Statistics/Data analysis

```
//  Program:    02-dataPrep-hhMemberDemogVars.do
//  Task:       cleans household members' demographic variables using Kenyan survey
> data (2015/16 KIHBS)
//  Project:    Kenya Fical MicroSim
//  Author:     Yared Seid - 2023Jan11


// #1
// Loading the data
use "${dataSources}\Kenya_2015_16_KIHBS\HH_Members_Information", clear


// #2
// Generating vars

* Generating IDs
codebook clid hhid

qui tostring clid,      gen(clusterID) format(%04.0f)                    // cluster i
> d
qui tostring hhid,      gen(hhid_STR) format(%02.0f)                     // houdehold
>  id, within a cluster
qui tostring b01,       gen(b01_STR) format(%02.0f)                              // h
> oudehold member id, within a household

label var clusterID "Cluster ID"

gen hhID = clusterID + hhid_STR
label var hhID "Household ID"

gen memberID = hhID + b01_STR
label var memberID "HH Member ID"
isid memberID


* age
gen date_birth = mdy(b06_mm, b06_dd, b06_yy)                                 // d
> ate of birth
format date_birth %td

merge m:1 hhID using "${DataCleaningPath}\01-dataPrep-hhDemogVars", keepusing(county
>  date_interview)              // importing interview date

gen age = round((date_interview - date_birth)/365, 0.1)                  // age as th
> e diff b/n survay/interview date and birth date. I didn't make use of ages in mont
> hs since its value contradicts with that reported in age in years.
label var age "Age in years"
replace age = b05_yy if missing(age)                     // if a more accurate age (i
> .e., YY MM YY) is missing, replace it by less accurate age (i.e., YY)

* gender
recode b04 (1=0) (2=1), gen(female)
label var female "=1 if female"

* hh member relationship to the head
clonevar relationship = b03
label var relationship "HH member's relationship to the head"



// #3
// Saving the data
local keepVars county clid clusterID hhID memberID age female relationship
```

```
/* To assign Yes/No value lable to dummy variables */
local yesnoVars          female

order `keepVars'
keep `keepVars'

sort  memberID
codebook hhID memberID

* saving details
local dta_name           02-dataPrep-hhMemberDemogVars
local dta_note          "Household members' demographic variables"

include "${DataCleaningPath}\i-dta-savingDetails.do"    // saving the data with addi
> tional info on notes, labels, etc
*************************


*******************************************************************************
// #4
// Post-data-saving checking

/* Checking for the hh size (we count only those who are present) */
bysort hhID: egen double hh_size_test = count(memberID)
        assert hh_size == hh_size_test // if !mi(ind_weight)

*******************************************************************************
```