The Mexican Health and Aging Study (MHAS/ENASEM):
Master Follow-up File 2001, 2003, 2012, 2015 and 2018

Version 1

July 22, 2020

Table of Contents

I. Introduction

This document describes the follow-up master file of the Mexican Health and Aging Study (MHAS/ENASEM) with information regarding the identifiers, the outcome of the interviews, weights, and related variables for the 2001, 2003, 2012, 2015, and 2018 waves of the study.

This descriptive file enhances but does not preclude the information provided in the previous "Follow-Up Master File For The Mexican Health And Aging Study (MHAS/ENASEM) 2001-2003" file (available here) and "Follow-Up Master File For The Mexican Health And Aging Study (MHAS/ENASEM) 2001-2012" (available here).

This file was designed to provide information at the individual level, starting from the individuals selected for the baseline survey in 2001 and including the new cohorts added in 2012 and 2018. The file includes the new samples added in the third and fifth waves. The file also includes a record for individuals that were selected but never interviewed, as well as those eligible and not eligible for re-interview in 2003, 2012, 2015 or 2018.

II. Study Description[1]

The 2001 baseline survey of the Mexican Health and Aging Study (MHAS/ENASEM) is a national representative survey of individuals born prior to 1951—that is, the population aged 50 or older as of the year 2001. The study protocols and survey instruments are highly comparable to the U.S. Health and Retirement Study (HRS).

The MHAS is designed to examine the aging process and evaluate the impact of disease on health, function, and mortality of adults over the age of 50 living in Mexico. The baseline survey was conducted in the summer of 2001, and a follow-up visit to the same individuals was carried out in the summer of 2003. The MHAS baseline sample was selected from residents of both rural and urban areas, from the National Employment Survey (Encuesta Nacional de Empleo, ENE), carried out by the INEGI (Instituto Nacional de Estadística y Geografía) in Mexico. Households with at least one resident of ages 50 or older were eligible to be part of the MHAS baseline sample. If more than one person was age-eligible in the selected households, then one person was selected at random for the study. If the selected MHAS person was married or in a

consensual union, with the spouse residing in the same household, then the spouse or partner was also recruited to be part of the MHAS regardless of his/her age.

The 2012 survey was conducted in the fall of 2012, and a follow-up visit was carried out in the fall of 2015. In addition, the 2012 sample was refreshed by adding a representative sample of the population from the 1952-1962 birth cohorts, as well as their spouses/partners regardless of age. Similar to the baseline interview, the sampling frame for the new cohort sample was the Mexican National Employment and Occupation Survey (ENOE, previously named National Employment Survey, ENE) 2012.

The 2018 survey was completed in the fall of 2018, and a follow-up visit will be completed in 2021. Similar to the previous surveys, the 2018 sample was refreshed by adding a representative sample of the population from the 1963-1968 birth cohorts, as well as their spouses/partners regardless of age. The sampling frame for the new cohort sample was INEGI's National Directory of Households.

Figure 1. MHAS 2001-2018 Timeline Sample Size and Response Rates Across Waves
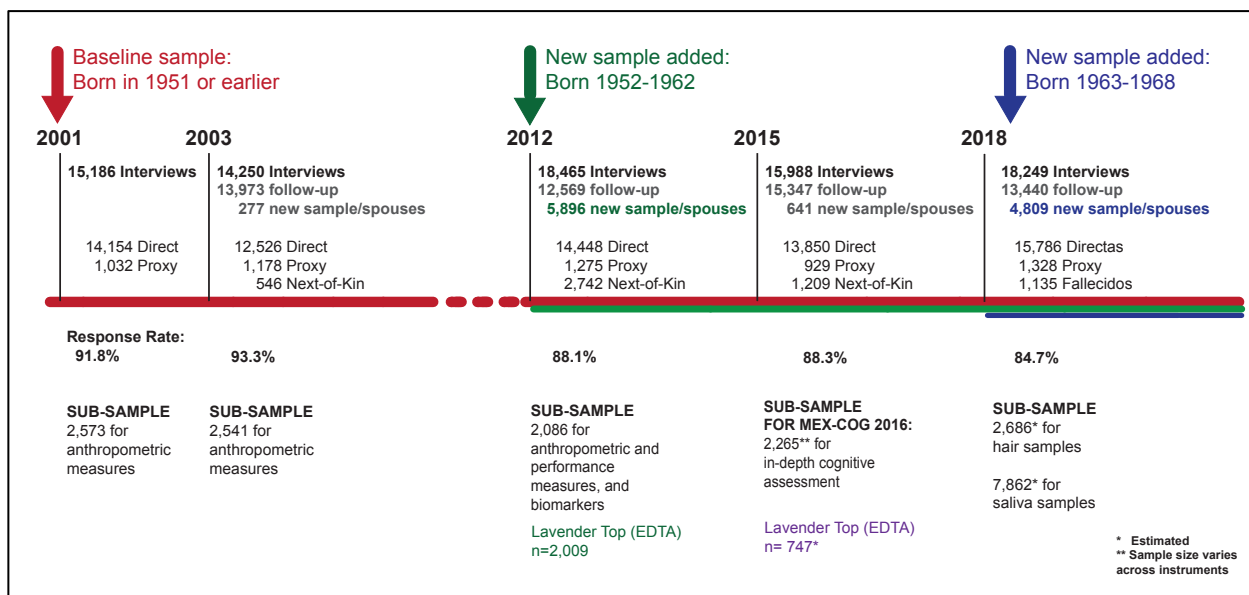
III. Structure of the MHAS Master Follow-up File

The 2018 Master Follow-up File contains a record for every subject included in the MHAS sample (n=28,303), all the identifiers used in each wave and the cover variables. The sample includes

a record of selected individuals eligible for re-interview in 2003, 2012, 2015, or 2018, those not eligible for re-interview in 2003, 2012, or 2015, in addition to samples added in the third and fifth wave (2012 and 2018, respectively). The file contains all the identifiers used in each wave and survey variables including: outcome of the interviews, sampling weights, and other useful variables. Thus the 2018 Master Follow-up File serves as main file to merge any data within and across all the MHAS waves.

There are 106 variables in 2018 Master Follow-up File. The file contents and codebook are included in Appendix 1 and 2, respectively. Below we provide a brief description of the variables included in the data file.

A. Identifiers

The first set of variables in the Master Follow-up File are the household and person identifiers used in all the waves. The following is a short description of each household and individual identifier for each wave, and an explanation on how to use them to merge the data file within and across waves.

1. Unique Household ID (Variable name: `CUNICAH`, also called `UNHHID`)

In 2001, a random sequential number ranging from 1 to 11,000 was created to identify each household at baseline (`CUNICAH` also called `UNHHID`).

The new sample added in 2012 was assigned a value starting from 11,001 to 15,130 to identify each new household, while the new sample added in 2018 was assigned a value starting from 15,131 to 20,218 to identify each new household.

2. Person Identifiers (Variable name: `NP, CODENT01` in 2001, and `CODENT03` in 2003)

Within each household, in the initial interview, there are up to two persons under study (the selected person of eligible age, and spouse regardless of age). Once interviewed, the two individuals become part of the longitudinal study.

2001 Person Identification Code (Variable name: `CODENT01`)

In the baseline interview, the unique person ID (`CODENT01`) is assigned in the household as follows (this distinction is important, as it determines the sample weight for each person):

1 = Selected individual,
2 = Spouse.

2003 Person Identification Code (Variable name: `CODENT03`)

In 2003, these two values were dragged from 2001 with each respondent, and we added two possible values to create `CODENT03`:

3 = New spouse of the person who was = 1 in 2001, and
4 = New spouse of the person who was = 2 in 2001

In combination with the unique household ID in 2001, and the updated household ID in 2003, these codes uniquely identify individual subjects of study.

TIPS: How to use identifiers to merge 2001 and 2003 data files

In 2001, the variable `CUNICAH` serves as unique household identifier. This variable, in combination with the person identifier `CODENT01` (also called `PS3`) serve as a unique person identifier for the first wave.

In 2003, the unique household identifier `CUNICAH` used in 2001 is supplemented with `ACTHOG` to form the unique household identifier. Also, these variables in combination with the person identifier for 2003 given by `CODENT03` (also called `ENT2`) serve as a unique person identifier for the second wave.

New Person identifier (Variable name: `NP`)

Starting in 2012, the variable `NP` was created to identify each person in each original household, as follows:

`NP`=010 (selected person), `NP`=020 (spouse of selected person)

Figure 1. Assignment of NP values within the 010 family
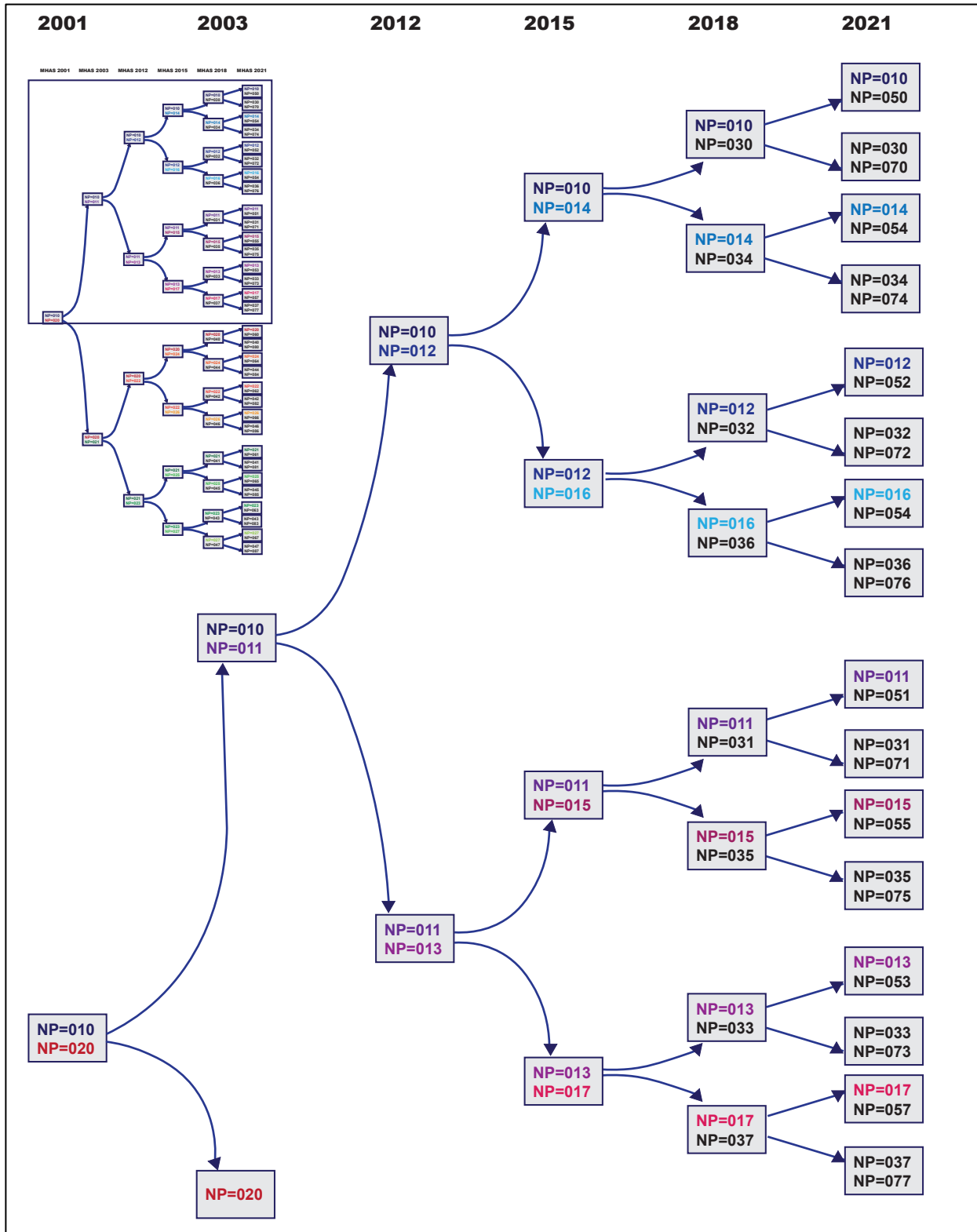
Figure 2. Assignment of NP values within the 020 family
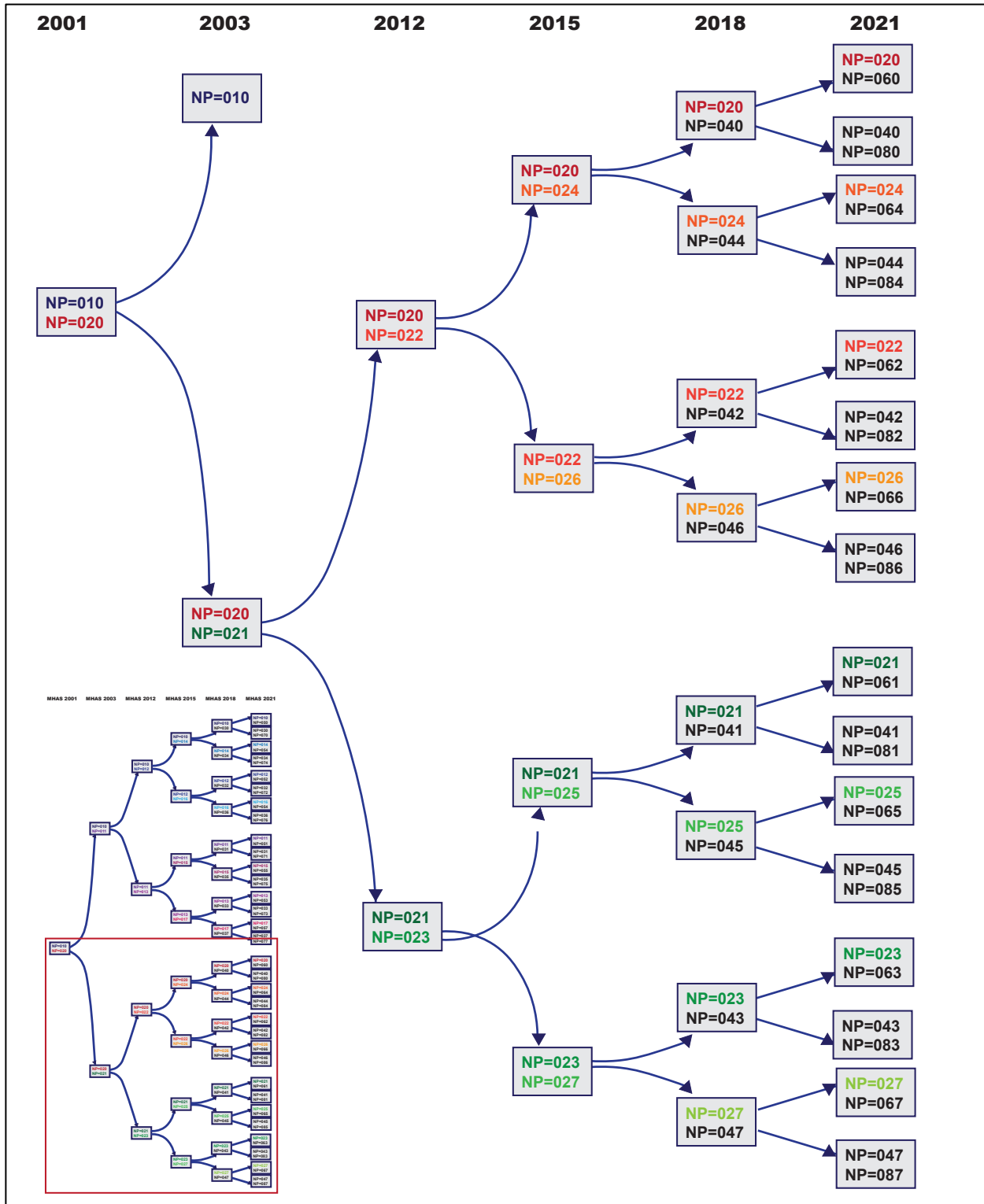
From then on, if these subjects get new spouses, the spouses will be assigned codes that correspond to the person they are married to. If they are a spouse of the baseline subject NP=010, the new spouse is assigned 011 (for the first new spouse if there is one), 012 (for the second new spouse if there is one), etc. If they are spouses of selected person NP=020; the new spouse is assigned 021 (for the first new spouse if there is one), 022 (for the second new spouse if there is one), etc. The NP key assignment would be done in a fixed way. Figure 1 illustrates how NP is assigned in the "family" of the selected subject NP=010, while Figure 2 does it for the "family" of the selected subject NP=020.

In combination with the unique household ID ($CUNICAH$), this code uniquely identifies individual subjects of study.

Unique Person ID (Variable name: $UNHHIDNP$)

Starting in 2018, we added a new string variable that combines the unique household ID ($CUNICAH$) and the person ID (NP).

3. Sub-household ID (Variable name: $ACTHOG$ and $SUBHOG\_XX$)

Updated Household ID (Variable name: $ACTHOG$)

In 2003, a household code was created to capture changes in the situation of the individual or couple interviewed in 2001. This is referred to as "updated household" ($ACTHOG$) and the codes reflect the type of change experienced, including divorce/separation, death, or new spouse. In the case of split couples as mentioned above, an interview was sought with both baseline respondents in their respective households and their new spouses if applicable. The updated-household codes also capture whether the household observed in 2003 contains the baseline *sampled* respondent, or the baseline *spouse* of the selected person.

Sub-household ID (Variable name: $SUBHOG\_XX$)

A sub-household ID ($SUBHOG\_XX$) was created to follow the modifications of the original household and new households that result from the changes occurred over time. These changes include: divorce/separation, death, or new spouse of the original subjects. The last two digits of the variable indicate the year of the respective survey. Thus, we constructed one variable for

each wave (`SUBHOG01`, `SUBHOG03`, `SUBHOG12`, `SUBHOG15`, and `SUBHOG18`), each of them reflecting the changes in the household recorded each wave. This new identifier was created to replace the "updated household" ID (`ACTHOG`) used in 2003.

Table 1. Description of SUBHOG

BASELINE HOUSEHOLD	
00	For all households that are new sample (in their baseline year)
NO CHANGES IN THE HOUSEHOLD	
01	The household contains an individual **NP = 010** (first selected subject that entered the sample)
02	The household contains an individual **NP=020** (subject in this household that entered the sample first)
03	The household contains an individual **NP=011** (subject in this household that entered the sample first)
04	The household contains an individual **NP=021** (subject in this household that entered the sample first)

Note: Starting from 05 up to 10 we could assign new IDs in future waves.

CHANGES IN THE HOUSEHOLD

Note: The change can indicate a dissolution (due to the death of the spouse or separation – with or without a new union) or a new union. In case of a death in the household, this ID is assigned to the survivor.

11	The new household contains an individual **NP=010** (subject in this new household that entered the sample first)
12	The new household contains an individual **NP=020** (subject in this new household that entered the sample first)
13	The new household contains an individual **NP=011** (subject in this new household that entered the sample first)
14	The new household contains an individual **NP=021** (subject in this new household that entered the sample first)
15	The new household contains an individual **NP=012** (subject in this new household that entered the sample first)
16	The new household contains an individual **NP=022** (subject in this new household that entered the sample first)
17	The new household contains an individual **NP=013** (subject in this new household that entered the sample first)
18	The new household contains an individual **NP=023** (subject in this new household that entered the sample first)
19	The new household contains an individual **NP=014** (subject in this new household that entered the sample first)
20	The new household contains an individual **NP=024** (subject in this new household that entered the sample first)

Note: Starting from 19 up to 30 we could assign new IDs in future waves.

DECEASED

Note: This ID is assigned to the deceased; to the survivor we assigned the ID from CHANGES IN THE HOUSEHOLD.

31	The deceased is **NP=010**
32	The deceased is **NP=020**
33	The deceased is **NP=011**
34	The deceased is **NP=021**
35	The deceased is **NP=012**
36	The deceased is **NP=022**
37	The deceased is **NP=013**
38	The deceased is **NP=023**
37	The deceased is **NP=014**
38	The deceased is **NP=024**

Note: Starting from 39 up to 69 we could assign new IDs in future waves.

RE-UNION

70	Two individuals (NP=010 ando NP=020) part of the study, they were together, separated, and got back together

| 71 | Two individuals, from the family of 10, they were together, separated, and got back together |
| 72 | Two individuals, from the family of 20, they were together, separated, and got back together |

NEW INDIVIDUAL	
99	To the new individuals we assign 99 in the previous waves to the one when they enter the study
DECEASED	
88	To the deceased we assign 88 in the posterior waves to the one when he/she was reported dead
NOT CONTACTED	
77	To the individuals not contacted (whole household)

Source: MHAS (2013), The Mexican Health and Aging Study: "MHAS 2012 Data Files Description, Version1, September 2013." Retrieved from www.MHASweb.org on (February 2015).

At baseline, the variable takes the value 00 to reflect no change in the household. Thus SUBHOG_01=00 for all households because they are all original households, new to the study in this wave. In the same way, for the new refresher sample of households added in 2012 and 2018, we assigned the same value (SUBHOG_12=00 and SUBHOG_18=00) to indicate that these are new to the study in this wave. For more information on how the values of SUBHOG_XX are assigned see "MHAS 2012 Data Files Description" available for download here.

B. Outcomes of the Interview, Sample and Study Status

The file includes the following variables to indicate the outcome of the interviews at different levels:

1. Eligible (Variable name: ELEGIBLE_XX)

Starting in 2003, ELEGIBLE_XX indicates whether the individual is eligible for a follow-up interview in any given wave. As general rule, individuals were considered eligible for a re-interview if they completed an interview in the previous wave and they reported age was within the required age to be part of the MHAS sample.

2. Follow-New person (Variable name: NEW_SAMPLE_XX):

Starting in 2003, NEW_SAMPLE_XX indicates whether the interview is for a new person or a follow-up subject. In 2003, the new sample includes either a new spouse from a follow-up subject or a sample subject in 2001 who did not complete the interview in Wave 1.

In 2012, the new sample includes either a new spouse of a follow-up subject, or a sample subject

in 2001 who did not complete the interview in Wave 1. In addition, it includes the new sample of those born between 1952-1962.

Similar to 2003, in 2015 the new sample includes either a new spouse from a follow-up subject or a subject from the new 2012 sample who did not complete the interview in Wave 3.
In 2018, the new sample includes either a new spouse of a follow-up subject or the new sample added of those born between 1963-1968.

3. Type of non-interview (Variable name: `TIPNE_XX`)

Indicates whether a respondent was interviewed as well as the reason of non-interview. This is a key variable for users as it indicates the respondent's status in any given wave.

In 2001 and 2003, it is considered a household variable because if at least one member in the household completed the interview it is considered an "Interviewed" household. In 2001 and 2003 (`TIPNE_01` and `TIPNE_03`) include the following options:

VALUE	LABEL	DESCRIPTION
0	Interviewed	At least one interview was completed in the household.
2	Nobody at time of interview	If nobody was present at the time of the visits (minimum of 5 visits).
3	Temporarily absent	A temporarily absent subject, is a person that at the time of the visit is not in the house, but intends to return during the collection period or before three months, regardless of the reason of the absence.
4	Refused to provide information	No interviews were completed in the household because the subject and the spouse (if any) refused to provide information and didn't allow others (proxy) to provide it either.
5	Other (occupied house)	If the interview was not completed for other reasons different from options 2 to 4.
6	Unoccupied, fit for living	If the house, with or without furniture, is unoccupied, but fit for living, for sale or rent.
7	Temporary housing	No permanent residents were found in the dwelling.
8	Unoccupied, inadequate to live in due to bad conditions or remodeling	If the house is unoccupied and not considered appropriate for living, that is, there is unacceptable structural conditions due either to its damage or other reasons like lack of windows and doors.
9	Unoccupied, temporarily used for other means different than housing	If the house is temporarily occupied by businesses, workshops, factories, etc.
10	Demolished	If the house has been destroyed or is in demolition process.

11	Tent, trailer, mobile home, etc.	If the tent, trailer, mobile home (etc.) has changed location by the time of the interview.
12	Business or permanent warehouse	If the house has been transformed or adapted from their normal living use to a commercial one, for business, or to store harvests, machinery, office supplies and other uses.
13	Other (specify)	If the house can't be classified in any of the previous options.
14	Selected (and spouse) moved and unable to reach or locate	If the selected subject was not located (or lost to follow-up).
15	Selected (and spouse) not fit for interview and no substitute informant	If the selected subject and the spouse (if any) could not complete the interview, for different reasons, and no substitute informant was available to complete the interview.
16	Selected deceased and no spouse	The selected subject was deceased without a spouse.
17	Selected deceased and no spouse within required age	The selected subject was deceased, and the spouse was not within the required age (50 years or older).
18	Selected (and spouse) temporarily absent and there is no adequate or available substitute informant	A household is considered "temporarily absent" if the selected subject and the spouse (if any) were never present at the time of the visits but they were identified living in the household. In this case, not substitute informant was available to complete the interview.
19	Other (specify)	No interviews were completed in the household for other reason not included in the previous options.

In 2012, TIPNE_12 includes the following options:

VALUE	DESCRIPTION
1	Interviewed
2	Incomplete Interview
3	Postponed Interview
4	Absent and there is no informant
5	Empty household or informant not adequate
6	Subject unable to complete interview without informant
7	Deceased without informant
8	Refused
9	Refused direct interview
10	Refused proxy interview
11	Refused next-of-kin interview
12	Changed location
13	Subject could not be located
14	Empty housing
15	Housing with temporary use
16	Housing with non-residential use, demolished or uninhabitable
17	Address not located
18	Area not safe
19	Other
20	Deceased subject from the new sample added in 2012

In 2015, `TIPNE_15` includes the following options:

VALUE	DESCRIPTION
1	Complete Interview
2	Incomplete Interview
3	Postponed Interview
4	Absence of the subject or the proxy or next-of-kin informant
5	Absence of occupants or adequate informant
6	Subject not fit for interview and without proxy informant
7	Subject deceased without next-of-kin informant
8	Refusal
9	Refusal from the subject or the proxy or next-of-kin informant
10	Change of residence
11	Subject could not be located
12	Empty residence
13	Dwelling for temporary use
14	Dwelling with non-residential use, demolished or uninhabitable
15	Address not located
16	Area not safe
17	Other

In 2018, `TIPNE_18` includes the following options:

VALUE	DESCRIPTION
1	Complete Interview
2	Incomplete Interview
3	Postponed Interview
4	Absence of the subject, or proxy or next-of-kin informant
5	Absence of occupants or adequate informant
6	Refusal
7	Subject not fit for interview and w/o proxy informant
8	Subject deceased without next-of-kin informant
9	Age out of the range for new sample
10	Change of residence
11	Subject could not be located
12	Residence could not be located
13	Other situation
14	Inhabited residence
15	Dwelling for temporary use
14	Dwelling with non-residential use, demolished or uninhabitable
16	Area not safe

4. Type of interview (Variable name: `TIPENT_XX`)

Indicates the type of individual interview. The MHAS makes every effort to obtain core interviews directly with respondents themselves. In cases where the respondent is unavailable or unable to participate, a proxy interview is sought with a knowledgeable informant. Starting in 2003, if the respondent was deceased an exit interview is sought with a next-of-kin. `TIPENT_XX`

indicates direct interviews or proxy interviews.

In 2001 and 2003, it also indicates the order of the interview in the household – first or second interview. The options included the following:

VALUE	DESCRIPTION	
	2001	2003
11	Direct, first interview	
12	Direct, second interview	
21	Proxy, first interview	
22	Proxy, second interview	
31		Next-of-kin interview, first interview
32		Next-of-kin interview, second interview
33		Next-of-kin interview, third interview
42		Non-response, second interview

Starting in 2012, TIPENT_XX uses the following options:

VALUE	DESCRIPTION
1	Direct, follow-up interview
2	Direct, new sample interview
3	Proxy, follow-up interview
4	Proxy, new sample interview
5	Next-of-kin

5. Reason for Proxy (Variable name: REASON_PROXY_XX)

In cases where the respondent is unable to complete a direct interview due to (1) health reason, (2) language reason, or (3) temporary absence a proxy interview is sought with a knowledgeable informant. The variable REASON_PROXY_XX indicates the reason a proxy interview was completed.

6. Interview Date (Variable name: INT_DATE_XX)

Indicates the date of the interview.

7. Wave Status – Alive or Deceased (Variable name: FALLECIDO_XX)

For all the eligible respondents in a given wave, the variable FALLECIDO_XX indicates whether a next-of-kin interview for a deceased subject was completed.

C. Other

1. Same dwelling (Variable name: `MISMA_V_XX`)

Indicates whether the household was located in the same dwelling or in a different dwelling or a group quarter (asylum, penitentiary, etc.) as in the previous wave.

2. Interview Technique – Electronic versus Paper (Variable name: `C_PAP_XX`)

Starting in 2012, the interviews were completed using a CAPI (Computer Assisted Personal Interview). However, in some cases due to security or technological issues the interview was completed using paper and pencil. `C_PAP_XX` indicates whether the interview was completed using the electronic questionnaire or paper and pencil.

3. High Migration States (Variable name: `EAM_XX`)

Starting in 2012, it indicates whether the subject is part of a high migration state that was oversampled.

4. Age and Sex (Variable name: `AGE_XX` and `SEX_XX`)

In each wave, the variable `AGE_XX` is calculated for the new sample using the self-reported birth date and the interview date. For the follow-up sample, the variable `AGE_XX` is the self-reported age unless the subject refuses to answer and doesn't know; in those cases the age is calculated similar to how it's done for the new sample.

The variable `SEX_XX` is the self-report given in each wave. In 2015, we observed differences in the reported sex across waves using the variables `SEXO_01`, `SEXO_03`, `SEX_12`, and `SEX_15`. These differences were carefully studied by the MHAS team and INEGI in Mexico and a new variable `SEX_VALIDATED_15` was constructed with the validated sex variable. The user can use this variable instead, from 2001 to 2015.

D. Weights

The file also contains the sampling weights that were constructed for every wave of the study. MHAS has individual (`FACTORI_XX`), household (`FACTORH_XX`), anthropometric and

biomarkers (`FACTORA_XX`) factors. The individual weights are based on the birth cohort, household composition, and the place of residence (in urban and rural areas) and geographic areas. Also, the household weights are based on the sample design and sample selection criteria, and the calibration variables of all community dwelling households with at least one resident 50 years or older, based on the household composition, and the place of residence (urban and rural areas) and geographic areas. The sample selection involves 3 stages: selection of Primary Sampling Units (PSU), then the selection of households within each PSU, and finally the selection of subjects in those households. The sampling weights provided reflect this sample selection as well as post-stratification adjustment for non-response. They are used to expand the sample to the national population. For more information regarding the calculation of the weights please see the 2012 or the 2015 Methodological Document available in our website: http://mhasweb.org/Resources/DOCUMENTS/2012/Methodological_Document_2012%E2%80%93SEC.pdf and

http://mhasweb.org/Resources/DOCUMENTS/2015/Methodological_Document_2015.pdf

E. Sub-Samples and Ancillary Study

1. Selected for Anthropometric Measures 2001 & 2003 (Variable name: `ANTRO_XX`)

In 2001 and 2003, `ANTRO_XX` indicates whether the subject was selected to be part of the sub-sample for anthropometric measures.

2. Result for Anthropometric Measures 2001 & 2003 (Variable name: `RES_ANTRO_XX`)

In 2001 and 2003, `RES_ANTRO_XX` indicates whether the subject selected to be part of the sub-sample for anthropometrics measures participated in the collection of the measurements.

3. Sub-sample 2012 (Variable name: `SUBSAMPLE_12`)

`SUBSAMPLE_12` indicates whether the subject was selected to be part of the 2012 sub-sample for anthropometric, biomarkers, and performance measures.

4. Result for Biomarkers Measures 2012 (Variable name: `RES_BIOMARKERS_12`)

`RES_BIOMARKERS_12` indicates whether the subject was selected to be part of the 2012 sub-sample for anthropometric, biomarkers, and performance measures participated in the collection

of the measurements.

5. Ancillary Study Variables: Cognitive Aging Study (Mex-Cog)

Mex-Cog 2016 Sub-sample (Variable name: `SUBSAMPLE_16`)

In 2016, `SUBSAMPLE_16` indicates whether the subject was selected to be part the Cognitive Aging Study (Mex-Cog) linked to the MHAS. The Mex-Cog 2016 sample was selected using the MHAS wave 4 (2015) as sampling frame. The criteria for eligibility for Mex-Cog were: first, aged 55 and older in MHAS 2015; and second, having completed a direct interview or a proxy interview for health reasons in the MHAS 2015. Only 8 of the 32 states were selected using stratified sampling procedures. The Mex-Cog contains three parts: Cognitive Assessment, Adequate Informant Survey, and Anthropometric and Performance Measures

Mex-Cog 2016 Phase (Variable name: `PHASE_MXCOG_16`)

`PHASE_MXCOG_16` indicates whether the information was collected during Phase 1 in the spring (March-April) or Phase 2 in the fall (October-November) of 2016.

Mex-Cog 2016 Result for Each Component (Variables name: `RES_MXCOG_16`, `RES_COGNITIVO_16`, `RES_INFORMANTE_16`, `RES_ANTRO_16`)

	Frequency	COMPONENTS			
		COG	INF	ANTRO	BIO
0000	983				
0010	2			✓	
0100	200		✓		
0101	2		✓		✓
0110	11		✓	✓	
0111	10		✓	✓	✓
1000	4	✓			
1010	127	✓		✓	
1011	62	✓		✓	✓
1100	9	✓	✓		
1101	1	✓	✓		✓
1110	1,162	✓	✓	✓	
1111	677	✓	✓	✓	✓
Total	**3,250**				

`RES_MXCOG_16` is a string variable that indicates whether or not a component of the study was completed; from left to right each digit indicates each of following components respectively:

cognitive assessment (COG), adequate informant survey (INF), anthropometric and performance measures (ANTRO), and biomarkers (BIO). The table above illustrates the different combinations.

`RES_COGNITIVO_16`, `RES_INFORMANTE_16`, and `RES_ANTRO_16` indicate whether the subject selected to be part of Mex-Cog 2016 completed the cognitive assessment, adequate informant survey, and/or anthropometric and performance measures portion, respectively. The variables also indicated the reason for not completing each component including refusal, loss to follow-up, deceased, or other.

Mex-Cog 2016 Other Variables (Variables name: `RES_BIOMARKERS_16, EDTA_16, HBA1C_16`)

`RES_BIOMARKERS_16` indicates whether the subject completed the biomarkers portion of Mex-Cog 2016. While `HBA1C_16` and `EDTA_16` indicate if Glycosylated Hemoglobin was measured and if a Lavender Tube with EDTA was collected, respectively.

C. Accessing the MHAS Data

All study databases and documentation can be accessed free of charge from the MHAS study website after registration is completed. The platform is in English www.MHASweb.org and in Spanish www.ENASEM.org. The website also features a dynamic searchable database of publications using MHAS data and a discussion forum. **We ask all users to please inform us to info@mhasweb.com of any published work using the MHAS data.**

D. Merging the MHAS Data

Correctly merging data sets is critical to every research project. We always recommend starting every merge with the Master Follow-up File, in particular when merging data files across waves. As described above, the Master Follow-up File contains a record of every subject that has participated in the study and includes all the identifiers used in each wave and the result of the interview in each wave. In this process it is essential you identify the right variables to uniquely identify household and individual. Below we provide a short description on which identifiers to use in each wave.

How to use identifiers to merge data files

To merge data files at the <u>household</u> level you need to use the following variables:

1) `UNHHID` (also named `CUNICAH`) in 2001
2) `CUNICAH`+`ACTHOG` in 2003
3) `CUNICAH`+`SUBHOG_XX` in each wave from 2012 to 2018

Check that the identifiers use the same names both in the 'master' and the 'using dataset'. For example, the unique household ID `CUNICAH` is also named `UNHHID` in some data files.

To merge data files at the <u>individual</u> level you need to use the following variables:

1) `UNHHID` (also named `CUNICAH`)+`CODENT01` (also named `PS3`) in 2001
2) `CUNICAH`+`ACTHOG`+`CODENT03` (also named `ENT2`) in 2003
3) `CUNICAH`+`NP` in each wave from 2012 to 2018

Check that the identifiers use the same names both in the 'master' and the 'using dataset'. For example, the 2001 person ID `CODENT01` is also named `PS3` and the 2003 person ID `CODENT03` is also named `ENT2` in some data files.

Appendix 1. Master Follow-up File Content

Number of Observations	28,303	Version 2. July 2020 (includes weights)
Number of Variables	106	

Variable Name	Type	Variable Label
unhhid	float	Unique Household identification code (=cunicah)
UNHHID	str5	Unique Household ID (=CUNICAH) - STRING
cunicah	float	Clave Unica del Hogar (=unhhid)
CUNICAH	str5	Clave Unica del Hogar (=UNHHID) - STRING
np	int	Person Number/ Numero de Persona
NP	str3	Person Number/Numero de Persona - STRING
UNHHIDNP	str8	Unique Person ID/Numero de Identificador de Persona Unico - STRING
codent01	float	Person identification code 2001 (=ps3)
subhog_01	int	2001 sub-household identifier
acthog	byte	Update household code 2003
codent03	byte	Person identification code 2003 (=ent2)
subhog_03	long	2003 sub-household identifier
subhog_12	byte	2012 sub-household identifier
subhog_15	byte	2015 sub-household identifier
subhog_18	byte	2018 sub-household identifier
res_01	int	Outcome of the household interview 2001
tipne_01	byte	Type of household non-interview 2001
tipent_01	byte	Type of interview 2001
reason_proxy_01	float	Reason for Proxy Interview 2001
tamloc_01	byte	Locality size 2001
factorh_01	long	Household weight 2001
factori_01	long	Individual weight 2001
factora_01	long	Biomarkers weight 2001
edad_01	int	Age 2001
sexo_01	float	Sex 2001 (Male=1)
antro_01	byte	Selected for anthropometric measurement 2001
res_antro_01	byte	Result for Anthropometric Measures 2003
elegible_03	byte	Eligible for interview in 2003
new_sample_03	byte	Follow-up/new person
misma_v_03	byte	Same dwelling in 2003
fallecido_03	byte	Died between 2001 and 2003
res_03	int	Outcome for the household interview 2003
res_ent_03	byte	Outcome for individual interview 2003
tipne_03	byte	Type of non-interview 2003
tipent_03	byte	Type of interview 2003

```
reason_proxy_03      byte      Reason for Proxy Interview 2003
factori_03           long      Individual weight 2003
factorh_03           long      Household weight 2003
factora_03           long      Biomarkers weight 2003
antro_03             byte      Selected for anthropometric measurement 2003
res_antro_03         byte      Result for Anthropometric Measures 2003
edad_03              int       Age 2003
sexo_03              byte      Sex 2003 (Male=1)
mes_03               byte      Month of birth 2003
a_o_03               int       Year of birth 2003
elegible_12          float     Eligible for interview in 2012
new_sample_12        byte      Follow-up and new sample 2012
misma_v_12           byte      Same dwelling in 2012
tipne_12             byte      Type of non-interview 2012
tipent_12            float     Type of interview 2012
reason_proxy_12      byte      Reason for Proxy Interview 2012
int_date_12          str10     Interview date 2012
order_12             byte      Order of interview 2012 (in the household)
fallecido_12         float     Died between 2003 and 2012
tam_loc_12           float     Locality size 2012
eam_12               byte      High migration states 2012
factorh_12           long      Household weight 2012
factori_12           long      Individual weight 2012
factora_12           long      Biomarkers/Anthropometrics weight 2012
age_12               int       Age 2012
sex_12               byte      Sex 2012 (Male=1)
subsample_12         float     Selected subsample for Biomarkers/Anthropometrics 2012
res_biomarke~12      byte      Result for Biomarkers/Anthropometrics 2012
biomarkers_12        float     Sample with Biomarkers 2012
edta_12              float     2012 Completed EDTA-Lavender Tube
hba1c_12             float     2012 Completed HbA1c
elegible_15          float     Eligible for interview in 2015
new_sample_15        float     Follow-up and new sample/spouses 2015
misma_v_15           byte      Same dwelling as in 2012
tipne_15             byte      Type of non-interview 2015
tipent_15            float     Type of interview 2015
reason_proxy_15      byte      Reason for Proxy Interview 2015
int_date_15          str10     Interview date 2015
c_pap_15             byte      Interview technique - CAPI vs Paper
fallecido_15         float     Died between 2012 and 2015
```

```
tam_loc_15        byte     Locality size 2015
eam_15            float    Current Residence in High Migration States from 2012
factorh_15        long     Household weight 2015
factori_15        long     Individual weight 2015
age_15            float    Age 2015
sex_15            byte     Sex (Male=1)
sex_validate~15   float    Validated Sex 2015 (Male=1)
subsample_16      float    Selected subsample for Mex-Cog 2016
phase_mxcog_16    float    Mex-Cog 2016 Phase
res_mxcog_co~16   str4     Mex-Cog Result for Each Component (Character): COG/INF/ANTRO/BIOM
res_mxcog_16      float    Mex-Cog Result All Four Components
res_cognitiv~16   byte     Mex-Cog Result of Cognitive Assessment
res_informan~16   byte     Mex-Cog Result of Informant Interview
res_antro_16      byte     Mex-Cog Result for Anthropometrics 2016
biomarkers_16     float    Mex-Cog 2016 Sample with Biomarkers
edta_16           float    Mex-Cog 2016 Sample with EDTA-Lavender Tube
hba1c_16          float    Mex-Cog 2016 Sample with HbA1c
elegible_18       float    Eligible for interview in 2018
new_sample_18     float    Follow-up and new sample/spouses 2018
tipne_18          byte     Type of non-interview 2018
misma_v_18        byte     Same dwelling as in 2018
resul_hh_18       float    Result of Sections at Household Level 2018
tipent_18         float    Type of interview 2018
reason_proxy_18   byte     Reason for Proxy Interview 2018
int_date_18       str8     Interview date 2018
c_pap_18          byte     Interview technique - Electronic vs Paper
fallecido_18      float    Died between 2015 and 2018
tam_loc_18        byte     Locality size 2018
eam_18            byte     Current Residence in High Migration States from 2018
age_18            float    Age 2018
sex_18            float    Sex (Male=1)
```

Appendix 2. Master Follow-up File Codebook

```
--------------------------------------------------------------------------------------------------
unhhid                                           Unique Household identification code (=cunicah)
--------------------------------------------------------------------------------------------------

                type:  numeric (float)

              range:  [1,20218]                          units:  1
      unique values:  17,664                    missing .:  0/28,303

    Variable |        Obs         Mean    Std. Dev.        Min        Max
-------------+---------------------------------------------------------
      unhhid |     28,303      9794.42    5666.713          1      20218

--------------------------------------------------------------------------------------------------
UNHHID                                                             Unique Household ID
(=CUNICAH) - STRING
--------------------------------------------------------------------------------------------------

                type:  string (str5)

      unique values:  17,664                    missing "":  0/28,303

           examples:  "03926"
                      "07891"
                      "11727"
                      "15182"

--------------------------------------------------------------------------------------------------
cunicah                                                      Clave Unica del Hogar (=unhhid)
--------------------------------------------------------------------------------------------------

                type:  numeric (float)

              range:  [1,20218]                          units:  1
      unique values:  17,664                    missing .:  0/28,303

    Variable |        Obs         Mean    Std. Dev.        Min        Max
-------------+---------------------------------------------------------
     cunicah |     28,303      9794.42    5666.713          1      20218

--------------------------------------------------------------------------------------------------
CUNICAH                                          Clave Unica del Hogar (=UNHHID) - STRING
--------------------------------------------------------------------------------------------------

                type:  string (str5)

      unique values:  17,664                    missing "":  0/28,303

           examples:  "03926"
                      "07891"
                      "11727"
                      "15182"

--------------------------------------------------------------------------------------------------
np                                                          Person Number/ Numero de Persona
--------------------------------------------------------------------------------------------------

                type:  numeric (int)

              range:  [10,44]                            units:  1
      unique values:  13                        missing .:  0/28,303

     Person |
    Number/ |
```

```
  Numero de |
    Persona |      Freq.      Percent        Cum.
------------+-----------------------------------
         10 |     17,621        62.26       62.26
         11 |        236         0.83       63.09
         12 |          4         0.01       63.11
         13 |          5         0.02       63.12
         14 |        110         0.39       63.51
         15 |          2         0.01       63.52
         20 |     10,234        36.16       99.68
         21 |         56         0.20       99.88
         22 |          1         0.00       99.88
         24 |         30         0.11       99.99
         34 |          2         0.01       99.99
         40 |          1         0.00      100.00
         44 |          1         0.00      100.00
------------+-----------------------------------
      Total |     28,303       100.00
```

```
-------------------------------------------------------------------------------------------------
NP                                                       Person Number/Numero de Persona - STRING
-------------------------------------------------------------------------------------------------

               type:  string (str3)

      unique values:  13                          missing "":  0/28,303

            examples:  "010"
                       "010"
                       "010"
                       "020"
```

```
-------------------------------------------------------------------------------------------------
UNHHIDNP                           Unique Person ID/Numero de Identificador de Persona Unico - STRING
-------------------------------------------------------------------------------------------------

               type:  string (str8)

      unique values:  28,303                      missing "":  0/28,303

            examples:  "03926010"
                       "07891020"
                       "11727010"
                       "15182010"
```

```
-------------------------------------------------------------------------------------------------
codent01                                                   Person identification code 2001 (=ps3)
-------------------------------------------------------------------------------------------------

               type:  numeric (float)
              label:  codent01

              range:  [1,2]                          units:  1
      unique values:  2                          missing .:  12,901/28,303

 Person identification code |
            2001 (=ps3) |      Freq.      Percent        Cum.
---------------------------+-----------------------------------
        1.Selected person |      9,812        63.71       63.71
2.Spouse of selected person |      5,590        36.29      100.00
---------------------------+-----------------------------------
                    Total |     15,402       100.00
```

```
-------------------------------------------------------------------------------------------------
subhog_01                                                          2001 sub-household identifier
-------------------------------------------------------------------------------------------------
```

```
                 type:  numeric (int)
                label:  clavesubhog

                range:  [0,99]                      units:  1
        unique values:  3                         missing .:  0/28,303

        2001 sub-household identifier |     Freq.      Percent        Cum.
---------------------------------------+-----------------------------------
                     00.Baseline HH |    15,402        54.42       54.42
  77.HH lost to follow-up/Not contacted |      446         1.58       55.99
        99.Not yet part of the study |    12,455        44.01      100.00
---------------------------------------+-----------------------------------
                               Total |    28,303       100.00


-------------------------------------------------------------------------------------------
acthog                                                         Update household code 2003
-------------------------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  acthog

                range:  [0,32]                      units:  1
        unique values:  9                         missing .:  13,027/28,303

           Update household code 2003 |     Freq.      Percent        Cum.
---------------------------------------+-----------------------------------
00.Neither separation nor death, no new |    14,100        92.30       92.30
01.Neither separation nor death, new sp |      198         1.30       93.60
10.Had a separation and contains sample |       63         0.41       94.01
11.Had a separation and contains sample |        8         0.05       94.06
20.Had a separation and contains spouse |       61         0.40       94.46
21.Had a separation and contains spouse |       12         0.08       94.54
30.One 2001 respondent died, no new spo |      790         5.17       99.71
31.One 2001 respondent died, new spouse |       16         0.10       99.82
         32.Both 2001 respondents died |       28         0.18      100.00
---------------------------------------+-----------------------------------
                               Total |    15,276       100.00


-------------------------------------------------------------------------------------------
codent03                                                 Person identification code 2003 (=ent2)
-------------------------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  codent03

                range:  [1,4]                       units:  1
        unique values:  4                         missing .:  13,028/28,303

Person identification code 2003 (=ent2) |     Freq.      Percent        Cum.
---------------------------------------+-----------------------------------
         1.Selected person in 2001 |     9,675        63.34       63.34
   2.Spouse of selected person in 2001 |     5,498        35.99       99.33
3.New spouse of selected person in 2001 |       95         0.62       99.95
4.New spouse-of-spouse of selected pers |        7         0.05      100.00
---------------------------------------+-----------------------------------
                               Total |    15,275       100.00


-------------------------------------------------------------------------------------------
subhog_03                                                        2003 sub-household identifier
-------------------------------------------------------------------------------------------

                 type:  numeric (long)
                label:  clavesubhog

                range:  [0,99]                      units:  1
        unique values:  11                        missing .:  0/28,303
```

```
          2003 sub-household identifier |     Freq.    Percent       Cum.
----------------------------------------+-----------------------------------
                         00.Baseline HH |       804       2.84       2.84
     01.No change, HH contains NP=010 |    13,272      46.89      49.73
     02.No change, HH contains NP=020 |         1       0.00      49.74
           11.New HH contains NP=010 |       247       0.87      50.61
           12.New HH contains NP=020 |        59       0.21      50.82
           13.New HH contains NP=011 |         6       0.02      50.84
           14.New HH contains NP=021 |         1       0.00      50.84
                    31.NP=010 Deceased |       534       1.89      52.73
                    32.NP=020 Deceased |       278       0.98      53.71
  77.HH lost to follow-up/Not contacted |     1,163       4.11      57.82
       99.Not yet part of the study |    11,938      42.18     100.00
----------------------------------------+-----------------------------------
                                 Total |    28,303     100.00
```

----
subhog_12   2012 sub-household identifier

```
          2012 sub-household identifier |     Freq.    Percent       Cum.
----------------------------------------+-----------------------------------
                         00.Baseline HH |     5,754      20.33      20.33
     01.No change, HH contains NP=010 |     8,222      29.05      49.38
     02.No change, HH contains NP=020 |       117       0.41      49.79
           11.New HH contains NP=010 |       832       2.94      52.73
           12.New HH contains NP=020 |       766       2.71      55.44
           13.New HH contains NP=011 |        20       0.07      55.51
           14.New HH contains NP=021 |         3       0.01      55.52
                    31.NP=010 Deceased |     1,938       6.85      62.37
                    32.NP=020 Deceased |       792       2.80      65.17
                    33.NP=011 Deceased |        12       0.04      65.21
  70.NP=010 & NP=020 separated, reunited |         4       0.01      65.22
71.Subsequent NP=010 HH separated, reun |         2       0.01      65.23
  77.HH lost to follow-up/Not contacted |     3,253      11.49      76.72
       88.Deceased before current wave |       546       1.93      78.65
       99.Not yet part of the study |     6,042      21.35     100.00
----------------------------------------+-----------------------------------
                                 Total |    28,303     100.00
```

----
subhog_15   2015 sub-household identifier

```
              type:  numeric (byte)
             label:  clavesubhog

             range:  [0,99]                    units:  1
      unique values:  15                  missing .:  0/28,303

          2015 sub-household identifier |     Freq.    Percent       Cum.
----------------------------------------+-----------------------------------
                         00.Baseline HH |       644       2.28       2.28
     01.No change, HH contains NP=010 |    12,347      43.62      45.90
     02.No change, HH contains NP=020 |        91       0.32      46.22
           11.New HH contains NP=010 |     1,308       4.62      50.84
           12.New HH contains NP=020 |     1,097       3.88      54.72
           13.New HH contains NP=011 |        20       0.07      54.79
           14.New HH contains NP=021 |         7       0.02      54.81
           15.New HH contains NP=012 |        29       0.10      54.92
                    31.NP=010 Deceased |       794       2.81      57.72
                    32.NP=020 Deceased |       403       1.42      59.15
                    33.NP=011 Deceased |        11       0.04      59.18
                    34.NP=021 Deceased |         2       0.01      59.19
  70.NP=010 & NP=020 separated, reunited |        44       0.16      59.35
  77.HH lost to follow-up/Not contacted |     5,713      20.19      79.53
```

```
          99.Not yet part of the study |    5,793      20.47      100.00
-----------------------------------------+------------------------------------
                                   Total |   28,303     100.00


--------------------------------------------------------------------------------------------
subhog_18                                                        2018 sub-household identifier
--------------------------------------------------------------------------------------------

               type:  numeric (byte)

              range:  [0,77]                        units:  1
      unique values:  18                        missing .:  6,764/28,303

               mean:   6.82599
           std. dev:   16.294

          2018 sub-household identifier |    Freq.     Percent       Cum.
-----------------------------------------+------------------------------------
                        00.Baseline HH |    5,766      26.77       26.77
        01.No change, HH contains NP=010 |   10,913      50.67       77.44
        02.No change, HH contains NP=020 |       74       0.34       77.78
           11.New HH contains NP=010 |    1,397       6.49       84.27
           12.New HH contains NP=020 |    1,268       5.89       90.15
           13.New HH contains NP=011 |       32       0.15       90.30
           14.New HH contains NP=021 |        6       0.03       90.33
           15.New HH contains NP=012 |       19       0.09       90.42
          19. New HH contains NP=014 |       14       0.06       90.48
           20.New HH contains NP=024 |        4       0.02       90.50
                    31.NP=010 Deceased |      784       3.64       94.14
                    32.NP=020 Deceased |      360       1.67       95.81
                    33.NP=011 Deceased |        7       0.03       95.84
                    34.NP=021 Deceased |        2       0.01       95.85
                    39.NP=014 Deceased |        4       0.02       95.87
                    40.NP=024 Deceased |        3       0.01       95.89
    70.NP=010 & NP=020 separated, reunited |      53       0.25       96.13
     77.HH lost to follow-up/Not contacted |      833       3.87      100.00
-----------------------------------------+------------------------------------
                                   Total |   21,539     100.00


--------------------------------------------------------------------------------------------
res_01                                                        Outcome of the household interview 2001
--------------------------------------------------------------------------------------------

               type:  numeric (int)
              label:  res01

              range:  [110,233]                      units:  1
      unique values:  12                        missing .:  12,901/28,303

Outcome of the household interview 2001 |    Freq.     Percent       Cum.
-----------------------------------------+------------------------------------
110.Selected person, information provid |    3,904      25.35       25.35
130.Selected person, information provid |      362       2.35       27.70
202.Couple, Information on selected per |      232       1.51       29.20
203.Couple, Information on selected per |        4       0.03       29.23
210.Couple, Information on selected per |      246       1.60       30.83
211.Couple, Information about both prov |      422       2.74       33.57
212.Couple, Information provided by eac |    9,460      61.42       94.99
213.Couple, Information on selected per |       78       0.51       95.49
222.Couple, Information on both by spou |      470       3.05       98.55
230.Couple, Information on selected per |        6       0.04       98.58
232.Couple, Information on selected per |       86       0.56       99.14
233.Couple, Information on both provide |      132       0.86      100.00
-----------------------------------------+------------------------------------
                                   Total |   15,402     100.00

--------------------------------------------------------------------------------------------
```

```
                     type:  numeric (byte)
                     label:  tipne01

                     range:  [0,0]                         units:  1
             unique values:  1                    missing .:  12,901/28,303
```

Type of household non-interview 2001	Freq.	Percent	Cum.
0.Interviewed	15,402	100.00	100.00
Total	15,402	100.00	

----
tipent_01   Type of interview 2001

```
                     type:  numeric (byte)
                     label:  tipent01

                     range:  [11,22]                       units:  1
             unique values:  4                    missing .:  12,901/28,303
```

Type of interview 2001	Freq.	Percent	Cum.
11.Direct, first interview	9,424	61.19	61.19
12.Direct, second interview	4,946	32.11	93.30
21.Proxy, first interview	438	2.84	96.14
22.Proxy, second interview	594	3.86	100.00
Total	15,402	100.00	

----
reason_proxy_01   Reason for Proxy Interview 2001

```
                     type:  numeric (float)
                     label:  reason

                     range:  [1,3]                         units:  1
             unique values:  3                    missing .:  27,271/28,303
```

Reason for Proxy Interview 2001	Freq.	Percent	Cum.
1.Health reasons	321	31.10	31.10
2.Language reasons	21	2.03	33.14
3.Temporary absence	690	66.86	100.00
Total	1,032	100.00	

----
tamloc_01   Locality size 2001

```
                     type:  numeric (byte)
                     label:  tamloc

                     range:  [1,4]                         units:  1
             unique values:  4                    missing .:  12,776/28,303
```

Locality size 2001	Freq.	Percent	Cum.
1.100,000 - +	9,148	58.92	58.92
2. 15,000 - 99,999	2,325	14.97	73.89

27

```
  3.  2,500 - 14,999 |        1,414          9.11        83.00
         4.< 2,500 |        2,640         17.00       100.00
      ------------------+---------------------------------
               Total |       15,527        100.00


------------------------------------------------------------------------------------------------
factorh_01                                                           Household weight 2001
------------------------------------------------------------------------------------------------

                  type:  numeric (long)

                 range:  [16,35360]                     units:  1
         unique values:  1,840                        missing .:  12,901/28,303

                  mean:  1014.07
              std. dev:  1578.02


      Variable |        Obs          Mean     Std. Dev.       Min          Max
    -------------+-------------------------------------------------------------
      factorh_01 |     15,402      1014.068    1578.016         16        35360

------------------------------------------------------------------------------------------------
factori_01                                                           Individual weight 2001
------------------------------------------------------------------------------------------------

                  type:  numeric (long)

                 range:  [0,44177]                      units:  1
         unique values:  2,125                        missing .:  12,901/28,303

                  mean:  948.301
              std. dev:  1686.75

      Variable |        Obs          Mean     Std. Dev.       Min          Max
    -------------+-------------------------------------------------------------
      factori_01 |     15,402      948.3007    1686.748          0        44177


------------------------------------------------------------------------------------------------
factora_01                                                           Biomarkers weight 2001
------------------------------------------------------------------------------------------------

                  type:  numeric (long)

                 range:  [0,108420]                     units:  1
         unique values:  1,290                        missing .:  12,901/28,303

                  mean:  948.301
              std. dev:  4341.15

      Variable |        Obs          Mean     Std. Dev.       Min          Max
    -------------+-------------------------------------------------------------
      factora_01 |     15,402      948.3013    4341.146          0       108420


------------------------------------------------------------------------------------------------
edad_01                                                                           Age 2001
------------------------------------------------------------------------------------------------

                  type:  numeric (int)

                 range:  [18,105]                       units:  1
         unique values:  84                           missing .:  13,117/28,303

                  mean:  60.2629
              std. dev:  10.8379

      Variable |        Obs          Mean     Std. Dev.       Min          Max
```

```
-------------+--------------------------------------------------------------
     edad_01 |     15,186     60.26294     10.83792          18          105

----------------------------------------------------------------------------
sexo_01                                                     Sex 2001 (Male=1)
----------------------------------------------------------------------------

                 type:  numeric (float)
                label:  sex

                range:  [1,2]                            units:  1
        unique values:  2                        missing .:  12,901/28,303

   Sex 2001 |
   (Male=1) |      Freq.      Percent        Cum.
------------+-----------------------------------
     1.Male |      6,650        43.18       43.18
   2.Female |      8,752        56.82      100.00
------------+-----------------------------------
      Total |     15,402       100.00

----------------------------------------------------------------------------
antro_01                          Selected for anthropometric measurement 2001
----------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  yesno

                range:  [0,1]                            units:  1
        unique values:  2                        missing .:  12,901/28,303

    Selected for |
   anthropometric |
measurement 2001 |      Freq.      Percent        Cum.
-----------------+-----------------------------------
          0.No |     12,458        80.89       80.89
         1.Yes |      2,944        19.11      100.00
-----------------+-----------------------------------
          Total |     15,402       100.00

----------------------------------------------------------------------------
res_antro_01                       Result for Anthropometric Measures 2003
----------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  yesno

                range:  [0,1]                            units:  1
        unique values:  2                        missing .:  12,901/28,303

    Result for |
 Anthropometric |
 Measures 2003 |      Freq.      Percent        Cum.
--------------+-----------------------------------
          0.No |     12,829        83.29       83.29
         1.Yes |      2,573        16.71      100.00
--------------+-----------------------------------
          Total |     15,402       100.00

----------------------------------------------------------------------------
elegible_03                                     Eligible for interview in 2003
----------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  elegible03

                range:  [0,1]                            units:  1
```

```
            unique values:  2                        missing .:  12,901/28,303

   Eligible for interview in 2003 |     Freq.     Percent        Cum.
----------------------------------+---------------------------------
0.Not eligible for 2003 interview |       252        1.64        1.64
    1.Eligible for 2003 interview |    15,150       98.36      100.00
----------------------------------+---------------------------------
                            Total |    15,402      100.00

--------------------------------------------------------------------------------------------------
new_sample_03                                                              Follow-up/new person
--------------------------------------------------------------------------------------------------

               type:  numeric (byte)
              label:  entrevista03

              range:  [1,2]                          units:  1
       unique values:  2                        missing .:  14,463/28,303

    Follow-up/new |
          person |     Freq.     Percent        Cum.
------------------+---------------------------------
1.Follow-up person |    13,620       98.41       98.41
      2.New person |       220        1.59      100.00
------------------+---------------------------------
            Total |    13,840      100.00

--------------------------------------------------------------------------------------------------
misma_v_03                                                               Same dwelling in 2003
--------------------------------------------------------------------------------------------------

               type:  numeric (byte)
              label:  misma_v03

              range:  [1,3]                          units:  1
       unique values:  3                        missing .:  13,028/28,303

             Same dwelling in 2003 |     Freq.     Percent        Cum.
-----------------------------------+---------------------------------
                 1.Same dwelling |    14,475       94.76       94.76
            2.Different dwelling |       798        5.22       99.99
3.Group quarters (asylum, penitentiary) |     2        0.01      100.00
-----------------------------------+---------------------------------
                           Total |    15,275      100.00

--------------------------------------------------------------------------------------------------
fallecido_03                                                          Died between 2001 and 2003
--------------------------------------------------------------------------------------------------

               type:  numeric (byte)
              label:  fallecido03

              range:  [0,1]                          units:  1
       unique values:  2                        missing .:  13,918/28,303

  Died between 2001 and |
                  2003 |     Freq.     Percent        Cum.
-----------------------+---------------------------------
      0.Alive in 2003 |    13,839       96.20       96.20
1.Died between 2001-2003 |     546        3.80      100.00
-----------------------+---------------------------------
                 Total |    14,385      100.00

--------------------------------------------------------------------------------------------------
res_03                                                      Outcome for the household interview 2003
--------------------------------------------------------------------------------------------------
```

```
                    type:  numeric (int)
                    label:  res03

                    range:  [1100,3231]                    units:  1
            unique values:  26                         missing .:  13,028/28,303

       Outcome for the household interview |
                                      2003 |   Freq.    Percent      Cum.
-----------------------------------------+----------------------------------
           1100.One, direct 1st int. |    3,428      22.44      22.44
             1200.One, proxy 1st int. |      331       2.17      24.61
           1300.One, next of kin int. |      257       1.68      26.29
               1400.One, no 1st int. |      279       1.83      28.12
2101.Two, direct 1st interview, direct |      236       1.55      29.66
2102.Two, direct 1st int., proxy int. n |       68       0.45      30.11
2104.Two, direct 1st int., no new perso |       10       0.07      30.17
2110.Two, direct 1st int., direct 2nd i |    7,880      51.59      81.76
2120.Two, direct 1st int., proxy 2nd in |    1,156       7.57      89.33
2130.Two, direct 1st int., next of kin |      456       2.99      92.31
 2140.Two, direct 1st int., no 2nd int. |      234       1.53      93.85
2201.Two, proxy 1st int., no 2nd int.,  |       16       0.10      93.95
2202.Two, proxy 1st int., proxy int. ne |       12       0.08      94.03
2204.Two, proxy 1st int., no new person |        2       0.01      94.04
2220.Two, proxy 1st int., proxy 2nd int |      188       1.23      95.27
2230.Two, proxy 1st int., next of kin i |       48       0.31      95.59
  2240.Two, proxy 1st int., no 2nd int. |        4       0.03      95.61
2301.Two, next of kin int., direct int. |        2       0.01      95.63
2304.Two, next of kin int., no 2nd int. |        2       0.01      95.64
2330.Two, next of kin int., next of kin |       24       0.16      95.80
2340.Two, next of kin int., no 2nd int. |       14       0.09      95.89
2401.Two, no 1st int., direct int. new  |        6       0.04      95.93
2404.Two, no 1st int., no 2nd int., no  |        6       0.04      95.97
    2440.Two, no 1st int., no 2nd int. |      604       3.95      99.92
3131.Three, direct 1st int.,next of kin |        9       0.06      99.98
3231.Three, proxy 1st int., next of kin |        3       0.02     100.00
-----------------------------------------+----------------------------------
                                   Total |   15,275     100.00
```

----
res_ent_03   Outcome for individual interview 2003

```
                    type:  numeric (byte)
                    label:  res_ent03

                    range:  [1,20]                         units:  1
            unique values:  2                          missing .:  14,053/28,303

Outcome for individual |
       interview 2003 |   Freq.    Percent      Cum.
-----------------------+----------------------------------
   1.Complete interview |   14,215      99.75      99.75
20.Incomplete interview |       35       0.25     100.00
-----------------------+----------------------------------
                  Total |   14,250     100.00
```

----
tipne_03   Type of non-interview 2003

```
                    type:  numeric (byte)
                    label:  tipne03

                    range:  [0,29]                         units:  1
            unique values:  9                          missing .:  13,028/28,303
```

```
          Type of non-interview 2003 |     Freq.     Percent      Cum.
----------------------------------------+-----------------------------------
                       0.Interviewed |    14,386       94.18      94.18
21.Changed location and cannot be locat |      317        2.08      96.26
            22.Refused Core Interview |       241        1.58      97.83
           23.Refused Proxy Interview |         5        0.03      97.87
     24.Refused Next-of-Kin Interview |         6        0.04      97.91
          25.Unable without informant |        24        0.16      98.06
   26.Absent and there is no informant |       213        1.39      99.46
    27.Deceased without an informant |        16        0.10      99.56
                          29.Other |        67        0.44     100.00
----------------------------------------+-----------------------------------
                          Total |    15,275      100.00
```

----
tipent_03   Type of interview 2003

```
           type:  numeric (byte)
           label:  tipent03

           range:  [11,42]                       units:  1
     unique values:  8                  missing .:  13,917/28,303

           Type of interview 2003 |     Freq.     Percent      Cum.
----------------------------------------+-----------------------------------
         11.Direct, first interview |     8,465       58.84      58.84
        12.Direct, second interview |     4,061       28.23      87.07
          21.Proxy, first interview |       457        3.18      90.25
         22.Proxy, second interview |       721        5.01      95.26
31.Next-of-kin interview, first intervi |      277        1.93      97.18
32.Next-of-kin interview, second interv |      265        1.84      99.03
33.Next-of-kin interview, third intervi |        4        0.03      99.05
     42.Non-response, second interview |      136        0.95     100.00
----------------------------------------+-----------------------------------
                          Total |    14,386      100.00
```

----
reason_proxy_03   Reason for Proxy Interview 2003

```
           type:  numeric (byte)
           label:  reason

           range:  [1,4]                         units:  1
     unique values:  4                  missing .:  27,125/28,303

  Reason for Proxy |
   Interview 2003 |     Freq.     Percent      Cum.
-------------------+-----------------------------------
   1.Health reasons |      358       30.39      30.39
 2.Language reasons |       54        4.58      34.97
3.Temporary absence |      763       64.77      99.75
    4.Invalid code |        3        0.25     100.00
-------------------+-----------------------------------
          Total |    1,178      100.00
```

----
factori_03   Individual weight 2003

```
           type:  numeric (long)

           range:  [0,46315]                     units:  1
     unique values:  2,114              missing .:  12,776/28,303

           mean:   907.109
```

```
            std. dev:   1746.56

   Variable |        Obs        Mean    Std. Dev.        Min        Max
-------------+--------------------------------------------------------
  factori_03 |     15,527    907.1087    1746.563          0      46315


----------------------------------------------------------------------------------------------
factorh_03                                                          Household weight 2003
----------------------------------------------------------------------------------------------

              type:  numeric (long)

             range:  [0,46315]                  units:  1
     unique values:  2,063                   missing .:  12,776/28,303

              mean:  900.397
          std. dev:   1731.9

   Variable |        Obs        Mean    Std. Dev.        Min        Max
-------------+--------------------------------------------------------
  factorh_03 |     15,527    900.3974    1731.898          0      46315


----------------------------------------------------------------------------------------------
factora_03                                                          Biomarkers weight 2003
----------------------------------------------------------------------------------------------

              type:  numeric (long)

             range:  [0,167630]                 units:  1
     unique values:  1,193                   missing .:  12,776/28,303

              mean:  900.398
          std. dev:   4391.99

   Variable |        Obs        Mean    Std. Dev.        Min        Max
-------------+--------------------------------------------------------
  factora_03 |     15,527    900.3983    4391.992          0     167630


----------------------------------------------------------------------------------------------
antro_03                                            Selected for anthropometric measurement 2003
----------------------------------------------------------------------------------------------

              type:  numeric (byte)
             label:  yesno

             range:  [0,1]                      units:  1
     unique values:  2                       missing .:  13,028/28,303

    Selected for |
   anthropometric |
measurement 2003 |      Freq.     Percent        Cum.
-----------------+---------------------------------
          0.No |     12,370       80.98       80.98
         1.Yes |      2,905       19.02      100.00
-----------------+---------------------------------
          Total |     15,275      100.00


----------------------------------------------------------------------------------------------
res_antro_03                                        Result for Anthropometric Measures 2003
----------------------------------------------------------------------------------------------

              type:  numeric (byte)
             label:  yesno

             range:  [0,1]                      units:  1
     unique values:  2                       missing .:  14,463/28,303
```

```
       Result for |
    Anthropometric |
    Measures 2003 |      Freq.       Percent        Cum.
   ---------------+---------------------------------
            0.No |     11,499         83.09       83.09
           1.Yes |      2,341         16.91      100.00
   ---------------+---------------------------------
           Total |     13,840        100.00
```

----
edad_03   Age 2003

```
              type:  numeric (int)

             range:  [21,999]                    units:  1
     unique values:  85                      missing .:  13,382/28,303

              mean:  62.8848
          std. dev:  27.5392
```

```
     Variable |        Obs         Mean    Std. Dev.       Min         Max
   -----------+-------------------------------------------------------------
      edad_03 |     14,910     62.19416     10.58009        21         107
```

----
sexo_03   Sex 2003 (Male=1)

```
              type:  numeric (byte)
             label:  sex

             range:  [1,2]                       units:  1
     unique values:  2                       missing .:  14,599/28,303
```

```
     Sex 2003 |
     (Male=1) |      Freq.       Percent        Cum.
   -----------+---------------------------------
       1.Male |      5,827         42.52       42.52
     2.Female |      7,877         57.48      100.00
   -----------+---------------------------------
        Total |     13,704        100.00
```

----
mes_03   Month of birth 2003

```
              type:  numeric (byte)

             range:  [1,99]                      units:  1
     unique values:  13                      missing .:  28,096/28,303

              mean:  16.8309
          std. dev:  29.3187
```

```
     Variable |        Obs         Mean    Std. Dev.       Min         Max
   -----------+-------------------------------------------------------------
       mes_03 |        184     6.559783     3.601261         1          12
```

----
a_o_03   Year of birth 2003

```
              type:  numeric (int)

             range:  [1910,9999]                 units:  1
```

```
        unique values:  51                          missing .:  28,096/28,303

                 mean:   2373.21
             std. dev:   1810.98

      Variable |        Obs        Mean    Std. Dev.       Min        Max
  -------------+---------------------------------------------------------
        a_o_03 |        196    1945.235    11.29357       1910       1974


----------------------------------------------------------------------------------------------
elegible_12                                                  Eligible for interview in 2012
----------------------------------------------------------------------------------------------

                 type:  numeric (float)
                label:  elegible

                range:  [0,2]                        units:  1
        unique values:  3                          missing .:  13,028/28,303

   Eligible for interview in |
                       2012 |      Freq.     Percent        Cum.
  --------------------------+---------------------------------
  0.Not eligible for interview |      415        2.72        2.72
    1.Eligible for interview |   14,291       93.56       96.27
                2.Fallecido |      569        3.73      100.00
  --------------------------+---------------------------------
                      Total |   15,275      100.00


----------------------------------------------------------------------------------------------
new_sample_12                                               Follow-up and new sample 2012
----------------------------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  sample

                range:  [1,2]                        units:  1
        unique values:  2                          missing .:  7,305/28,303

    Follow-up |
      and new |
  sample 2012 |      Freq.     Percent        Cum.
  ------------+---------------------------------
   1.Follow-up |   14,286       68.04       68.04
  2.New person |    6,712       31.96      100.00
  ------------+---------------------------------
        Total |   20,998      100.00


----------------------------------------------------------------------------------------------
misma_v_12                                                       Same dwelling in 2012
----------------------------------------------------------------------------------------------

                 type:  numeric (byte)
                label:  dwelling

                range:  [1,2]                        units:  1
        unique values:  2                          missing .:  9,013/28,303

    Same dwelling in |
                2012 |      Freq.     Percent        Cum.
  -------------------+---------------------------------
     1.Same dwelling |   18,788       97.40       97.40
  2.Different dwelling |      502        2.60      100.00
  -------------------+---------------------------------
               Total |   19,290      100.00


----------------------------------------------------------------------------------------------
```

```
              type:  numeric (byte)
              label:  nointerview

              range:  [1,20]                          units:  1
      unique values:  20                        missing .:  7,305/28,303
```

Type of non-interview 2012	Freq.	Percent	Cum.
1.Interviewed	18,465	87.94	87.94
2.Incomplete Interview	158	0.75	88.69
3.Postponed Interview	52	0.25	88.94
4.Absent and there is no informant	183	0.87	89.81
5.Empty household or informant not adeq	253	1.20	91.01
6.Subject unable to complete interview	12	0.06	91.07
7.Deceased without informant	120	0.57	91.64
8.Refused	403	1.92	93.56
9.Refused direct interview	236	1.12	94.69
10.Refused proxy interview	5	0.02	94.71
11.Refused next-of-kin interview	11	0.05	94.76
12.Changed location	113	0.54	95.30
13.Subject could not be located	690	3.29	98.59
14.Empty housing	95	0.45	99.04
15.Housing with temporary use	25	0.12	99.16
16.Housing with non-residential use, de	16	0.08	99.23
17.Address not located	89	0.42	99.66
18.Area not safe	11	0.05	99.71
19.Other	51	0.24	99.95
20.Deceased subject from the new sample	10	0.05	100.00
Total	20,998	100.00	


```
              type:  numeric (float)
              label:  interview

              range:  [1,5]                           units:  1
      unique values:  5                         missing .:  9,838/28,303
```

Type of interview 2012	Freq.	Percent	Cum.
1.Direct, follow-up interview	8,868	48.03	48.03
2.Direct, new sample interview	5,580	30.22	78.25
3.Proxy, follow-up interview	959	5.19	83.44
4.Proxy, new sample interview	316	1.71	85.15
5.Next-of-kin	2,742	14.85	100.00
Total	18,465	100.00	


```
              type:  numeric (byte)
              label:  reason

              range:  [1,4]                           units:  1
      unique values:  4                         missing .:  27,028/28,303
```

  Reason for Proxy |

```
      Interview 2012 |      Freq.     Percent        Cum.
-------------------+----------------------------------
  1.Health reasons |        657       51.53       51.53
 2.Language reasons |       111        8.71       60.24
3.Temporary absence |       506       39.69       99.92
     4.Invalid code |         1        0.08      100.00
-------------------+----------------------------------
            Total |      1,275      100.00
```

----
int_date_12   Interview date 2012

```
              type:  string (str10)

      unique values:  86                        missing "":  9,633/28,303

           examples:  ""
                     "06/10/2012"
                     "14/11/2012"
                     "23/11/2012"
```

----
order_12   Order of interview 2012 (in the household)

```
               type:  numeric (byte)
              label:  o_interview

              range:  [1,3]                      units:  1
      unique values:  3                        missing .:  8,222/28,303

Order of interview |
     2012 (in the |
        household) |      Freq.     Percent        Cum.
-------------------+----------------------------------
 1.First interview |     14,571       72.56       72.56
2.Second interview |      5,508       27.43       99.99
 3.Third interview |          2        0.01      100.00
-------------------+----------------------------------
            Total |     20,081      100.00
```

----
fallecido_12   Died between 2003 and 2012

```
               type:  numeric (float)

              range:  [0,1]                      units:  1
      unique values:  2                        missing .:  9,838/28,303

       Died |
    between |
   2003 and |
       2012 |      Freq.     Percent        Cum.
------------+----------------------------------
         0 |     15,723       85.15       85.15
         1 |      2,742       14.85      100.00
------------+----------------------------------
      Total |     18,465      100.00
```

----
tam_loc_12   Locality size 2012

```
               type:  numeric (float)
               label:  tam_loc

               range:  [1,4]                        units:  1
       unique values:  4                          missing .:  7,325/28,303

          Locality size 2012 |      Freq.      Percent       Cum.
----------------------------+----------------------------------
      1.Population = 100,000+ |     12,673       60.41       60.41
2.Population = 15,000 - 99,999 |      2,282       10.88       71.29
 3.Population = 2,500 - 14,999 |      2,250       10.73       82.01
        4.Population <2,500 |      3,773       17.99      100.00
----------------------------+----------------------------------
                     Total |     20,978      100.00


------------------------------------------------------------------------------------------------
eam_12                                                          High migration states 2012
------------------------------------------------------------------------------------------------

               type:  numeric (byte)

               range:  [0,1]                        units:  1
       unique values:  2                          missing .:  9,838/28,303

        High |
   migration |
 states 2012 |      Freq.      Percent       Cum.
------------+----------------------------------
          0 |     12,486       67.62       67.62
          1 |      5,979       32.38      100.00
------------+----------------------------------
       Total |     18,465      100.00


------------------------------------------------------------------------------------------------
factorh_12                                                         Household weight 2012
------------------------------------------------------------------------------------------------

               type:  numeric (long)

               range:  [0,72628]                    units:  1
       unique values:  3,049                      missing .:  7,305/28,303

               mean:   1390.29
           std. dev:   2837.92



     Variable |        Obs        Mean     Std. Dev.       Min         Max
-------------+----------------------------------------------------------
   factorh_12 |     20,998    1390.289      2837.92          0       72628

------------------------------------------------------------------------------------------------
factori_12                                                         Individual weight 2012
------------------------------------------------------------------------------------------------

               type:  numeric (long)

               range:  [0,97546]                    units:  1
       unique values:  3,209                      missing .:  7,305/28,303

               mean:   1365.85
           std. dev:   3022.62

     Variable |        Obs        Mean     Std. Dev.       Min         Max
-------------+----------------------------------------------------------
   factori_12 |     20,998    1365.852     3022.617          0       97546
```

```
--------------------------------------------------------------------------------------------------
factora_12                                                    Biomarkers/Anthropometrics weight 2012
--------------------------------------------------------------------------------------------------

              type:  numeric (long)

             range:  [0,133613]                      units:  1
     unique values:  848                           missing .:  7,305/28,303

              mean:   360.481
          std. dev:   2127.92

    Variable |        Obs         Mean     Std. Dev.       Min         Max
-------------+-----------------------------------------------------------
  factora_12 |     20,998     360.4812    2127.915          0      133613

--------------------------------------------------------------------------------------------------
age_12                                                                               Age 2012
--------------------------------------------------------------------------------------------------

              type:  numeric (int)

             range:  [21,999]                        units:  1
     unique values:  90                            missing .:  9,445/28,303

              mean:   84.8225
          std. dev:     132.1

    Variable |        Obs         Mean     Std. Dev.       Min         Max
-------------+-----------------------------------------------------------
      age_12 |     18,474     65.85049    11.51231         21         115

--------------------------------------------------------------------------------------------------
sex_12                                                                        Sex 2012 (Male=1)
--------------------------------------------------------------------------------------------------

              type:  numeric (byte)

             range:  [1,2]                           units:  1
     unique values:  2                             missing .:  9,508/28,303

   Sex 2012 |
   (Male=1) |      Freq.      Percent        Cum.
------------+-----------------------------------
          1 |      8,170        43.47       43.47
          2 |     10,625        56.53      100.00
------------+-----------------------------------
      Total |     18,795       100.00


--------------------------------------------------------------------------------------------------
subsample_12                           Selected subsample for Biomarkers/Anthropometrics 2012
--------------------------------------------------------------------------------------------------

              type:  numeric (float)

             range:  [0,1]                           units:  1
     unique values:  2                             missing .:  9,835/28,303


  Selected subsample |
     for Biomarkers/ |
Anthropometrics 2012 |      Freq.      Percent        Cum.
---------------------+-----------------------------------
                   0 |     16,160        87.50       87.50
                   1 |      2,308        12.50      100.00
```

```
--------------------+----------------------------------
           Total |     18,468      100.00
```

---
res_biomarkers_12   Result for Biomarkers/Anthropometrics 2012

```
                type:  numeric (byte)
                label:  biomarkers_res

                range:  [1,13]                        units:  1
        unique values:  8                    missing .:  25,999/28,303

  Result for Biomarkers/Anthropometrics |
                                  2012 |     Freq.      Percent        Cum.
---------------------------------------+---------------------------------
                 1.Complete Interview |     2,086       90.54        90.54
   3.Subject not fit for the interview |         1        0.04        90.58
                 5.Absence of subject |        68        2.95        93.53
                            6.Refusal |        92        3.99        97.53
                    7.Subject Deceased |         1        0.04        97.57
               8.Change of residence |        11        0.48        98.05
                             12.Other |        28        1.22        99.26
13.Household refused to provide informa |        17        0.74       100.00
---------------------------------------+---------------------------------
                               Total |     2,304      100.00
```

---
biomarkers_12   Sample with Biomarkers 2012

```
                type:  numeric (float)

                range:  [0,1]                          units:  1
        unique values:  2                    missing .:  25,995/28,303

Sample with |
 Biomarkers |
     2012 |     Freq.      Percent        Cum.
------------+---------------------------------
         0 |       222        9.62         9.62
         1 |     2,086       90.38       100.00
------------+---------------------------------
     Total |     2,308      100.00
```

---
edta_12   2012 Completed EDTA-Lavender
Tube

```
                type:  numeric (float)

                range:  [0,1]                          units:  1
        unique values:  2                    missing .:  25,982/28,303

         2012 |
Completed EDTA- |
  Lavender Tube |     Freq.      Percent        Cum.
----------------+---------------------------------
             0 |       310       13.36        13.36
             1 |     2,011       86.64       100.00
----------------+---------------------------------
         Total |     2,321      100.00
```

---
hba1c_12   2012 Completed HbA1c

```
```

```
                    type:  numeric (float)

                   range:  [0,1]                          units:  1
           unique values:  2                           missing .:  25,995/28,303


        2012 |
   Completed |
       HbA1c |      Freq.      Percent        Cum.
------------+-----------------------------------
          0 |        272        11.79        11.79
          1 |      2,036        88.21       100.00
------------+-----------------------------------
      Total |      2,308       100.00


--------------------------------------------------------------------------------------------
elegible_15                                                    Eligible for interview in 2015
--------------------------------------------------------------------------------------------

                    type:  numeric (float)
                   label:  elegible

                   range:  [0,2]                          units:  1
           unique values:  3                           missing .:  7,305/28,303

   Eligible for interview in |
                        2015 |      Freq.      Percent        Cum.
----------------------------+-----------------------------------
0.Not eligible for interview |        417         1.99         1.99
   1.Eligible for interview |     17,839        84.96        86.94
               2.Fallecido |      2,742        13.06       100.00
----------------------------+-----------------------------------
                     Total |     20,998       100.00


--------------------------------------------------------------------------------------------
new_sample_15                                             Follow-up and new sample/spouses 2015
--------------------------------------------------------------------------------------------

                    type:  numeric (float)
                   label:  sample15

                   range:  [1,3]                          units:  1
           unique values:  3                           missing .:  10,317/28,303

  Follow-up and new sample/spouses |
                              2015 |      Freq.      Percent        Cum.
----------------------------------+-----------------------------------
            1.Follow-up sample |     16,983        94.42        94.42
2.New spouse from follow-up sample |        306         1.70        96.12
 3.New person from 2012 new sample |        697         3.88       100.00
----------------------------------+-----------------------------------
                          Total |     17,986       100.00


--------------------------------------------------------------------------------------------
misma_v_15                                                       Same dwelling as in 2012
--------------------------------------------------------------------------------------------

                    type:  numeric (byte)
                   label:  yesno

                   range:  [1,2]                          units:  1
           unique values:  2                           missing .:  12,315/28,303
```

```
      Same |
dwelling as |
   in 2012 |      Freq.       Percent        Cum.
------------+-----------------------------------
     1.Yes |     15,788        98.75       98.75
      2.No |        200         1.25      100.00
------------+-----------------------------------
     Total |     15,988       100.00
```

----
tipne_15   Type of non-interview 2015

```
              type:  numeric (byte)
              label:  nointerview15

              range:  [1,17]                       units:  1
      unique values:  17                     missing .:  10,317/28,303

         Type of non-interview 2015 |     Freq.      Percent        Cum.
------------------------------------+-----------------------------------
                1.Complete Interview |    15,884        88.31       88.31
              2.Incomplete Interview |       125         0.69       89.01
               3.Postponed Interview |        50         0.28       89.29
4.Absence of the subject or the proxy o |    119         0.66       89.95
5.Absence of occupants or adequate info |    157         0.87       90.82
6.Subject not fit for interview and wit |      4         0.02       90.84
7.Subject deceased without next-of-kin |        14         0.08       90.92
                           8.Refusal |       274         1.52       92.44
9.Refusal from the subject or the proxy |    278         1.55       93.99
               10.Change of residence |        35         0.19       94.18
        11.Subject could not be located |      786         4.37       98.55
                 12.Empty residence |        68         0.38       98.93
          13.Dwelling for temporary use |       14         0.08       99.01
14.Dwelling with non-residential use, d |      6         0.03       99.04
               15.Address not located |        69         0.38       99.43
                   16.Area not safe |        10         0.06       99.48
                            17.Other |        93         0.52      100.00
------------------------------------+-----------------------------------
                              Total |    17,986       100.00
```

----
tipent_15   Type of interview 2015

```
              type:  numeric (float)
              label:  interview2

              range:  [1,5]                        units:  1
      unique values:  5                      missing .:  12,315/28,303

        Type of interview 2015 |     Freq.      Percent        Cum.
------------------------------+-----------------------------------
 1.Direct, follow-up interview |    13,254        82.90       82.90
 2.Direct, new spouse interview |       596         3.73       86.63
  3.Proxy, follow-up interview |       884         5.53       92.16
  4.Proxy, new spouse interview |        45         0.28       92.44
              5.Next-of-kin |     1,209         7.56      100.00
------------------------------+-----------------------------------
                       Total |    15,988       100.00
```
----
reason_proxy_15   Reason for Proxy Interview 2015

```
              type:  numeric (byte)
              label:  reason
```

```
                 range:  [1,3]                          units:  1
         unique values:  3                          missing .:  27,374/28,303

   Reason for Proxy |
     Interview 2015 |      Freq.      Percent        Cum.
-------------------+-----------------------------------
   1.Health reasons |        567        61.03        61.03
  2.Language reasons |         55         5.92        66.95
3.Temporary absence |        307        33.05       100.00
-------------------+-----------------------------------
              Total |        929       100.00
```

----
int_date_15   Interview date 2015

```
                  type:  string (str10)

         unique values:  74                         missing "":  10,629/28,303

              examples:  ""
                         "03/11/2015"
                         "13/10/2015"
                         "22/10/2015"
```

----
c_pap_15   Interview technique - CAPI vs Paper

```
                  type:  numeric (byte)
                 label:  capi

                 range:  [0,1]                          units:  1
         unique values:  2                          missing .:  13,524/28,303


   Interview |
 technique - |
    CAPI vs |
      Paper |      Freq.      Percent        Cum.
------------+-----------------------------------
     0.CAPI |     14,576        98.63        98.63
    1.Paper |        203         1.37       100.00
------------+-----------------------------------
      Total |     14,779       100.00
```

----
fallecido_15   Died between 2012 and 2015

```
                  type:  numeric (float)

                 range:  [0,1]                          units:  1
         unique values:  2                          missing .:  10,317/28,303

Died between |
2012 and 2015 |      Freq.      Percent        Cum.
------------+-----------------------------------
          0 |     16,777        93.28        93.28
          1 |      1,209         6.72       100.00
------------+-----------------------------------
      Total |     17,986       100.00
```

----
tam_loc_15   Locality size 2015

```
            type:  numeric (byte)
            label:  tam_loc

            range:  [1,4]                    units:  1
    unique values:  4                   missing .:  10,317/28,303

        Locality size 2015 |     Freq.     Percent        Cum.
-------------------------------+-----------------------------------
     1.Population = 100,000+ |    10,658       59.26       59.26
2.Population = 15,000 - 99,999 |     2,343       13.03       72.28
 3.Population = 2,500 - 14,999 |     1,624        9.03       81.31
        4.Population <2,500 |     3,361       18.69      100.00
-------------------------------+-----------------------------------
                     Total |    17,986      100.00


--------------------------------------------------------------------------------------------------
eam_15                                        Current Residence in High Migration States from 2012
--------------------------------------------------------------------------------------------------

            type:  numeric (float)

            range:  [0,1]                     units:  1
    unique values:  2                    missing .:  10,317/28,303


    Current |
   Residence |
    in High |
   Migration |
 States from |
       2012 |     Freq.     Percent        Cum.
------------+-----------------------------------
         0 |    11,126       61.86       61.86
         1 |     6,860       38.14      100.00
------------+-----------------------------------
     Total |    17,986      100.00


--------------------------------------------------------------------------------------------------
factorh_15                                                                   Household weight 2015
--------------------------------------------------------------------------------------------------

            type:  numeric (long)

            range:  [0,51137]                units:  1
    unique values:  2,893               missing .:  10,317/28,303

             mean:  1488.52
         std. dev:  2760.26

   Variable |       Obs        Mean     Std. Dev.      Min         Max
------------+--------------------------------------------------------
 factorh_15 |    17,986    1488.517    2760.262        0       51137

--------------------------------------------------------------------------------------------------
factori_15                                                                  Individual weight 2015
--------------------------------------------------------------------------------------------------

            type:  numeric (long)

            range:  [0,90984]                units:  1
    unique values:  3,186               missing .:  10,317/28,303

             mean:  1254.83
         std. dev:  2641.32

   Variable |        Obs        Mean     Std. Dev.      Min          Max
```

```
-------------+-------------------------------------------------------------
 factori_15 |     17,986     1254.832    2641.321            0      90984


-------------------------------------------------------------------------------------------------
age_15                                                                                   Age 2015
-------------------------------------------------------------------------------------------------

                type:  numeric (float)

               range:  [22,999]                        units:  1
       unique values:  85                            missing .:  13,510/28,303

                mean:  66.3865
            std. dev:  18.6787

      Variable |        Obs        Mean    Std. Dev.       Min        Max
 -------------+-------------------------------------------------------------
        age_15 |     14,789    66.13422    10.66113        22        113


-------------------------------------------------------------------------------------------------
sex_15                                                                                Sex (Male=1)
-------------------------------------------------------------------------------------------------

                type:  numeric (byte)
               label:  sex

               range:  [1,2]                          units:  1
       unique values:  2                            missing .:  10,342/28,303

         Sex |
     (Male=1) |      Freq.     Percent        Cum.
 ------------+-----------------------------------
     1.Male |      7,666       42.68       42.68
   2.Female |     10,295       57.32      100.00
 ------------+-----------------------------------
       Total |     17,961      100.00


-------------------------------------------------------------------------------------------------
sex_validated_15                                                       Validated Sex 2015 (Male=1)
-------------------------------------------------------------------------------------------------

                type:  numeric (float)
               label:  sex

               range:  [1,2]                          units:  1
       unique values:  2                            missing .:  5,909/28,303

   Validated |
   Sex 2015 |
     (Male=1) |      Freq.     Percent        Cum.
 ------------+-----------------------------------
     1.Male |      9,860       44.03       44.03
   2.Female |     12,534       55.97      100.00
 ------------+-----------------------------------
       Total |     22,394      100.00


-------------------------------------------------------------------------------------------------
subsample_16                                              Selected subsample for Mex-Cog 2016
-------------------------------------------------------------------------------------------------

                type:  numeric (float)

               range:  [0,1]                          units:  1
       unique values:  2                            missing .:  13,524/28,303
```

```
   Selected |
  subsample |
 for Mex-Cog |
       2016 |      Freq.        Percent        Cum.
------------+-----------------------------------
          0 |     11,529          78.01         78.01
          1 |      3,250          21.99        100.00
------------+-----------------------------------
      Total |     14,779         100.00
```

---
phase_mxcog_16   Mex-Cog 2016 Phase

```
               type:  numeric (float)
              label:  phase

              range:  [1,2]                     units:  1
      unique values:  2             missing .:  25,053/28,303

     Mex-Cog |
  2016 Phase |      Freq.        Percent        Cum.
------------+-----------------------------------
  1.Phase 1 |      1,821          56.03         56.03
  2.Phase 2 |      1,429          43.97        100.00
------------+-----------------------------------
      Total |      3,250         100.00
```

---
res_mxcog_comp_16   Mex-Cog Result for Each Component (Character): COG/INF/ANTRO/BIOM

```
               type:  string (str4)

      unique values:  13            missing "":  25,053/28,303

Mex-Cog Result for |
    Each Component |
      (Character): |
COG/INF/ANTRO/BIOM |      Freq.        Percent        Cum.
-------------------+-----------------------------------
              0000 |        983          30.25         30.25
              0010 |          2           0.06         30.31
              0100 |        200           6.15         36.46
              0101 |          2           0.06         36.52
              0110 |         11           0.34         36.86
              0111 |         10           0.31         37.17
              1000 |          4           0.12         37.29
              1010 |        127           3.91         41.20
              1011 |         62           1.91         43.11
              1100 |          9           0.28         43.38
              1101 |          1           0.03         43.42
              1110 |      1,162          35.75         79.17
              1111 |        677          20.83        100.00
-------------------+-----------------------------------
             Total |      3,250         100.00
```

---
res_mxcog_16   Mex-Cog Result All Four Components

```
               type:  numeric (float)
              label:  result

              range:  [1,6]                     units:  1
```

```
      Mex-Cog Result All Four Components |     Freq.     Percent       Cum.
----------------------------------------+-----------------------------------
                            1.Complete |    1,839       56.58       56.58
                          2.Incomplete |      428       13.17       69.75
                             3.Refusal |      267        8.22       77.97
                     4.Lost to Follow-up |      647       19.91       97.88
                           5.Deceased |       67        2.06       99.94
6.Subject not fit for interview & witho |        2        0.06      100.00
----------------------------------------+-----------------------------------
                                 Total |    3,250      100.00
```

```
-----------------------------------------------------------------------------------------
res_cognitivo_16                                    Mex-Cog Result of Cognitive Assessment
-----------------------------------------------------------------------------------------
```

```
            type:  numeric (byte)
           label:  result

           range:  [1,6]                        units:  1
    unique values:  5                       missing .:  25,053/28,303
```

```
Mex-Cog Result of Cognitive Assessment |     Freq.     Percent       Cum.
----------------------------------------+-----------------------------------
                            1.Complete |    2,042       62.83       62.83
                             3.Refusal |      306        9.42       72.25
                     4.Lost to Follow-up |      807       24.83       97.08
                           5.Deceased |       67        2.06       99.14
6.Subject not fit for interview & witho |       28        0.86      100.00
----------------------------------------+-----------------------------------
                                 Total |    3,250      100.00
```

```
-----------------------------------------------------------------------------------------
res_informante_16                                    Mex-Cog Result of Informant Interview
-----------------------------------------------------------------------------------------
```

```
            type:  numeric (byte)
           label:  result

           range:  [1,6]                        units:  1
    unique values:  5                       missing .:  25,053/28,303
```

```
  Mex-Cog Result of Informant Interview |     Freq.     Percent       Cum.
----------------------------------------+-----------------------------------
                            1.Complete |    2,072       63.75       63.75
                             3.Refusal |      295        9.08       72.83
                     4.Lost to Follow-up |      806       24.80       97.63
                           5.Deceased |       67        2.06       99.69
6.Subject not fit for interview & witho |       10        0.31      100.00
----------------------------------------+-----------------------------------
                                 Total |    3,250      100.00
```

```
-----------------------------------------------------------------------------------------
res_antro_16                                      Mex-Cog Result for Anthropometrics 2016
-----------------------------------------------------------------------------------------
```

```
            type:  numeric (byte)
           label:  result

           range:  [1,6]                        units:  1
    unique values:  5                       missing .:  25,053/28,303
```

```
Mex-Cog Result for Anthropometrics 2016 |     Freq.     Percent       Cum.
```

```
-----------------------------------------+----------------------------------
                        1.Complete |      2,051        63.11        63.11
                         3.Refusal |        303         9.32        72.43
                4.Lost to Follow-up |        804        24.74        97.17
                        5.Deceased |         67         2.06        99.23
6.Subject not fit for interview & witho |         25         0.77       100.00
-----------------------------------------+----------------------------------
                            Total |      3,250       100.00
-------------------------------------------------------------------------------------
biomarkers_16                                         Mex-Cog 2016 Sample with Biomarkers
-------------------------------------------------------------------------------------

                type:  numeric (float)

               range:  [0,1]                         units:  1
       unique values:  2                         missing .:  25,053/28,303

    Mex-Cog |
2016 Sample |
       with |
 Biomarkers |      Freq.      Percent         Cum.
------------+-----------------------------------
          0 |      2,502        76.98        76.98
          1 |        748        23.02       100.00
------------+-----------------------------------
      Total |      3,250       100.00


-------------------------------------------------------------------------------------
edta_16                                          Mex-Cog 2016 Sample with EDTA-Lavender Tube
-------------------------------------------------------------------------------------

                type:  numeric (float)

               range:  [0,1]                         units:  1
       unique values:  2                         missing .:  25,053/28,303

       Mex-Cog |
 2016 Sample |
    with EDTA- |
Lavender Tube |      Freq.      Percent         Cum.
--------------+-----------------------------------
            0 |      2,500        76.92        76.92
            1 |        750        23.08       100.00
--------------+-----------------------------------
        Total |      3,250       100.00


-------------------------------------------------------------------------------------
hba1c_16                                               Mex-Cog 2016 Sample with HbA1c
-------------------------------------------------------------------------------------

                type:  numeric (float)
               label:  EDTA, but label does not exist

               range:  [0,1]                         units:  1
       unique values:  2                         missing .:  25,053/28,303

    Mex-Cog |
2016 Sample |
 with HbA1c |      Freq.      Percent         Cum.
------------+-----------------------------------
          0 |      2,109        64.89        64.89
          1 |      1,141        35.11       100.00
------------+-----------------------------------
      Total |      3,250       100.00


-------------------------------------------------------------------------------------
elegible_18                                                  Eligible for interview in 2018
```

```
----------------------------------------------------------------------------------------------------+
                    type:  numeric (float)
                   label:  elegible

                   range:  [0,2]                             units:  1
           unique values:  3                              missing .:  10,328/28,303

  Eligible for interview in |
                       2018 |      Freq.     Percent        Cum.
---------------------------+-----------------------------------
0.Not eligible for interview |      1,054        5.86        5.86
   1.Eligible for interview |     15,698       87.33       93.20
              2.Fallecido |      1,223        6.80      100.00
---------------------------+-----------------------------------
                     Total |     17,975      100.00

----------------------------------------------------------------------------------------------------
new_sample_18                                          Follow-up and new sample/spouses 2018
----------------------------------------------------------------------------------------------------

                    type:  numeric (float)
                   label:  sample18

                   range:  [1,4]                             units:  1
           unique values:  4                              missing .:  6,764/28,303

  Follow-up and new sample/spouses 2018 |      Freq.     Percent        Cum.
-----------------------------------------+-----------------------------------
              1.R from follow-up sample |     15,699       72.89       72.89
2.R from follow-up sample w/o informati |         36        0.17       73.05
   3.New spouse from follow-up sample |         77        0.36       73.41
    4.New person from 2018 new sample |      5,727       26.59      100.00
-----------------------------------------+-----------------------------------
                                 Total |     21,539      100.00

----------------------------------------------------------------------------------------------------
tipne_18                                                        Type of non-interview 2018
----------------------------------------------------------------------------------------------------

                    type:  numeric (byte)
                   label:  nointerview18

                   range:  [1,16]                            units:  1
           unique values:  15                             missing .:  6,764/28,303

         Type of non-interview 2018 |      Freq.     Percent        Cum.
-----------------------------------------+-----------------------------------
              1.Complete Interview |     18,249       84.73       84.73
            2.Incomplete Interview |          7        0.03       84.76
            3.Postponed Interview |        180        0.84       85.59
4.Absence of the subject, or proxy or n |        111        0.52       86.11
5.Absence of occupants or adequate info |        722        3.35       89.46
                         6.Refusal |        932        4.33       93.79
7.Subject not fit for interview and w/o |          3        0.01       93.80
8.Subject deceased without next-of-kin |         11        0.05       93.85
            10.Change of residence |        147        0.68       94.54
         11.Subject could not be located |        546        2.53       97.07
       12.Residence could not be located |         61        0.28       97.35
                 13.Other situation |        425        1.97       99.33
             14.Inhabited residence |         96        0.45       99.77
         15.Dwelling for temporary use |         40        0.19       99.96
14.Dwelling with non-residential use, d |          9        0.04      100.00
-----------------------------------------+-----------------------------------
                             Total |     21,539      100.00

----------------------------------------------------------------------------------------------------
```

```
                    type:  numeric (byte)
                   label:  yesno

                   range:  [1,2]                        units:  1
           unique values:  2                         missing .:  11,189/28,303
```

```
      Same |
dwelling as |
   in 2018 |      Freq.     Percent        Cum.
------------+-----------------------------------
    1.Yes |     16,600       97.00       97.00
     2.No |        514        3.00      100.00
------------+-----------------------------------
    Total |     17,114      100.00
```

----
resul_hh_18   Result of Sections at Household Level 2018

```
                    type:  numeric (float)
                   label:  resul_h

                   range:  [1,3]                        units:  1
           unique values:  3                         missing .:  6,764/28,303
```

```
Result of Sections at |
 Household Level 2018 |      Freq.     Percent        Cum.
----------------------+-----------------------------------
         1.Complete |     17,806       82.67       82.67
       2.Incomplete |        873        4.05       86.72
3.Without Information |      2,860       13.28      100.00
----------------------+-----------------------------------
              Total |     21,539      100.00
```

----
tipent_18   Type of interview 2018

```
                    type:  numeric (float)
                   label:  interview

                   range:  [1,5]                        units:  1
           unique values:  5                         missing .:  10,054/28,303
```

```
      Type of interview 2018 |      Freq.     Percent        Cum.
-----------------------------+-----------------------------------
 1.Direct, follow-up interview |     11,183       61.28       61.28
2.Direct, new sample interview |      4,603       25.22       86.50
  3.Proxy, follow-up interview |      1,122        6.15       92.65
  4.Proxy, new sample interview |        206        1.13       93.78
              5.Next-of-kin |      1,135        6.22      100.00
-----------------------------+-----------------------------------
                      Total |     18,249      100.00
```

----
reason_proxy_18   Reason for Proxy Interview 2018

```
                    type:  numeric (byte)
```

```
            label:  reason

           range:  [1,3]                      units:  1
    unique values:  3                    missing .:  26,975/28,303

   Reason for Proxy |
     Interview 2018 |      Freq.      Percent        Cum.
-------------------+----------------------------------------
  1.Health reasons |        709        53.39       53.39
 2.Language reasons |         45         3.39       56.78
3.Temporary absence |        574        43.22      100.00
-------------------+----------------------------------------
             Total |      1,328       100.00
```

```
--------------------------------------------------------------------------------------------
int_date_18                                                      Interview date 2018
--------------------------------------------------------------------------------------------

             type:  string (str8)

    unique values:  91                   missing "":  11,160/28,303

         examples:  ""
                    "01/12/18"
                    "12/12/18"
```

```
--------------------------------------------------------------------------------------------
c_pap_18                                           Interview technique - Electronic vs Paper
--------------------------------------------------------------------------------------------

             type:  numeric (byte)
            label:  capi18

           range:  [1,2]                      units:  1
    unique values:  2                    missing .:  11,190/28,303

   Interview |
 technique - |
  Electronic |
    vs Paper |      Freq.      Percent        Cum.
-------------+----------------------------------------
1.Electronic |      8,969        52.41       52.41
    2.Paper |      8,144        47.59      100.00
-------------+----------------------------------------
      Total |     17,113       100.00
```

```
--------------------------------------------------------------------------------------------
fallecido_18                                               Died between 2015 and 2018
--------------------------------------------------------------------------------------------

             type:  numeric (float)

           range:  [0,1]                      units:  1
    unique values:  2                    missing .:  6,764/28,303

       Died |
    between |
   2015 and |
       2018 |      Freq.      Percent        Cum.
-------------+----------------------------------------
          0 |     20,404        94.73       94.73
          1 |      1,135         5.27      100.00
-------------+----------------------------------------
```

```
          Total |     21,539       100.00
```

----
tam_loc_18   Locality size 2018

```
              type:  numeric (byte)
              label:  tam_loc

              range:  [1,4]                        units:  1
      unique values:  4                          missing .:  6,764/28,303

        Locality size 2018 |      Freq.      Percent       Cum.
-----------------------------+-----------------------------------
      1.Population = 100,000+ |     12,558       58.30       58.30
2.Population = 15,000 - 99,999 |      2,786       12.93       71.24
 3.Population = 2,500 - 14,999 |      2,199       10.21       81.45
         4.Population <2,500 |      3,996       18.55      100.00
-----------------------------+-----------------------------------
                      Total |     21,539      100.00
```
----
eam_18   Current Residence in High Migration States from 2018

```
              type:  numeric (byte)

              range:  [0,1]                        units:  1
      unique values:  2                          missing .:  6,764/28,303

    Current |
  Residence |
    in High |
  Migration |
States from |
       2018 |      Freq.      Percent       Cum.
------------+-----------------------------------
          0 |     15,140       70.29       70.29
          1 |      6,399       29.71      100.00
------------+-----------------------------------
      Total |     21,539      100.00
```

----
age_18   Age 2018

```
              type:  numeric (float)

              range:  [17,999]                     units:  1
      unique values:  90                         missing .:  10,028/28,303

               mean:   67.2678
          std. dev:   47.6274

    Variable |        Obs        Mean     Std. Dev.       Min        Max
-------------+----------------------------------------------------------
      age_18 |     18,229    64.95315    11.91534         17        115
```

----
sex_18   Sex (Male=1)

```
              type:  numeric (float)
              label:  sex

              range:  [1,2]                        units:  1
      unique values:  2                          missing .:  6,925/28,303
```

```
     Sex |
 (Male=1) |       Freq.      Percent        Cum.
----------+-----------------------------------
   1.Male |       9,165        42.87       42.87
 2.Female |      12,213        57.13      100.00
----------+-----------------------------------
    Total |      21,378       100.00
```