

Anexo 2.2. Creacion de variables ENIGH

Juan Carlos Martinez-Ovando

ITAM

Se presenta y comenta el código para la creación de las variables de análisis del modelo MUM y propuesto a partir de los datos de la ENIGH 2016 compatibles con la información proporcionada por los datos CUIS.

Parte I: Creacion de Variables

Preamble

Las bases de datos fueron obtenidas del INEGI (www.inegi.org.mx) en formato *.dbf de la liga:

<http://www.beta.inegi.org.mx/proyectos/enchogares/regulares/enigh/nc/2016/>

Nota.- Los scripts de este documento fueron probados en R version > 3.4.x. (Linux y Windows x64).

Los scripts de este documento hacen uso de las siguientes tablas:

- Concentrado hogar: `concentradohogar.dbf`
- Base hogares: `hogares.dbf`
- Base poblacion: `poblacion.dbf`
- Base laboral: `trabajos.dbf`
- Base de viviendas: `viviendas.dbf`

Las rutas de las bases de datos correspondientes al proyecto `Modelo.Distribucion.Ingreso.2017` son:

1. Bases originales: `./BasesDatos/ENIGH.2016`
2. Bases generadas: `./Bases4Analysis`

Paquetes de R

Se cargan los paquetes necesarios de R de manera oculta. En el archivo *.Rmd se encuentra el script para la verificación e instalación de los paquetes.

Fijamos, adicionalmente, la ruta de la carpeta donde los datos fuente del INEGI se localizan.

=====

Preparacion de los datos

La preparación de los datos sigue las especificaciones y reglas definidas por el CONEVAL (www.coneval.org.mx).

Parte I. Preparacion

Importancion de las tablas

```
concentradohogar <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/concentradohogar.dbf",
                             as.is = TRUE)
dim(concentradohogar)
```

```
## [1] 70311 127
```

Creacion de tabla base

```
aux_var <- c("tam_loc",
            "sexo_jefe", "edad_jefe", "educa_jefe",
            "tot_integ", "p12_64", "p65mas", "remesas",
            "ing_cor")
hogares_agr <- concentradohogar[,c("folioviv", "foliohog", aux_var)]
colnames(hogares_agr) <- c("FOLIOVIV", "FOLIOHOG", aux_var)
dim(hogares_agr)
```

```
## [1] 70311 11
```

```
write.csv(hogares_agr, file="../Bases.Enigh/2016_a_hogares_enigh.csv")
rm(concentradohogar)
rm(aux_var)
```

Preparacion de tablas

Jefe/conyuge

```
poblacion <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/poblacion.dbf",
                      as.is = TRUE)
dim(poblacion)
```

```
## [1] 257805 178
```

```
colnames(poblacion) <- toupper(colnames(poblacion))

# Jefe y conyuge
poblacion$es_jefe <- 0
poblacion$es_cony <- 0
poblacion[which(poblacion$PARENTESCO==101), "es_jefe"] <- 1
poblacion[which(poblacion$PARENTESCO==201), "es_cony"] <- 1
table(poblacion[,c("es_jefe", "es_cony")])
```

```
##          es_cony
## es_jefe      0      1
##          0 139654  47840
##          1  70311      0
```

```
poblacion_jefe <- poblacion[which( poblacion$es_jefe==1 | poblacion$es_cony==1),]
dim(poblacion_jefe)
```

```
## [1] 118151      180
```

```
# Escolaridad Jefe y Conyuge
poblacion_jefe$p_esc3 <- 0
poblacion_jefe[which( (poblacion$PARENTESCO==101 | poblacion$PARENTESCO==201) &
                      poblacion$EDAD>=12 &
                      ( poblacion$NIVELAPROB>=2 & poblacion$NIVELAPROB<3 ) ), "p_esc3"] <- 1

poblacion_jefe$p_esc5b <- 0
poblacion_jefe[which( (poblacion$PARENTESCO==101 | poblacion$PARENTESCO==201) &
                      poblacion$EDAD>=12 &
                      poblacion$NIVELAPROB>=3 ), "p_esc5b"] <- 1

# Agregago
aux_var <- c("p_esc3", "p_esc5b")
poblacion_jefe <- aggregate( x = poblacion_jefe[,aux_var],
                             by = list(poblacion_jefe$FOLIOVIV, poblacion_jefe$FOLIOHOG),
                             FUN = max )
colnames(poblacion_jefe) <- c("FOLIOVIV", "FOLIOHOG", aux_var)
dim(poblacion_jefe)
```

```
## [1] 70311      4
```

```
rm(poblacion)
rm(aux_var)
```

Poblacion

```
poblacion <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/poblacion.dbf",
                      as.is = TRUE)
dim(poblacion)
```

```
## [1] 257805      178
```

```
colnames(poblacion) <- toupper(colnames(poblacion))
```

```
# Grupos de edad
```

```

poblacion$int0a12 <- 0
poblacion$int12a64 <- 0
poblacion$int65a98 <- 0
poblacion[which(poblacion$EDAD>0 & poblacion$EDAD<12 ), "int0a12"] <- 1
poblacion[which(poblacion$EDAD>=12 & poblacion$EDAD<=64 ), "int12a64"] <- 1
poblacion[which(poblacion$EDAD>=65 & poblacion$EDAD<=98 ), "int65a98"] <- 1

# Mujeres
poblacion$muj12a49 <- 0
poblacion[which( poblacion$SEXO==2 &
                 (poblacion$EDAD>=12 & poblacion$EDAD<49) ), "muj12a49"] <- 1

# Seguro Popular
poblacion$seg_pop <- 0
poblacion[which( poblacion$SEGPPOP==1 ), "seg_pop"] <- 1

# Atencion Medica
poblacion$ss <- 0
poblacion[which( poblacion$ATEMED==1 ), "ss"] <- 1

# Agregago
aux_var <- c("int0a12", "int12a64", "int65a98", "muj12a49", "seg_pop", "ss")
poblacion_agr <- aggregate( x = poblacion[,aux_var],
                           by = list(poblacion$FOLIOVIV, poblacion$FOLIOHOG),
                           FUN = sum )
colnames(poblacion_agr) <- c("FOLIOVIV", "FOLIOHOG", aux_var)
dim(poblacion_agr)

```

```
## [1] 70311      8
```

```
rm(poblacion)
rm(aux_var)
```

Hogares

```

hogares <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/hogares.dbf",
                  as.is = TRUE)
dim(hogares)

```

```
## [1] 70311    137
```

```

colnames(hogares) <- toupper(colnames(hogares))

# Seguridad Alimentaria 3 (Desayuno, comida o cena)
table(hogares$ACC ALIM5)

```

```
##
##      1      2
## 10195 60116
```

```
hogares$seg_alim2 <- 0
hogares[which(hogares$ACC_ALIM5==1), "seg_alim2"] <- 1

# Seguridad Alimentaria 3 (Comida)
table(hogares$ACC_ALIM2)
```

```
##
##      1      2
##  8895 61416
```

```
hogares$seg_alim3 <- 0
hogares[which(hogares$ACC_ALIM2==1), "seg_alim3"] <- 1

# Seguridad Alimentaria a (Comida)
table(hogares$ACC_ALIM8)
```

```
##
##      1      2
##  7788 26881
```

```
hogares$seg_alim_a <- 0
hogares[which(hogares$ACC_ALIM8==1), "seg_alim_a"] <- 1

# Refrigerador
table(hogares$NUM_REFRI)
```

```
##
##      0      1      2      3      4      5      7      8     14     15
## 10021 59467   767   38      7      4      3      1      2      1
```

```
hogares$sin_refri <- 0
hogares[which( hogares$NUM_REFRI==0 ), "sin_refri"] <- 1
table(hogares$sin_refri)
```

```
##
##      0      1
## 60290 10021
```

```
# Vehiculos
hogares$num_vehi <- hogares$NUM_AUTO + hogares$NUM_VAN + hogares$NUM_PICKUP
table(hogares$num_vehi)
```

```
##
##      -1      0      1      2      3      4      5      6      7      8      9     10
##      1 38655 23890 6123  1229   268    83    30    11     7     4     5
##      12     14     20
##      3      1      1
```

```
hogares$sin_vehi <- 0
hogares[which( hogares$num_vehi==0 ), "sin_vehi"] <- 1
table(hogares$sin_vehi)
```

```
##
##      0      1
## 31656 38655
```

```
# Computadora
table(hogares$NUM_COMPU)
```

```
##
##      0      1      2      3      4      5      6      7      8      9     15
## 53011 14437  1979   661   149   42    15     9     4     3     1
```

```
hogares$sin_compu <- 0
hogares[which( hogares$NUM_COMPU==0 ), "sin_compu"] <- 1
table(hogares$sin_compu)
```

```
##
##      0      1
## 17300 53011
```

```
# Video & DVD
hogares$num_vidvd <- hogares$NUM_DVD + hogares$NUM_VIDEO
table(hogares$num_vidvd)
```

```
##
##      0      1      2      3      4      5      6      7      8      9     12     15
## 45855 21282  2668   389   80    24     7     1     2     1     1     1
```

```
hogares$sin_vidvd <- 0
hogares[which( hogares$num_vidvd==0 ), "sin_vidvd"] <- 1
table(hogares$sin_vidvd)
```

```
##
##      0      1
## 24456 45855
```

```
# Telefono
table(hogares$TELEFONO)
```

```
##
##      1      2
## 20668 49643
```

```
hogares$sin_telef <- 1
hogares[which( hogares$TELEFONO==1 ), "sin_telef"] <- 0
table(hogares$sin_telef)
```

```
##
##      0      1
## 20668 49643
```

```
# Horno (microondas)
table(hogares$NUM_MICRO)
```

```
##
##      0      1      2      3      4      5      9     11     13     15
## 41067 29066   161    10     1     1     1     1     2     1
```

```
hogares$sin_horno <- 0
hogares[which( hogares$NUM_MICRO==0 ), "sin_horno"] <- 1
table(hogares$sin_horno)
```

```
##
##      0      1
## 29244 41067
```

```
# Agregado
aux_var <- c("seg_alim2", "seg_alim3", "seg_alim_a", "sin_refri", "sin_vehi", "sin_vidvd", "sin_compu")
hogares_agr_agr <- aggregate( x = hogares[,aux_var],
                             by = list(hogares$FOLIOVIV, hogares$FOLIOHOG),
                             FUN = max )
colnames(hogares_agr_agr) <- c("FOLIOVIV", "FOLIOHOG", aux_var)
summary(hogares_agr_agr$seg_alim2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000   0.000   0.000   0.145   0.000   1.000
```

```
summary(hogares_agr_agr$sin_refri)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  0.0000  0.1425  0.0000  1.0000
```

```
summary(hogares_agr_agr$sin_vehi)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000  0.0000  1.0000  0.5498  1.0000  1.0000
```

```
summary(hogares_agr_agr$sin_compu)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.000  1.000  1.000  0.754  1.000  1.000
```

```
summary(hogares_agr_agr$sin_vidvd)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000  0.0000  1.0000  0.6522  1.0000  1.0000
```

```
summary(hogares_agr_agr$sin_telef)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.000  0.000  1.000  0.706  1.000  1.000
```

```
summary(hogares_agr_agr$sin_horno)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000  0.0000  1.0000  0.5841  1.0000  1.0000
```

```
rm(hogares)
```

Trabajos

```
trabajos <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/trabajos.dbf",
                    as.is = TRUE)
```

```
dim(trabajos)
```

```
## [1] 132095    55
```

```
colnames(trabajos) <- toupper(colnames(trabajos))
```

```
poblacion <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/poblacion.dbf",
                    as.is = TRUE)
```

```
dim(poblacion)
```

```
## [1] 257805    178
```



```

colnames(poblacion) <- toupper(colnames(poblacion))

trabajos <- merge(trabajos,
                  poblacion[,c("FOLIOVIV", "FOLIOHOG",
                               "NUMREN", "EDAD", "PARENTESCO")],
                  by=c("FOLIOVIV", "FOLIOHOG", "NUMREN"),
                  all.x=TRUE)

# Trabajo subordinado
trabajos$trab_sub <- 0
trabajos[which( trabajos$EDAD >= 12 &
               trabajos$ID_TRABAJO==1 &
               #( trabajos$ID_TRABAJO==1 / trabajos$ID_TRABAJO==2 ) &
               trabajos$SUBOR==1 ), "trab_sub"] <- 1

table(trabajos$trab_sub)

##
##      0      1
## 39705 92390

# Trabajo independiendte
trabajos$trab_ind <- 0
trabajos[which( trabajos$EDAD >= 12 &
               trabajos$ID_TRABAJO==1 &
               #( trabajos$ID_TRABAJO==1 / trabajos$ID_TRABAJO==2 ) &
               trabajos$INDEP==1 ), "trab_ind"] <- 1

table(trabajos$trab_ind)

##
##      0      1
## 104849 27246

# Trabajo subordinado sin pago
trabajos$trab_s_pago <- 0
trabajos[which( trabajos$EDAD >= 12 &
               trabajos$ID_TRABAJO==1 &
               #( trabajos$ID_TRABAJO==1 / trabajos$ID_TRABAJO==2 ) &
               trabajos$SUBOR==1 &
               (trabajos$PAGO==2 | trabajos$PAGO==3) ), "trab_s_pago"] <- 1

table(trabajos$trab_s_pago)

##
##      0      1
## 126331 5764

```

```
# Trabajo independiendte / Jefe del hogar
trabajos$jtrab_ind <- 0
trabajos[which( trabajos$EDAD >= 12 &
               trabajos$PARENTESCO==101 &
               trabajos$ID_TRABAJO==1 &
               #( trabajos$ID_TRABAJO==1 / trabajos$ID_TRABAJO==2 ) &
               trabajos$INDEP==1 ), "jtrab_ind"] <- 1

table(trabajos$jtrab_ind)
```

```
##
##      0      1
## 116462 15633
```

```
# Agregago
aux_var <- c("trab_sub", "trab_ind", "jtrab_ind", "trab_s_pago")
trabajos_agr <- aggregate( x = trabajos[,aux_var],
                           by = list(trabajos$FOLIOVIV, trabajos$FOLIOHOG),
                           FUN = sum )
colnames(trabajos_agr) <- c("FOLIOVIV", "FOLIOHOG", aux_var)
summary(trabajos_agr$trab_sub)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000   1.000   1.000   1.441   2.000   12.000
```

```
summary(trabajos_agr$trab_ind)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000   0.000   0.000   0.425   1.000   6.000
```

```
summary(trabajos_agr$trab_s_pago)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.00000 0.00000 0.08991 0.00000 10.00000
```

```
rm(trabajos)
rm(poblacion)
```

Viviendas

```
viviendas <- read.dbf("../Bases.Datos.Original/ENIGH2016_Coneval/Microdatos/viviendas.dbf",
                     as.is = TRUE)
dim(viviendas)
```

```
## [1] 69169    64
```

```
colnames(viviendas) <- toupper(colnames(viviendas))
```

```
# Vivienda propia
```

```
table(viviendas$TENENCIA)
```

```
##
```

```
##      1      2      3      4      5      6
## 8440  9534  6768 42855  1057   515
```

```
viviendas$viv_prop <- 0
```

```
viviendas[which( viviendas$TENENCIA==3 | viviendas$TENENCIA==4 ),"viv_prop"] <- 1
```

```
# Vivienda rentada
```

```
viviendas$viv_rent <- 0
```

```
viviendas[which( viviendas$TENENCIA==1 ),"viv_rent"] <- 1
```

```
# Piso firme
```

```
table(viviendas$MAT_PISOS)
```

```
##
```

```
##      &      1      2      3
##      4  2221 37986 28958
```

```
viviendas$piso_fir <- 0
```

```
viviendas[which( viviendas$MAT_PISOS==2 ),"piso_fir"] <- 1
```

```
# Piso con recubrimiento
```

```
viviendas$piso_rec <- 0
```

```
viviendas[which( viviendas$MAT_PISOS==3 ),"piso_rec"] <- 1
```

```
# Piso con recubrimiento
```

```
viviendas$tot_cuar <- viviendas$NUM_CUARTO
```

```
# Combustible
```

```
table(viviendas$COMBUSTIBL)
```

```
##
```

```
##      1      2      3      4      5      6
## 11489   209 52571  3613   760   527
```

```
viviendas$combustible <- 0
```

```
viviendas[which( viviendas$COMBUSTIBL==1 | viviendas$COMBUSTIBL==2 ),"combustible"] <- 1
```

```
table(viviendas$combustible)
```

```
##
```

```
##      0      1
## 57471 11698
```

```
#Escusado exclusivo con descarga de agua
table(viviendas[,c("EXCUSADO", "USO_COMPAR", "SANIT_AGUA")])
```

```
## , , SANIT_AGUA = 1
##
##      USO_COMPAR
## EXCUSADO      1      2
##      1  1984 43394
##      2      0      0
##
## , , SANIT_AGUA = 2
##
##      USO_COMPAR
## EXCUSADO      1      2
##      1  2105 14878
##      2      0      0
##
## , , SANIT_AGUA = 3
##
##      USO_COMPAR
## EXCUSADO      1      2
##      1   772 4007
##      2      0      0
```

```
viviendas$bao13 <- 0
viviendas[which( viviendas$EXCUSADO==1 &
                  viviendas$USO_COMPAR==2 &
                  viviendas$SANIT_AGUA== 1 ), "bao13"] <- 1
table(viviendas$bao13)
```

```
##
##      0      1
## 25775 43394
```

```
# Agregacion
aux_var <- c("viv_prop", "viv_rent", "piso_fir", "piso_rec", "tot_cuar", "combustible", "bao13")

viviendas_agr <- viviendas[,c("FOLIOVIV", aux_var)]

rm(viviendas)
```

Parte II. Calculo de variables para modelacion

1. Dependencia demografica (numerica)

depdemog

```
# v01 depdemog

# Se obtiene directamente de 'poblacion_agr'
poblacion_agr$depdemog <- 0
poblacion_agr[which(poblacion_agr$int12a64>0),"depdemog"] <-
  (poblacion_agr[which(poblacion_agr$int12a64>0),"int0a12"] +
    poblacion_agr[which(poblacion_agr$int12a64>0),"int65a98"]) /
  poblacion_agr[which(poblacion_agr$int12a64>0),"int12a64"]
summary(poblacion_agr$depdemog)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000 0.0000 0.2500 0.4199 0.6667 6.0000

aux_var <- c("int0a12","int12a64","int65a98","depdemog")

hogares_agr <- merge(hogares_agr,
  poblacion_agr[,c("FOLIOVIV","FOLIOHOG",aux_var)],
  by=c("FOLIOVIV","FOLIOHOG"))
dim(hogares_agr)

## [1] 70311    15

colnames(hogares_agr)

## [1] "FOLIOVIV" "FOLIOHOG" "tam_loc"   "sexo_jefe" "edad_jefe"
## [6] "educa_jefe" "tot_integ" "p12_64"   "p65mas"   "remesas"
## [11] "ing_cor" "int0a12" "int12a64" "int65a98" "depdemog"
```

2. Mujeres entre 12 y 49 años (numero de personas)

muj12a49

```
# v02 muj12a49
# Trabajamos con la tabla 'poblacion_agr'
summary(poblacion_agr$muj12a49)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00     0.00     1.00     1.06     2.00     8.00

aux_var <- c("muj12a49")

hogares_agr <- merge(hogares_agr,
  poblacion_agr[,c("FOLIOVIV","FOLIOHOG",aux_var)],
  by=c("FOLIOVIV","FOLIOHOG"))
dim(hogares_agr)

## [1] 70311    16
```

```
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV" "FOLIOHOG" "tam_loc" "sexo_jefe" "edad_jefe"
## [6] "educa_jefe" "tot_integ" "p12_64" "p65mas" "remesas"
## [11] "ing_cor" "int0a12" "int12a64" "int65a98" "depdemog"
## [16] "muj12a49"
```

3. Total de personas (numero de personas en log-scale)

```
# v03 ltot_per
# Trabajamos con la tabla 'hogares_agr'
summary(hogares_agr$tot_integ)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.000   2.000   4.000   3.665   5.000  21.000
```

```
summary(hogares_agr$int0a12 + hogares_agr$int12a64 + hogares_agr$int65a98)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000   2.000   3.000   3.609   5.000  20.000
```

```
hogares_agr$tot_per <- hogares_agr$int0a12 +
  hogares_agr$int12a64 + hogares_agr$int65a98
hogares_agr$ltot_per <- log(as.numeric(hogares_agr$tot_per))
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV" "FOLIOHOG" "tam_loc" "sexo_jefe" "edad_jefe"
## [6] "educa_jefe" "tot_integ" "p12_64" "p65mas" "remesas"
## [11] "ing_cor" "int0a12" "int12a64" "int65a98" "depdemog"
## [16] "muj12a49" "tot_per" "ltot_per"
```

4. Educacion del jefe del hogar y conyuge

p_esc3

```
# v04 p_esc3, p_esc4 and p_esc5
# Trabajamos con la tabla 'poblacion_jefe'
table(poblacion_jefe$p_esc3)
```

```
##
##      0      1
## 54340 15971
```

```
hogares_agr <- merge(hogares_agr,
  poblacion_jefe[,c("FOLIOVIV", "FOLIOHOG", "p_esc3")],
  by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$p_esc3)
```

```
##
##      0      1
## 54340 15971
```

5. Educacion del jefe del hogar y conyuge

p_esc5b

```
# v05 p_esc5b
# Trabajamos con la tabla 'poblacion_jefe'
table(poblacion_jefe$p_esc5b)
```

```
##
##      0      1
## 43244 27067
```

```
hogares_agr <- merge(hogares_agr,
                     poblacion_jefe[,c("FOLIOVIV", "FOLIOHOG", "p_esc5b")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$p_esc5b)
```

```
##
##      0      1
## 43244 27067
```

6. Trabajo subordinado (numero de personas)

trab_sub

```
# v06 trab_sub (numero)
# Traajamos con la tabla 'trabajos_agr'
summary(trabajos_agr$trab_sub)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000   1.000   1.000   1.441   2.000   12.000
```

```
aux_var <- c("trab_sub")

hogares_agr <- merge(hogares_agr,
                     trabajos_agr[,c("FOLIOVIV", "FOLIOHOG", aux_var)],
                     by=c("FOLIOVIV", "FOLIOHOG"))
dim(hogares_agr)
```

```
## [1] 64108    21
```

```
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV" "FOLIOHOG" "tam_loc" "sexo_jefe" "edad_jefe"
## [6] "educa_jefe" "tot_integ" "p12_64" "p65mas" "remesas"
## [11] "ing_cor" "int0a12" "int12a64" "int65a98" "depdemog"
## [16] "muj12a49" "tot_per" "ltot_per" "p_esc3" "p_esc5b"
## [21] "trab_sub"
```

7. Trabajo independiente (numero de personas)

```
trab_ind
```

```
# v07 trab_ind (numero)
# Traajamos con la tabla 'trabajos_agr'
summary(trabajos_agr$trab_ind)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000   0.000   0.000   0.425   1.000   6.000
```

```
summary(trabajos_agr$trab_s_pago)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.00000 0.00000 0.08991 0.00000 10.00000
```

```
aux_var <- c("trab_ind")
```

```
hogares_agr <- merge(hogares_agr,
                     trabajos_agr[,c("FOLIOVIV", "FOLIOHOG", aux_var)],
                     by=c("FOLIOVIV", "FOLIOHOG"))
dim(hogares_agr)
```

```
## [1] 64108    22
```

```
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV" "FOLIOHOG" "tam_loc" "sexo_jefe" "edad_jefe"
## [6] "educa_jefe" "tot_integ" "p12_64" "p65mas" "remesas"
## [11] "ing_cor" "int0a12" "int12a64" "int65a98" "depdemog"
## [16] "muj12a49" "tot_per" "ltot_per" "p_esc3" "p_esc5b"
## [21] "trab_sub" "trab_ind"
```

8. Trabajo subordinado sin pago (numero de personas)

```
trab_s_pago
```



```
# v08 trab_s_pag (numero)
# Traajamos con la tabla 'trabajos_agr'
summary(trabajos_agr$trab_s_pago)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.00000 0.00000 0.08991 0.00000 10.00000
```

```
aux_var <- c("trab_s_pago")

hogares_agr <- merge(hogares_agr,
                     trabajos_agr[,c("FOLIOVIV", "FOLIOHOG", aux_var)],
                     by=c("FOLIOVIV", "FOLIOHOG"))
dim(hogares_agr)
```

```
## [1] 64108    23
```

```
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV"    "FOLIOHOG"    "tam_loc"     "sexo_jefe"   "edad_jefe"
## [6] "educa_jefe"  "tot_integ"   "p12_64"      "p65mas"      "remesas"
## [11] "ing_cor"     "int0a12"     "int12a64"    "int65a98"    "depdemog"
## [16] "muj12a49"    "tot_per"     "ltot_per"    "p_esc3"      "p_esc5b"
## [21] "trab_sub"    "trab_ind"    "trab_s_pago"
```

9. Seguridad alimentaria 1

```
seg_alim2
```

```
# v09 seg_alim2 (indicadora)
# Trabajamos con la tabla 'hogares_agr_agr'
table(hogares_agr_agr$seg_alim2)
```

```
##
##      0      1
## 60116 10195
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "seg_alim2")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$seg_alim2)
```

```
##
##      0      1
## 54841  9267
```

10. Seguridad alimentaria 2

seg_alim3

```
# v10 seg_alim3 (indicadora)
# Trabajamos con la tabla 'hogares_agr_agr'
table(hogares_agr_agr$seg_alim3)
```

```
##
##      0      1
## 61416  8895
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "seg_alim3")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$seg_alim3)
```

```
##
##      0      1
## 56089  8019
```

11. Seguridad alimentaria conjunta

seg_alim_a2

```
# v11 seg_alim_a2 (indicadora)
# Trabajamos con la tabla 'hogares_agr_agr'
# Trabajamos con la tabla 'hogares_agr_agr'
table(hogares_agr_agr$seg_alim_a)
```

```
##
##      0      1
## 62523  7788
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "seg_alim_a")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$seg_alim_a)
```

```
##
##      0      1
## 57032  7076
```

12. Acceso a seguro popular (numero de personas)

seg_pop

```
# v12 seg_pop (numero)
table(poblacion_agr$seg_pop)
```

```
##
##      0      1      2      3      4      5      6      7      8      9     10     11
## 29760  8997  8412  7672  7133  4575  2052   946   405   177    94    46
##      12     13     14     15
##      24     11      4      3
```

```
aux_var <- c("seg_pop")

hogares_agr <- merge(hogares_agr,
                     poblacion_agr[,c("FOLIOVIV", "FOLIOHOG", aux_var)],
                     by=c("FOLIOVIV", "FOLIOHOG"))
dim(hogares_agr)
```

```
## [1] 64108    27
```

```
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV" "FOLIOHOG" "tam_loc" "sexo_jefe" "edad_jefe"
## [6] "educa_jefe" "tot_integ" "p12_64" "p65mas" "remesas"
## [11] "ing_cor" "int0a12" "int12a64" "int65a98" "depdemog"
## [16] "muj12a49" "tot_per" "ltot_per" "p_esc3" "p_esc5b"
## [21] "trab_sub" "trab_ind" "trab_s_pago" "seg_alim2" "seg_alim3"
## [26] "seg_alim_a" "seg_pop"
```

13. Servicio medico

```
ss
```

```
# v13 ss (indicadora)
table(poblacion_agr$ss)
```

```
##
##      0      1      2      3      4      5      6      7      8      9     10     11
## 28405 11908 10239  7812  6830  3416  1133   375   112    53    16     6
##      12     13
##      3      3
```

```
aux_var <- c("ss")

hogares_agr <- merge(hogares_agr,
                     poblacion_agr[,c("FOLIOVIV", "FOLIOHOG", aux_var)],
                     by=c("FOLIOVIV", "FOLIOHOG"))
dim(hogares_agr)
```

```
## [1] 64108    28
```

```
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV"    "FOLIOHOG"    "tam_loc"     "sexo_jefe"   "edad_jefe"
## [6] "educa_jefe"  "tot_integ"   "p12_64"      "p65mas"      "remesas"
## [11] "ing_cor"     "int0a12"     "int12a64"    "int65a98"    "depdemog"
## [16] "muj12a49"    "tot_per"     "ltot_per"    "p_esc3"      "p_esc5b"
## [21] "trab_sub"    "trab_ind"    "trab_s_pago" "seg_alim2"   "seg_alim3"
## [26] "seg_alim_a"  "seg_pop"     "ss"
```

14. Seguridad social por trabajo independiente

```
ssjtrabind
```

```
# v14 ssjtrabind (indicadora)
table(hogares_agr$ss)
```

```
##
##      0      1      2      3      4      5      6      7      8      9     10     11
## 25730 10076  8891  7575  6737  3401  1130   375   112    53    16     6
##      12     13
##       3      3
```

```
table(trabajos_agr$jtrab_ind)
```

```
##
##      0      1
## 48475 15633
```

```
hogares_agr <- merge(hogares_agr,
                     trabajos_agr[,c("FOLIOVIV", "FOLIOHOG", "jtrab_ind")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
```

```
hogares_agr$ssjtrabind <- 0
hogares_agr[which( hogares_agr$ss==1 &
                  hogares_agr$jtrab_ind==1 ), "ssjtrabind"] <- 1
table(hogares_agr$ssjtrabind)
```

```
##
##      0      1
## 61734  2374
```

15. Recepcion de remesas

```
con_remasas
```

```
# v30 remesa
summary(hogares_agr$remesas)
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.
##      0.0     0.0     0.0    325.7    0.0 173152.2
```

```
hogares_agr$con_remesas <- 0
hogares_agr[which( hogares_agr$remesas>0 ), "con_remesas"] <- 1
table(hogares_agr$con_remesas)
```

```
##
##      0      1
## 60820 3288
```

16. Vivienda propia

viv_prop

```
# v16 viv_prop (indicadora)
# Trabajamos con 'viviendas_agr'
table(viviendas_agr$viv_prop)
```

```
##
##      0      1
## 19546 49623
```

```
hogares_agr <- merge(hogares_agr,
                     viviendas_agr[,c("FOLIOVIV", "viv_prop")],
                     by=c("FOLIOVIV"))
table(hogares_agr$viv_prop)
```

```
##
##      0      1
## 18291 45817
```

17. Vivienda rentada

viv_rent

```
# v17 viv_rent (indicadora)
# Trabajamos con 'viviendas_agr'
table(viviendas_agr$viv_rent)
```

```
##
##      0      1
## 60729 8440
```

```
hogares_agr <- merge(hogares_agr,
                     viviendas_agr[,c("FOLIOVIV", "viv_rent")],
                     by=c("FOLIOVIV"))
table(hogares_agr$viv_rent)
```

```
##
##      0      1
## 56022  8086
```

18. Total de cuartos (numero)

tot_cuar

```
# v18 tot_cuar (numero)
table(viviendas_agr$tot_cuar)
```

```
##
##      1      2      3      4      5      6      7      8      9     10     11     12
## 4012 10646 18565 18435 10334 4311 1635  750  249  143  41  26
##      13     14     15     16     17     23
##      10      3      3      4      1      1
```

```
hogares_agr <- merge(hogares_agr,
                     viviendas_agr[,c("FOLIOVIV", "tot_cuar")],
                     by=c("FOLIOVIV"))
table(hogares_agr$tot_cuar)
```

```
##
##      1      2      3      4      5      6      7      8      9     10     11     12
## 3688 9816 17203 17134 9586 4005 1536  696  233  132  36  23
##      13     14     15     16     17     23
##      9      2      3      4      1      1
```

19. Tipo de banio

```
# v19 bao1, bao2 and bao3 (indicadoras)
```

20. Banio exclusivo

bao13

```
# v20 bao13 (indicadoras)
# Trabajamos con la tabla 'viviendas_agr'
table(viviendas_agr$bao13)
```

```
##
##      0      1
## 25775 43394
```

```
hogares_agr <- merge(hogares_agr,
                     viviendas_agr[,c("FOLIOVIV", "bao13")],
                     by=c("FOLIOVIV"))
table(hogares_agr$bao13)
```

```
##
##      0      1
## 24338 39770
```

21. Piso firme

piso_fir

```
# v21 piso_fir & piso_rec (indicadoras)
# Trabajamos con 'viviendas_agr'
table(viviendas_agr$piso_fir)
```

```
##
##      0      1
## 31183 37986
```

```
hogares_agr <- merge(hogares_agr,
                     viviendas_agr[,c("FOLIOVIV", "piso_fir")],
                     by=c("FOLIOVIV"))
table(hogares_agr$piso_fir)
```

```
##
##      0      1
## 28455 35653
```

22. Piso con recubrimiento

piso_rec

```
# v22 piso_fir & piso_rec (indicadoras)
# Trabajamos con 'viviendas_agr'
table(viviendas_agr$piso_rec)
```

```
##
##      0      1
## 40211 28958
```

```
hogares_agr <- merge(hogares_agr,
                     viviendas_agr[,c("FOLIOVIV", " piso_rec")],
                     by=c("FOLIOVIV"))
table(hogares_agr$ piso_rec)
```

```
##
##      0      1
## 37706 26402
```

23. Uso de combustible

combustible

```
# v223< combustible (indicadoras)
# Trabajamos con la tabla 'viviendas_agr'
table(viviendas_agr$combustible)
```

```
##
##      0      1
## 57471 11698
```

```
hogares_agr <- merge(x=hogares_agr,
                     y=viviendas_agr[,c("FOLIOVIV", "combustible")],
                     by=c("FOLIOVIV"))
table(hogares_agr$combustible)
```

```
##
##      0      1
## 52969 11139
```

24. Hogar sin refrigerador

sin_refri

```
# v24< sin_refri (indicadoras)
# Trabajamos con la base 'hogares_agr_agr'
table(hogares_agr_agr$sin_refri)
```

```
##
##      0      1
## 60290 10021
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "sin_refri")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$sin_refri)
```

```
##
##      0      1
## 54990  9118
```


25. Hogar sin vehiculo

sin_vehi

```
# v25 sin_vehi (indicadoras)
# Trabajamos con la base 'hogares_agr_agr'
table(hogares_agr_agr$sin_vehi)
```

```
##
##      0      1
## 31656 38655
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "sin_vehi")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$sin_vehi)
```

```
##
##      0      1
## 29617 34491
```

26. Hogar sin computadora

sin_compu

```
# v26 sin_compu (indicadoras)
# Trabajamos con la base 'hogares_agr_agr'
table(hogares_agr_agr$sin_compu)
```

```
##
##      0      1
## 17300 53011
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "sin_compu")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$sin_compu)
```

```
##
##      0      1
## 16278 47830
```

27. Hogar sin video ni dvd

sin_vidvd

```
# v27 sin_vidvd (indicadoras)
# Trabajamos con la base 'hogares_agr_agr'
table(hogares_agr_agr$sin_vidvd)
```

```
##
##      0      1
## 24456 45855
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV","FOLIOHOG","sin_vidvd")],
                     by=c("FOLIOVIV","FOLIOHOG"))
table(hogares_agr$sin_vidvd)
```

```
##
##      0      1
## 23055 41053
```

28. Hogar sin telefono

sin_telef

```
# v28 sin_telef (indicadoras)
# Trabajamos con la base 'hogares_agr_agr'
table(hogares_agr_agr$sin_telef)
```

```
##
##      0      1
## 20668 49643
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV","FOLIOHOG","sin_telef")],
                     by=c("FOLIOVIV","FOLIOHOG"))
table(hogares_agr$sin_telef)
```

```
##
##      0      1
## 18133 45975
```

29. Hogar sin horno

sin_horno

```
# v29 sin_horno (indicadoras)
# Trabajamos con la base 'hogares_agr_agr'
table(hogares_agr_agr$sin_horno)
```

```
##
##      0      1
## 29244 41067
```

```
hogares_agr <- merge(hogares_agr,
                     hogares_agr_agr[,c("FOLIOVIV", "FOLIOHOG", "sin_horno")],
                     by=c("FOLIOVIV", "FOLIOHOG"))
table(hogares_agr$sin_horno)
```

```
##
##      0      1
## 26566 37542
```

```
dim(hogares_agr)
```

```
## [1] 64108    44
```

Ingreso corriente total

ict y ictpc

```
# ict - Ingresp Corriente Total
# Trabajamos con la base '2016_a_bienestar_ingreso'
ict_agr <- read.csv("../Bases.Enigh/2016_a_bienestar_ingreso.csv", header=TRUE)
dim(ict_agr)
```

```
## [1] 70311    10
```

```
colnames(ict_agr) <- c("FOLIOVIV", "FOLIOHOG", "factor", "tam_loc", "rururb", "tamhogesc", "ict", "ictpc")
table(ict_agr[,c("plb_m", "plb")])
```

```
##      plb
## plb_m    0      1
##      0 40728 19791
##      1      0  9792
```

```
aux_var <- c("factor", "rururb", "tamhogesc", "ict", "ictpc")
ict_agr$ict <- ict_agr$ict + 1
ict_agr$ictpc <- ict_agr$ictpc + 1

hogares_agr <- merge(hogares_agr,
                     ict_agr[,c("FOLIOVIV", "FOLIOHOG", aux_var)],
                     by=c("FOLIOVIV", "FOLIOHOG"),
                     x.all=TRUE)
summary(hogares_agr$ictpc)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##          1    1483    2454   3828   4222 1917475
```

Variables numericas y categoricas (p/ modelo)

```
hogares_agr[,var_enighcuis_num] <- lapply(hogares_agr[,var_enighcuis_num],as.numeric)
hogares_agr[,var_enighcuis_cat] <- lapply(hogares_agr[,var_enighcuis_cat],factor)
```

Exportacion de tabla de trabajo

```
write.csv(hogares_agr,file="../Bases.Enigh/hogares_enigh_agr.csv")
colnames(hogares_agr)
```

```
## [1] "FOLIOVIV"      "FOLIOHOG"      "tam_loc"       "sexo_jefe"     "edad_jefe"
## [6] "educa_jefe"    "tot_integ"     "p12_64"        "p65mas"        "remesas"
## [11] "ing_cor"       "int0a12"       "int12a64"      "int65a98"      "depdemog"
## [16] "muj12a49"      "tot_per"       "ltot_per"      "p_esc3"        "p_esc5b"
## [21] "trab_sub"      "trab_ind"      "trab_s_pago"   "seg_alim2"     "seg_alim3"
## [26] "seg_alim_a"    "seg_pop"       "ss"            "jtrab_ind"     "ssjtrabind"
## [31] "con_remesas"   "viv_prop"      "viv_rent"      "tot_cuar"      "bao13"
## [36] "piso_fir"      "piso_rec"      "combustible"   "sin_refri"     "sin_vehi"
## [41] "sin_compu"     "sin_vidvd"     "sin_telef"     "sin_horno"     "factor"
## [46] "rururb"        "tamhogesc"     "ict"           "ictpc"
```

```
dim(hogares_agr)
```

```
## [1] 43609    49
```

```
rm(list=ls())
gc()
```

```
##          used (Mb) gc trigger (Mb) max used (Mb)
## Ncells 1978327 105.7   3886542 207.6   3886542 207.6
## Vcells 2462228  18.8   93506304 713.4 146079537 1114.5
```