatrDHS2016_HHold <- wtd.table(DHS2016_HHReduced$hv001, y=DHS2016_HHReduced$hv201,
                        weights = DHS2016_HHReduced$hv005/1000000,
                         normwt = FALSE, na.rm = TRUE,
                         na.show = FALSE)

## Save WtdRegED_DHS2016_Wom2 (region-level wife-beating, weighted) to a folder as CSV
write.csv(prop.table (WtdWatrDHS2016_HHold, 1), file = "WeightedClusterWaterSourceProp.csv")

#############################################################################
#DHS2016_HHoldWTTlt, weighed cluster-level type of toilet
#WtdDHS2016_HHoldWTTlt: weighted data TOILET FACILITIES

WtdDHS2016_HHoldWTTlt <- wtd.table(DHS2016_HHReduced$hv001, y=DHS2016_HHReduced$hv205,
                         weights = DHS2016_HHReduced$hv005/1000000,
                        normwt = FALSE, na.rm = TRUE,
                        na.show = FALSE)

## Save WtdDHS2016_HHoldWTTlt (region-level toilet facilities to a folder as CSV
write.csv(prop.table (WtdDHS2016_HHoldWTTlt, 1), file = "WeightedClusterToiletTypeProp.csv")

#############################################################################
#WtdDHS2016_HHoldWTRadio, weighed cluster-level household has RADIO
#WtdDHS2016_HHoldWTRadio: weighted data Radio Ownership

WtdDHS2016_HHoldWTRadio <- wtd.table(DHS2016_HHReduced$hv001,
y=DHS2016_HHReduced$hv207,
                          weights = DHS2016_HHReduced$hv005/1000000,
                         normwt = FALSE, na.rm = TRUE,
                         na.show = FALSE)

## Save WtdDHS2016_HHoldWTTlt (region-level toilet facilities to a folder as CSV
write.csv(prop.table (WtdDHS2016_HHoldWTRadio, 1), file =
"WeightedClusterRadioOwnershipProp.csv")

#############################################################################
```

```
#DHS2016_HHoldWTElectric, weighed cluster-level household has Electricity

#DHS2016_HHoldWTElectric: weighted data TOILET FACILITIES
DHS2016_HHoldWTElectric <- wtd.table(DHS2016_HHReduced$hv001,
y=DHS2016_HHReduced$hv206,
                    weights = DHS2016_HHReduced$hv005/1000000,
                    normwt = FALSE, na.rm = TRUE,
                    na.show = FALSE)

## Save WtdDHS2016_HHoldWTTlt (region-level toilet facilities to a folder as CSV
write.csv(prop.table (DHS2016_HHoldWTElectric, 1), file = "WeightedClusterElectricityProp.csv")

##############################################################################
#DHS2016_HHoldWealthIndex, weighed cluster-level household WealthIndex

DHS2016_HHoldWealthIndex <- wtd.table(DHS2016_HHReduced$hv001,
y=DHS2016_HHReduced$hv270,
                     weights = DHS2016_HHReduced$hv005/1000,
                    normwt = FALSE, na.rm = TRUE,
                    na.show = FALSE)

## Save WtdDHS2016_HHoldWTTlt (region-level toilet facilities to a folder as CSV
write.csv(prop.table (DHS2016_HHoldWealthIndex, 1), file = "WeightedClusterWealthIndexProp.csv")

##############################################################################
```