

EDA_17-abr-24

Primera Revisada de Datos

Previo

Bibliotecas

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
v dplyr      1.1.4      v readr      2.1.5  
v forcats    1.0.0      v stringr    1.5.1  
v ggplot2    3.5.0      v tibble     3.2.1  
v lubridate  1.9.3      v tidyr      1.3.1  
v purrr      1.0.2
```

```
-- Conflicts ----- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
```

```
x dplyr::lag()     masks stats::lag()
```

```
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(dplyr)
```

```
library(lubridate)
```

```
library(janitor)
```

Attaching package: 'janitor'

The following objects are masked from 'package:stats':

```
chisq.test, fisher.test
```

```
library(ggplot2)
```

Carga de Datos

```
raw_data <- read.csv("../data/icm-00-23.csv")
glimpse(raw_data)
```

Rows: 70,741

Columns: 9

```
$ Año      <chr> "2020", "2020", "2020", "2020", "2020", "2020", "2020", "20~
$ Entidad  <chr> "NACIONAL", "AGUASCALIENTES", "BAJA CALIFORNIA", "BAJA CALI~
$ Población <chr> "Población total", "Población total", "Población total", "P~
$ Sexo     <chr> "Total", "Total", "Total", "Total", "Total", "Total", "Tota~
$ Edad     <chr> "0 a 17 años", "0 a 17 años", "0 a 17 años", "0 a 17 años",~
$ Cantidad <int> 38247958, 463335, 1061893, 236629, 286880, 982841, 209402, ~
$ Totales  <int> 126014024, 1425607, 3769020, 798447, 928363, 3146771, 73139~
$ Porcentaje <dbl> 0.3035214, 0.3250089, 0.2817425, 0.2963616, 0.3090171, 0.31~
$ Tasa     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
```

Revisión general

Revisamos años

```
unique(raw_data$Año)
```

```
[1] "2020"
[3] "2010"
[5] "2016"
[7] "2012"
[9] "2019"
[11] "2013"
[13] "2009"
[15] "1994"
[17] "2006"
[19] "2005"
[21] "2021"
[23] "2004"

"2015"
"2018"
"2014"
"2008"
"2017"
"2011"
"1.CIFRA SIN AÑO DE REFERENCIA"
"2002"
"2001"
"2007"
"2000"
"2003"
```

[25]	"1964"	"1971"
[27]	"1974"	"1975"
[29]	"1977"	"1981"
[31]	"1987"	"1988"
[33]	"1989"	"1992"
[35]	"1993"	"1995"
[37]	"1996"	"1997"
[39]	"1998"	"1999"
[41]	"2022"	"2023"

```
raw_data %>% filter(Año=="1.CIFRA SIN AÑO DE REFERENCIA")
```

	Año	Entidad
1	1.CIFRA SIN AÑO DE REFERENCIA	CHIAPAS
2	1.CIFRA SIN AÑO DE REFERENCIA	CHIAPAS
3	1.CIFRA SIN AÑO DE REFERENCIA	CHIAPAS
4	1.CIFRA SIN AÑO DE REFERENCIA	YUCATÁN
5	1.CIFRA SIN AÑO DE REFERENCIA	YUCATÁN
6	1.CIFRA SIN AÑO DE REFERENCIA	YUCATÁN
7	1.CIFRA SIN AÑO DE REFERENCIA	NACIONAL
8	1.CIFRA SIN AÑO DE REFERENCIA	NACIONAL
9	1.CIFRA SIN AÑO DE REFERENCIA	NACIONAL

	Población	Sexo	Edad	Cantidad	Totales
1	Personas desaparecidas y no localizadas	Hombre	0 a 17 años	0	NA
2	Personas desaparecidas y no localizadas	Mujer	0 a 17 años	3	NA
3	Personas desaparecidas y no localizadas	Total	0 a 17 años	3	NA
4	Personas desaparecidas y no localizadas	Hombre	0 a 17 años	4	NA
5	Personas desaparecidas y no localizadas	Mujer	0 a 17 años	10	NA
6	Personas desaparecidas y no localizadas	Total	0 a 17 años	14	NA
7	Personas desaparecidas y no localizadas	Hombre	0 a 17 años	1048	NA
8	Personas desaparecidas y no localizadas	Mujer	0 a 17 años	1787	NA
9	Personas desaparecidas y no localizadas	Total	0 a 17 años	2857	NA

	Porcentaje	Tasa
1	NA	NA
2	NA	NA
3	NA	NA
4	NA	NA
5	NA	NA
6	NA	NA
7	NA	NA
8	NA	NA
9	NA	NA

Filtramos estos datos

Igual hay que preguntar que onda con estos datos

```
raw_data_1 <- raw_data %>% filter(Año!="1.CIFRA SIN AÑO DE REFERENCIA") %>%  
  mutate(Año=as.integer(Año)) %>%  
  arrange(Año)  
unique(raw_data_1$Año)
```

```
[1] 1964 1971 1974 1975 1977 1981 1987 1988 1989 1992 1993 1994 1995 1996 1997  
[16] 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012  
[31] 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023
```

Revisar Estados

```
unique(raw_data_1$Entidad)
```

```
[1] "NACIONAL"           "CHIAPAS"           "AGUASCALIENTES"  
[4] "BAJA CALIFORNIA"    "BAJA CALIFORNIA SUR" "CAMPECHE"  
[7] "COAHUILA"           "COLIMA"             "CHIHUAHUA"  
[10] "CIUDAD DE MÉXICO"   "DURANGO"            "GUANAJUATO"  
[13] "GUERRERO"           "HIDALGO"            "JALISCO"  
[16] "MÉXICO"             "MICHOACÁN"          "MORELOS"  
[19] "NAYARIT"            "NUEVO LEÓN"         "OAXACA"  
[22] "PUEBLA"             "QUERÉTARO"          "QUINTANA ROO"  
[25] "SAN LUIS POTOSÍ"    "SINALOA"            "SONORA"  
[28] "TABASCO"            "TAMAULIPAS"         "TLAXCALA"  
[31] "VERACRUZ"           "YUCATÁN"            "ZACATECAS"
```

Revisar Género

```
unique(raw_data_1$Sexo)
```

```
[1] "Hombre"              "Mujer"  
[3] "Total"               "Viviendas particulares habitadas"
```

Revisamos viviendas

Separar viviendas a otra tabla

```
raw_viviendas <- raw_data_1 %>% filter(Sexo=="Viviendas particulares habitadas")
raw_data_2 <- raw_data_1 %>% filter(Sexo!="Viviendas particulares habitadas")
glimpse(raw_viviendas)
```

Rows: 561

Columns: 9

```
$ Año      <int> 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000,~
$ Entidad  <chr> "NACIONAL", "AGUASCALIENTES", "BAJA CALIFORNIA", "BAJA CALI~
$ Población <chr> "No disponen de energía eléctrica", "No disponen de energía~
$ Sexo     <chr> "Viviendas particulares habitadas", "Viviendas particulares~
$ Edad     <chr> "Viviendas particulares habitadas", "Viviendas particulares~
$ Cantidad <int> 988804, 3432, 13149, 5140, 13291, 8772, 2837, 89968, 41659,~
$ Totales  <int> 21512236, 199398, 559402, 104341, 156125, 539169, 124714, 7~
$ Porcentaje <dbl> 0.045964724, 0.017211808, 0.023505458, 0.049261556, 0.08513~
$ Tasa     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
```

```
unique(raw_viviendas$Año)      # Sólo hay tres años
```

```
[1] 2000 2010 2020
```

```
unique(raw_viviendas$Entidad)  # Todos los estados y a nivel Nacional
```

```
[1] "NACIONAL"          "AGUASCALIENTES"      "BAJA CALIFORNIA"
[4] "BAJA CALIFORNIA SUR" "CAMPECHE"            "COAHUILA"
[7] "COLIMA"            "CHIAPAS"             "CHIHUAHUA"
[10] "CIUDAD DE MÉXICO"  "DURANGO"             "GUANAJUATO"
[13] "GUERRERO"         "HIDALGO"             "JALISCO"
[16] "MÉXICO"           "MICHOACÁN"          "MORELOS"
[19] "NAYARIT"          "NUEVO LEÓN"         "OAXACA"
[22] "PUEBLA"           "QUERÉTARO"          "QUINTANA ROO"
[25] "SAN LUIS POTOSÍ"   "SINALOA"            "SONORA"
[28] "TABASCO"          "TAMAULIPAS"         "TLAXCALA"
[31] "VERACRUZ"         "YUCATÁN"            "ZACATECAS"
```

```
unique(raw_viviendas$Sexo)      # Quitar
```

```
[1] "Viviendas particulares habitadas"
```

```
unique(raw_viviendas$Edad)      # Quitar
```

```
[1] "Viviendas particulares habitadas"
```

```
unique(raw_viviendas$Tasa)      # Quitar
```

```
[1] NA
```

```
unique(raw_viviendas$Población) # Son 6 categorías
```

```
[1] "No disponen de energía eléctrica" "No disponen de agua entubada"  
[3] "No disponen de drenaje"           "No disponen de computadora"  
[5] "No disponen de televisor"         "No disponen de Internet"
```

Revisamos cómo es el cálculo de porcentaje:

```
raw_viviendas %>% select(Año, Entidad, Población, Cantidad, Totales, Porcentaje) %>%  
  mutate(revision = round(Porcentaje - Cantidad/Totales,3)) %>%  
  filter(revision!=0)
```

```
[1] Año      Entidad  Población  Cantidad  Totales  Porcentaje revision  
<0 rows> (or 0-length row.names)
```

Revisamos el total de viviendas:

```
raw_viviendas %>% filter(Entidad!="NACIONAL") %>%  
  group_by(Población) %>%  
  summarise(tot = sum(Totales))
```

```
# A tibble: 6 x 2
  Población      tot
  <chr>        <int>
1 No disponen de Internet 63295453
2 No disponen de agua entubada 84807689
3 No disponen de computadora 84807689
4 No disponen de drenaje 84807689
5 No disponen de energía eléctrica 84807689
6 No disponen de televisor 84807689
```

```
raw_viviendas %>% filter(Entidad=="NACIONAL") %>%
  group_by(Población) %>%
  summarise(tot = sum(Totales))
```

```
# A tibble: 6 x 2
  Población      tot
  <chr>        <int>
1 No disponen de Internet 63295453
2 No disponen de agua entubada 84807689
3 No disponen de computadora 84807689
4 No disponen de drenaje 84807689
5 No disponen de energía eléctrica 84807689
6 No disponen de televisor 84807689
```

Revisamos los datos de Internet

```
raw_viviendas %>% filter(Población=="No disponen de Internet") %>%
  group_by(Año) %>%
  summarise(tot = sum(Totales))
```

```
# A tibble: 2 x 2
  Año      tot
  <int>    <int>
1 2010 56277112
2 2020 70313794
```

No hay año 2000 en Internet

Hacemos los cambios mencionados:

```
raw_viviendas_1 <- raw_viviendas %>% select(Año, Entidad, Población, Cantidad, Totales) %>%
  mutate(Porcentaje = Cantidad/Totales)
glimpse(raw_viviendas_1)
```

Rows: 561

Columns: 6

```
$ Año      <int> 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, ~
$ Entidad  <chr> "NACIONAL", "AGUASCALIENTES", "BAJA CALIFORNIA", "BAJA CALI~
$ Población <chr> "No disponen de energía eléctrica", "No disponen de energía~
$ Cantidad <int> 988804, 3432, 13149, 5140, 13291, 8772, 2837, 89968, 41659, ~
$ Totales  <int> 21512236, 199398, 559402, 104341, 156125, 539169, 124714, 7~
$ Porcentaje <dbl> 0.045964724, 0.017211808, 0.023505458, 0.049261556, 0.08513~
```

Revisamos Población

```
unique(raw_data_2$Población)
```

```
[1] "Personas desaparecidas y no localizadas"
[2] "Población total"
[3] "Población sin afiliación a servicio de salud"
[4] "Población que no asiste a la escuela"
[5] "Población con hijos"
[6] "Población casada o en unión libre"
[7] "Tasa de mortalidad infantil"
[8] "Defunciones por suicidio"
[9] "Defunciones por agresiones"
[10] "Población con discapacidad"
[11] "Población hablante de lengua indígena"
[12] "Población indígena con discapacidad"
[13] "Población indígena sin afiliación a servicio de salud"
[14] "Población indígena que no asiste a la escuela"
[15] "Población indígena casada o en unión libre"
[16] "Población indígena con hijos"
[17] "Abandono escolar (primaria)"
[18] "Abandono escolar (secundaria)"
[19] "Población ocupada"
[20] "Población en pobreza (hasta 2018)"
[21] "Inasistencia escolar"
[22] "Rezago educativo (hasta 2018)"
```


- [23] "Carencia por acceso a servicios de salud (hasta 2018)"
- [24] "Carencia por acceso a la alimentación (hasta 2018)"
- [25] "Carencia por calidad y espacios de la vivienda (hasta 2018)"
- [26] "Carencia por servicios básicos de la vivienda (hasta 2018)"
- [27] "Población en pobreza extrema (hasta 2018)"
- [28] "Abandono escolar (media superior)"
- [29] "Repatriaciones desde EUA"
- [30] "Población con condición mental"
- [31] "Población con condición mental sin afiliación a servicio de salud"
- [32] "Población indígena con condición mental"
- [33] "Población con condición mental que no asiste a la escuela"
- [34] "Población con condición mental casada o en unión libre"
- [35] "Población con condición mental con hijos"
- [36] "Inseguridad alimentaria"
- [37] "Población con registro"
- [38] "Víctimas de feminicidio"
- [39] "Índice de calidad de leyes"
- [40] "Población no indígena en pobreza"
- [41] "Población indígena en pobreza"
- [42] "Población no indígena en pobreza extrema"
- [43] "Población indígena en pobreza extrema"
- [44] "Carencia por acceso a la alimentación nutritiva y de calidad (no indígena)"
- [45] "Carencia por acceso a la alimentación nutritiva y de calidad (indígena)"
- [46] "Rezago educativo (no indígena)"
- [47] "Rezago educativo (indígena)"
- [48] "Carencia por acceso a servicios de salud (no indígena)"
- [49] "Carencia por acceso a servicios de salud (indígena)"
- [50] "Carencia por acceso a la seguridad social (no indígena)"
- [51] "Carencia por acceso a la seguridad social (indígena)"
- [52] "Carencia por calidad y espacios de la vivienda (no indígena)"
- [53] "Carencia por calidad y espacios de la vivienda (indígena)"
- [54] "Carencia por servicios básicos de la vivienda (no indígena)"
- [55] "Carencia por servicios básicos de la vivienda (indígena)"
- [56] "Población en pobreza"
- [57] "Población en pobreza extrema"
- [58] "Rezago educativo"
- [59] "Carencia por acceso a servicios de salud"
- [60] "Carencia por acceso a la seguridad social"
- [61] "Carencia por calidad y espacios de la vivienda"
- [62] "Carencia por servicios básicos de la vivienda"
- [63] "Carencia por acceso a la alimentación nutritiva y de calidad"
- [64] "Trabajo infantil"
- [65] "Nacimientos de madres"

```
[66] "Población afrodescendiente"
[67] "Población que nació fuera de la entidad"
[68] "Población afrodescendiente con discapacidad"
[69] "Población afrodescendiente sin afiliación a servicio de salud"
[70] "Población afrodescendiente que no asiste a la escuela"
[71] "Población afrodescendiente casada o en unión libre"
[72] "Población afrodescendiente con hijos"
[73] "Población afrodescendiente con condición mental"
```

Clasificación

Vamos a ir filtrando por tema

Educación

```
raw_educacion <- raw_data_2 %>% filter(Población == "Población que no asiste a la escuela" |
  Población == "Población indígena que no asiste a la escuela" |
  Población == "Abandono escolar (primaria)" |
  Población == "Abandono escolar (secundaria)" |
  Población == "Inasistencia escolar" |
  Población == "Rezago educativo (hasta 2018)" |
  Población == "Abandono escolar (media superior)" |
  Población == "Población con condición mental que no asiste a la escuela" |
  Población == "Rezago educativo (no indígena)" |
  Población == "Rezago educativo (indígena)" |
  Población == "Rezago educativo" |
  Población != "Población afrodescendiente que no asiste a la escuela"
)

raw_data_3 <- raw_data_2 %>% filter(Población != "Población que no asiste a la escuela" &
  Población != "Población indígena que no asiste a la escuela" &
  Población != "Abandono escolar (primaria)" &
  Población != "Abandono escolar (secundaria)" &
  Población != "Inasistencia escolar" &
  Población != "Rezago educativo (hasta 2018)" &
  Población != "Abandono escolar (media superior)" &
  Población != "Población con condición mental que no asiste a la escuela" &
  Población != "Rezago educativo (no indígena)" &
  Población != "Rezago educativo (indígena)" &
  Población != "Rezago educativo" &
```

```
Población != "Población afrodescendiente que no asiste a la escuela"
)
```

Salud

Acá también entra seguridad social??

```
raw_salud <- raw_data_3 %>% filter(Población == "Población sin afiliación a servicio de sa
Población == "Población indígena sin afiliación a servicio de salud"
Población == "Carencia por acceso a servicios de salud (hasta 2018)"
Población == "Población con condición mental sin afiliación a servic
Población == "Carencia por acceso a servicios de salud (no indígena)"
Población == "Carencia por acceso a servicios de salud (indígena)" |
Población == "Carencia por acceso a servicios de salud" |
Población == "Población afrodescendiente sin afiliación a servicio d
Población == "Carencia por acceso a la seguridad social (no indígena)"
Población == "Carencia por acceso a la seguridad social (indígena)"
Población == "Carencia por acceso a la seguridad social"
)
```

```
raw_data_4 <- raw_data_3 %>% filter(Población != "Población sin afiliación a servicio de s
Población != "Población indígena sin afiliación a servicio de salud"
Población != "Carencia por acceso a servicios de salud (hasta 2018)"
Población != "Población con condición mental sin afiliación a servic
Población != "Carencia por acceso a servicios de salud (no indígena)"
Población != "Carencia por acceso a servicios de salud (indígena)" &
Población != "Carencia por acceso a servicios de salud" &
Población != "Población afrodescendiente sin afiliación a servicio d
Población != "Carencia por acceso a la seguridad social (no indígena)"
Población != "Carencia por acceso a la seguridad social (indígena)"
Población != "Carencia por acceso a la seguridad social"
)
```

Vivienda

```
raw_vivienda_2 <- raw_data_4 %>% filter(Población == "Carencia por calidad y espacios de l
Población == "Carencia por servicios básicos de la vivienda (hasta 2
Población == "Carencia por calidad y espacios de la vivienda (no ind
Población == "Carencia por calidad y espacios de la vivienda (indíge
```

```

Población == "Carencia por servicios básicos de la vivienda (no indí
Población == "Carencia por servicios básicos de la vivienda (indígen
Población == "Carencia por calidad y espacios de la vivienda" |
Población == "Carencia por servicios básicos de la vivienda"
)

raw_data_5 <- raw_data_4 %>% filter(Población != "Carencia por calidad y espacios de la vi
Población != "Carencia por servicios básicos de la vivienda (hasta 2
Población != "Carencia por calidad y espacios de la vivienda (no ind
Población != "Carencia por calidad y espacios de la vivienda (indíge
Población != "Carencia por servicios básicos de la vivienda (no indí
Población != "Carencia por servicios básicos de la vivienda (indígen
Población != "Carencia por calidad y espacios de la vivienda" &
Población != "Carencia por servicios básicos de la vivienda"
)

```

Alimentación

```

raw_alimentacion <- raw_data_5 %>% filter(Población == "Carencia por acceso a la alimentac
Población == "Inseguridad alimentaria" |
Población == "Carencia por acceso a la alimentación nutritiva y de c
Población == "Carencia por acceso a la alimentación nutritiva y de c
Población == "Carencia por acceso a la alimentación nutritiva y de c
)

raw_data_6 <- raw_data_5 %>% filter(Población != "Carencia por acceso a la alimentación (h
Población != "Inseguridad alimentaria" &
Población != "Carencia por acceso a la alimentación nutritiva y de c
Población != "Carencia por acceso a la alimentación nutritiva y de c
Población != "Carencia por acceso a la alimentación nutritiva y de c
)

```

Pobreza

```

raw_pobreza <- raw_data_6 %>% filter(Población == "Población en pobreza (hasta 2018)" |
Población == "Población en pobreza extrema (hasta 2018)" |
Población == "Población no indígena en pobreza" |
Población == "Población indígena en pobreza" |
Población == "Población no indígena en pobreza extrema" |

```

```

Población == "Población indígena en pobreza extrema" |
Población == "Población en pobreza" |
Población == "Población en pobreza extrema"
)

raw_data_7 <- raw_data_6 %>% filter(Población != "Población en pobreza (hasta 2018)" &
  Población != "Población en pobreza extrema (hasta 2018)" &
  Población != "Población no indígena en pobreza" &
  Población != "Población indígena en pobreza" &
  Población != "Población no indígena en pobreza extrema" &
  Población != "Población indígena en pobreza extrema" &
  Población != "Población en pobreza" &
  Población != "Población en pobreza extrema"
)

```

Violencia

Es correcto poner suicidio aquí?? Desaparecidos y no localizados???

```

raw_violencia <- raw_data_7 %>% filter(Población == "Defunciones por agresiones" |
  Población == "Tasa de mortalidad infantil" |
  Población == "Víctimas de feminicidio" |
  Población == "Defunciones por suicidio" |
  Población == "Personas desaparecidas y no localizadas"
)

raw_data_8 <- raw_data_7 %>% filter(Población != "Defunciones por agresiones" &
  Población != "Tasa de mortalidad infantil" &
  Población != "Víctimas de feminicidio" &
  Población != "Defunciones por suicidio" &
  Población != "Personas desaparecidas y no localizadas"
)

```

Madres adolescentes o Matrimonio Infantil

```

raw_m_infantil <- raw_data_8 %>% filter(Población == "Población casada o en unión libre" |
  Población == "Nacimientos de madres" |
  Población == "Población indígena casada o en unión libre" |
  Población == "Población con condición mental casada o en unión libre"
)

```

```

        Población == "Población afrodescendiente casada o en unión libre" |
        Población == "Población con hijos" |
        Población == "Población indígena con hijos" |
        Población == "Población con condición mental con hijos" |
        Población == "Población afrodescendiente con hijos"
    )

raw_data_9 <- raw_data_8 %>% filter(Población != "Población casada o en unión libre" &
    Población != "Nacimientos de madres" &
    Población != "Población indígena casada o en unión libre" &
    Población != "Población con condición mental casada o en unión libre" &
    Población != "Población afrodescendiente casada o en unión libre" &
    Población != "Población con hijos" &
    Población != "Población indígena con hijos" &
    Población != "Población con condición mental con hijos" &
    Población != "Población afrodescendiente con hijos"
)

```

Trabajo Infantil

```

raw_trabajo <- raw_data_9 %>% filter(Población == "Población ocupada" |
    Población == "Trabajo infantil"
)
raw_data_10 <- raw_data_9 %>% filter(Población != "Población ocupada" &
    Población != "Trabajo infantil"
)

```

Derecho a Nacionalidad

```

raw_nacionalidad <- raw_data_10 %>% filter(Población == "Repatriaciones desde EUA" |
    Población == "Población con registro"
)
raw_data_11 <- raw_data_10 %>% filter(Población != "Repatriaciones desde EUA" &
    Población != "Población con registro"
)

```

Interseccionalidad

```
unique(raw_data_11$Población)
```

```
[1] "Población total"
[2] "Población con discapacidad"
[3] "Población hablante de lengua indígena"
[4] "Población indígena con discapacidad"
[5] "Población con condición mental"
[6] "Población indígena con condición mental"
[7] "Índice de calidad de leyes"
[8] "Población afrodescendiente"
[9] "Población que nació fuera de la entidad"
[10] "Población afrodescendiente con discapacidad"
[11] "Población afrodescendiente con condición mental"
```

Qué es “Índice de calidad de leyes” ??? Como afecta “Población que nació fuera de la entidad”???

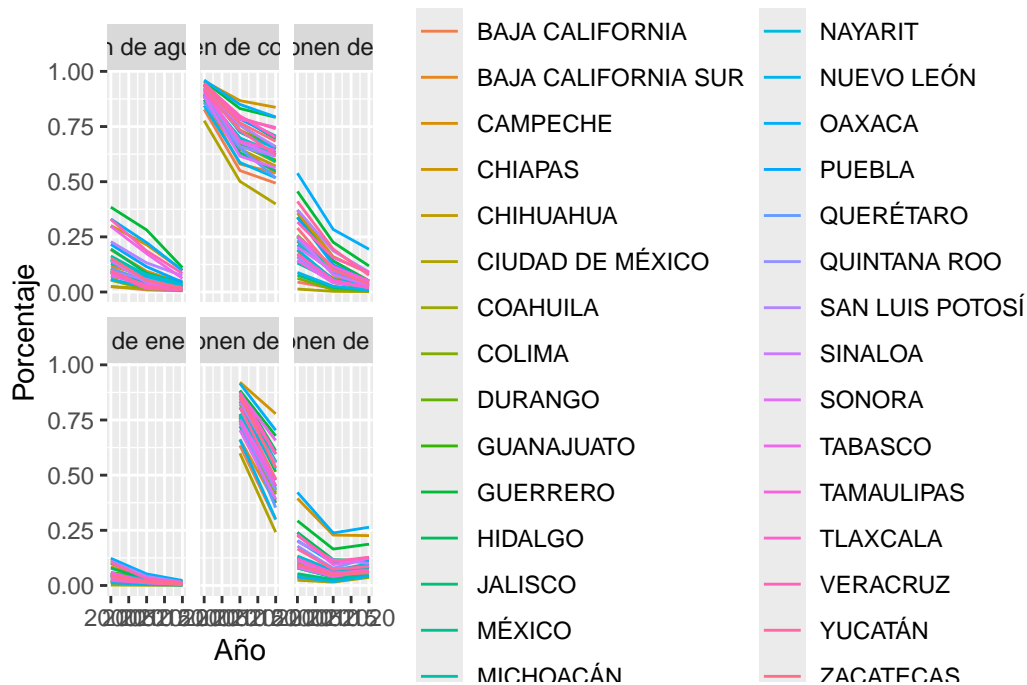
Revisión de cifras

raw_viviendas_1

```
unique(raw_viviendas_1$Población)
```

```
[1] "No disponen de energía eléctrica" "No disponen de agua entubada"
[3] "No disponen de drenaje"           "No disponen de computadora"
[5] "No disponen de televisor"         "No disponen de Internet"
```

```
ggplot(raw_viviendas_1, aes(x=Año, y = Porcentaje, colour = Entidad)) +
  geom_line() +
  facet_wrap(~Población)
```

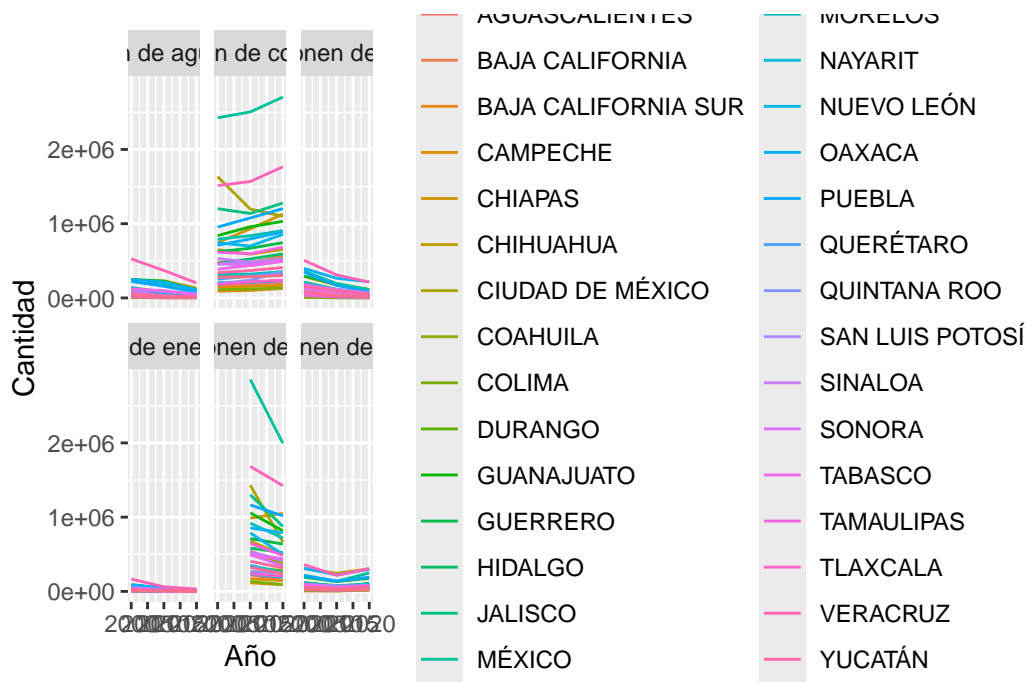


Porqué bajan??

```
unique(raw_viviendas_1$Población)
```

```
[1] "No disponen de energía eléctrica" "No disponen de agua entubada"
[3] "No disponen de drenaje"           "No disponen de computadora"
[5] "No disponen de televisor"         "No disponen de Internet"
```

```
raw_viviendas_1 %>%
  filter(Entidad != "NACIONAL") %>%
  ggplot(aes(x=Año, y = Cantidad, colour = Entidad)) +
  geom_line() +
  facet_wrap(~Población)
```

raw_vivienda_2

```
glimpse(raw_vivienda_2)
```

Rows: 6,748

Columns: 9

```
$ Año      <int> 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008, ~
$ Entidad  <chr> "NACIONAL", "NACIONAL", "AGUASCALIENTES", "BAJA CALIFORNIA"~
$ Población <chr> "Carencia por calidad y espacios de la vivienda (hasta 2018~
$ Sexo     <chr> "Total", "Total", "Total", "Total", "Total", "Total", "Tota~
$ Edad     <chr> "0 a 17 años", "0 a 1 años", "0 a 1 años", "0 a 1 años", "0~
$ Cantidad <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ Totales  <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ Porcentaje <dbl> 0.23018334, 0.24647287, 0.15043096, 0.10332704, 0.19061294, ~
$ Tasa     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
```

```
unique(raw_vivienda_2$Año) # Del 2008-2022 de dos en dos
```

```
[1] 2008 2010 2012 2014 2016 2018 2020 2022
```

```
unique(raw_vivienda_2$Entidad) # Completo con Nacional
```

```
[1] "NACIONAL"          "AGUASCALIENTES"    "BAJA CALIFORNIA"
[4] "BAJA CALIFORNIA SUR" "CAMPECHE"          "COAHUILA"
[7] "COLIMA"            "CHIAPAS"           "CHIHUAHUA"
[10] "CIUDAD DE MÉXICO"   "DURANGO"           "GUANAJUATO"
[13] "GUERRERO"          "HIDALGO"           "JALISCO"
[16] "MÉXICO"            "MICHOACÁN"         "MORELOS"
[19] "NAYARIT"           "NUEVO LEÓN"        "OAXACA"
[22] "PUEBLA"            "QUERÉTARO"         "QUINTANA ROO"
[25] "SAN LUIS POTOSÍ"    "SINALOA"           "SONORA"
[28] "TABASCO"           "TAMAULIPAS"        "TLAXCALA"
[31] "VERACRUZ"          "YUCATÁN"           "ZACATECAS"
```

```
unique(raw_vivienda_2$Edad) # Intervalos de 3-17 y 0-17 años usarlos para comprar
```

```
[1] "0 a 17 años" "0 a 1 años" "2 a 5 años" "6 a 11 años" "12 a 17 años"
[6] "3 a 17 años"
```

```
#unique(raw_vivienda_2$Cantidad) # Algunos NA, investigar donde no hay de: cantidad,
```

Vamos a ver dónde están los NA

```
raw_vivienda_2 %>% filter(is.na(Cantidad)) %>% {unique(.$Población)}
```

```
[1] "Carencia por calidad y espacios de la vivienda (hasta 2018)"
[2] "Carencia por servicios básicos de la vivienda (hasta 2018)"
```

Estas cantidades son la suma hasta 2018?

Limpiamos la parte que no es hasta 2018:

```
raw_vivienda_3 <- raw_vivienda_2 %>% filter(Población!="Carencia por calidad y espacios de
Población!="Carencia por servicios básicos de la vivienda (hasta
glimpse(raw_vivienda_3)
```

Rows: 808

Columns: 9

```
$ Año      <int> 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016,~
$ Entidad  <chr> "NACIONAL", "NACIONAL", "NACIONAL", "NACIONAL", "NACIONAL",~
$ Población <chr> "Carencia por calidad y espacios de la vivienda (no indígen~
$ Sexo     <chr> "Total", "Total", "Total", "Total", "Hombre", "Hombre", "Ho~
$ Edad     <chr> "3 a 17 años", "3 a 17 años", "3 a 17 años", "3 a 17 años",~
$ Cantidad <int> 4687888, 776555, 6353273, 1310777, 3318040, 18825, 65694, 2~
$ Totales  <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Porcentaje <dbl> 0.14467316, 0.46535521, 0.19606869, 0.78549093, 0.16443617,~
$ Tasa     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
```

```
unique(raw_vivienda_3$Totales) # Borrar
```

```
[1] NA
```

```
unique(raw_vivienda_3$Tasa) # Borrar
```

```
[1] NA
```

```
unique(raw_vivienda_3$Edad) #Ojo solo de 3-17 y 0-17
```

```
[1] "3 a 17 años" "0 a 17 años"
```

```
unique(raw_vivienda_3$Sexo) # Hombre Mujer y Total
```

```
[1] "Total" "Hombre" "Mujer"
```

```
unique(raw_vivienda_3$Entidad) # Todos los estados y Nacional
```

```
[1] "NACIONAL" "AGUASCALIENTES" "BAJA CALIFORNIA"
[4] "BAJA CALIFORNIA SUR" "CAMPECHE" "COAHUILA"
[7] "COLIMA" "CHIAPAS" "CHIHUAHUA"
[10] "CIUDAD DE MÉXICO" "DURANGO" "GUANAJUATO"
[13] "GUERRERO" "HIDALGO" "JALISCO"
```

```
[16] "MÉXICO"           "MICHOACÁN"       "MORELOS"
[19] "NAYARIT"          "NUEVO LEÓN"      "OAXACA"
[22] "PUEBLA"           "QUERÉTARO"       "QUINTANA ROO"
[25] "SAN LUIS POTOSÍ"  "SINALOA"         "SONORA"
[28] "TABASCO"          "TAMAULIPAS"      "TLAXCALA"
[31] "VERACRUZ"         "YUCATÁN"         "ZACATECAS"
```

```
unique(raw_vivienda_3$Año)      # 2016, 2018, 2020, 2022
```

```
[1] 2016 2018 2020 2022
```

```
raw_vivienda_4 <- raw_vivienda_3 %>% select(Año, Entidad, Población, Sexo, Edad, Cantidad,
```

limpiamos la parte hasta 2018

```
raw_vivienda_5 <- raw_vivienda_2 %>% filter(Población=="Carencia por calidad y espacios de
Población=="Carencia por servicios básicos de la vivienda (hasta
glimpse(raw_vivienda_5)
```

Rows: 5,940

Columns: 9

```
$ Año      <int> 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008, 2008,~
$ Entidad  <chr> "NACIONAL", "NACIONAL", "AGUASCALIENTES", "BAJA CALIFORNIA"~
$ Población <chr> "Carencia por calidad y espacios de la vivienda (hasta 2018~
$ Sexo     <chr> "Total", "Total", "Total", "Total", "Total", "Total", "Tota~
$ Edad     <chr> "0 a 17 años", "0 a 1 años", "0 a 1 años", "0 a 1 años", "0~
$ Cantidad <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Totales  <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Porcentaje <dbl> 0.23018334, 0.24647287, 0.15043096, 0.10332704, 0.19061294,~
$ Tasa     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
```

```
unique(raw_vivienda_5$Totales)  # Borrar
```

```
[1] NA
```

```
unique(raw_vivienda_5$Tasa)     # Borrar
```

```
[1] NA
```

```
unique(raw_vivienda_5$Cantidad)      # Borrar
```

```
[1] NA
```

```
unique(raw_vivienda_5$Edad)           # "0 a 17 años"  "0 a 1 años"  "2 a 5 años"  "6 a 11 años"
```

```
[1] "0 a 17 años"  "0 a 1 años"  "2 a 5 años"  "6 a 11 años"  "12 a 17 años"
```

```
unique(raw_vivienda_5$Sexo)           # Hombre Mujer y Total
```

```
[1] "Total"  "Hombre" "Mujer"
```

```
unique(raw_vivienda_5$Entidad)        # Todos los estados y Nacional
```

```
[1] "NACIONAL"          "AGUASCALIENTES"      "BAJA CALIFORNIA"
[4] "BAJA CALIFORNIA SUR" "CAMPECHE"             "COAHUILA"
[7] "COLIMA"            "CHIAPAS"              "CHIHUAHUA"
[10] "CIUDAD DE MÉXICO"   "DURANGO"              "GUANAJUATO"
[13] "GUERRERO"          "HIDALGO"              "JALISCO"
[16] "MÉXICO"            "MICHOACÁN"           "MORELOS"
[19] "NAYARIT"           "NUEVO LEÓN"           "OAXACA"
[22] "PUEBLA"            "QUERÉTARO"           "QUINTANA ROO"
[25] "SAN LUIS POTOSÍ"    "SINALOA"              "SONORA"
[28] "TABASCO"           "TAMAULIPAS"          "TLAXCALA"
[31] "VERACRUZ"          "YUCATÁN"              "ZACATECAS"
```

```
unique(raw_vivienda_5$Año)            # 2008 2010 2012 2014 2016 2018
```

```
[1] 2008 2010 2012 2014 2016 2018
```

```
raw_vivienda_6 <- raw_vivienda_5 %>% select(Año, Entidad, Población, Sexo, Edad, Porcentaj
```

Como se comparan estos dos?

```
raw_vivienda_4 %>% filter(Año==2016 | Año==2018) %>%
  filter(Edad=="0 a 17 años") %>%
  group_by(Población) %>%
  summarise(sum_p=sum(Porcentaje))
```

```
# A tibble: 2 x 2
  Población          sum_p
  <chr>          <dbl>
1 Carencia por calidad y espacios de la vivienda  30.5
2 Carencia por servicios básicos de la vivienda  44.1
```

```
raw_vivienda_6 %>% filter(Año==2016 | Año==2018) %>%
  filter(Edad=="0 a 17 años") %>%
  group_by(Población) %>%
  summarise(sum_p=sum(Porcentaje))
```

```
# A tibble: 2 x 2
  Población          sum_p
  <chr>          <dbl>
1 Carencia por calidad y espacios de la vivienda (hasta 2018)  30.5
2 Carencia por servicios básicos de la vivienda (hasta 2018)  44.1
```

Propuestas Sara

```
unique(raw_data_2$Población)
```

```
[1] "Personas desaparecidas y no localizadas"
[2] "Población total"
[3] "Población sin afiliación a servicio de salud"
[4] "Población que no asiste a la escuela"
[5] "Población con hijos"
[6] "Población casada o en unión libre"
[7] "Tasa de mortalidad infantil"
[8] "Defunciones por suicidio"
[9] "Defunciones por agresiones"
[10] "Población con discapacidad"
[11] "Población hablante de lengua indígena"
```

- [12] "Población indígena con discapacidad"
- [13] "Población indígena sin afiliación a servicio de salud"
- [14] "Población indígena que no asiste a la escuela"
- [15] "Población indígena casada o en unión libre"
- [16] "Población indígena con hijos"
- [17] "Abandono escolar (primaria)"
- [18] "Abandono escolar (secundaria)"
- [19] "Población ocupada"
- [20] "Población en pobreza (hasta 2018)"
- [21] "Inasistencia escolar"
- [22] "Rezago educativo (hasta 2018)"
- [23] "Carencia por acceso a servicios de salud (hasta 2018)"
- [24] "Carencia por acceso a la alimentación (hasta 2018)"
- [25] "Carencia por calidad y espacios de la vivienda (hasta 2018)"
- [26] "Carencia por servicios básicos de la vivienda (hasta 2018)"
- [27] "Población en pobreza extrema (hasta 2018)"
- [28] "Abandono escolar (media superior)"
- [29] "Repatriaciones desde EUA"
- [30] "Población con condición mental"
- [31] "Población con condición mental sin afiliación a servicio de salud"
- [32] "Población indígena con condición mental"
- [33] "Población con condición mental que no asiste a la escuela"
- [34] "Población con condición mental casada o en unión libre"
- [35] "Población con condición mental con hijos"
- [36] "Inseguridad alimentaria"
- [37] "Población con registro"
- [38] "Víctimas de feminicidio"
- [39] "Índice de calidad de leyes"
- [40] "Población no indígena en pobreza"
- [41] "Población indígena en pobreza"
- [42] "Población no indígena en pobreza extrema"
- [43] "Población indígena en pobreza extrema"
- [44] "Carencia por acceso a la alimentación nutritiva y de calidad (no indígena)"
- [45] "Carencia por acceso a la alimentación nutritiva y de calidad (indígena)"
- [46] "Rezago educativo (no indígena)"
- [47] "Rezago educativo (indígena)"
- [48] "Carencia por acceso a servicios de salud (no indígena)"
- [49] "Carencia por acceso a servicios de salud (indígena)"
- [50] "Carencia por acceso a la seguridad social (no indígena)"
- [51] "Carencia por acceso a la seguridad social (indígena)"
- [52] "Carencia por calidad y espacios de la vivienda (no indígena)"
- [53] "Carencia por calidad y espacios de la vivienda (indígena)"
- [54] "Carencia por servicios básicos de la vivienda (no indígena)"

```

[55] "Carencia por servicios básicos de la vivienda (indígena)"
[56] "Población en pobreza"
[57] "Población en pobreza extrema"
[58] "Rezago educativo"
[59] "Carencia por acceso a servicios de salud"
[60] "Carencia por acceso a la seguridad social"
[61] "Carencia por calidad y espacios de la vivienda"
[62] "Carencia por servicios básicos de la vivienda"
[63] "Carencia por acceso a la alimentación nutritiva y de calidad"
[64] "Trabajo infantil"
[65] "Nacimientos de madres"
[66] "Población afrodescendiente"
[67] "Población que nació fuera de la entidad"
[68] "Población afrodescendiente con discapacidad"
[69] "Población afrodescendiente sin afiliación a servicio de salud"
[70] "Población afrodescendiente que no asiste a la escuela"
[71] "Población afrodescendiente casada o en unión libre"
[72] "Población afrodescendiente con hijos"
[73] "Población afrodescendiente con condición mental"

```

```

datos_1 <-raw_data_2 %>%
  filter(
    Población == "Población afrodescendiente" |
    Población == "Población afrodescendiente que no asiste a la escuela" |
    Población == "Población afrodescendiente con hijos" |
    Población == "Población afrodescendiente sin afiliación a servicio de salud" |
    Población == "Población afrodescendiente casada o en unión libre" |
    Población == "Población afrodescendiente con condición mental" |

    Población == "Población hablante de lengua indígena" |
    Población == "Población indígena que no asiste a la escuela" |
    Población == "Población indígena con hijos" |
    Población == "Población indígena sin afiliación a servicio de salud" |
    Población == "Población indígena casada o en unión libre" |
    Población == "Población indígena con condición mental" |

    Población == "Población con condición mental" |
    Población == "Población con condición mental que no asiste a la escuela" |
    Población == "Población con condición mental con hijos" |
    Población == "Población con condición mental sin afiliación a servicio de salud" |
    Población == "Población con condición mental casada o en unión libre"
  ) %>%

```



```
select(Año, Entidad, Población, Sexo, Edad, Cantidad, Totales, Porcentaje)
```

```
datos_1 <- datos_1 %>% mutate(varios=paste(Población,Edad)) %>% arrange(varios)
```

```
unique(datos_1$varios)
```

```
[1] "Población afrodescendiente 0 a 17 años"
[2] "Población afrodescendiente 0 a 4 años"
[3] "Población afrodescendiente 10 a 14 años"
[4] "Población afrodescendiente 15 a 17 años"
[5] "Población afrodescendiente 5 a 9 años"
[6] "Población afrodescendiente casada o en unión libre 12 a 14 años"
[7] "Población afrodescendiente casada o en unión libre 12 a 17 años"
[8] "Población afrodescendiente casada o en unión libre 15 a 17 años"
[9] "Población afrodescendiente con condición mental 0 a 17 años"
[10] "Población afrodescendiente con condición mental 0 a 4 años"
[11] "Población afrodescendiente con condición mental 10 a 14 años"
[12] "Población afrodescendiente con condición mental 15 a 17 años"
[13] "Población afrodescendiente con condición mental 5 a 9 años"
[14] "Población afrodescendiente con hijos 12 a 14 años"
[15] "Población afrodescendiente con hijos 12 a 17 años"
[16] "Población afrodescendiente con hijos 15 a 17 años"
[17] "Población afrodescendiente que no asiste a la escuela 12 a 14 años"
[18] "Población afrodescendiente que no asiste a la escuela 15 a 17 años"
[19] "Población afrodescendiente que no asiste a la escuela 3 a 17 años"
[20] "Población afrodescendiente que no asiste a la escuela 3 a 5 años"
[21] "Población afrodescendiente que no asiste a la escuela 6 a 11 años"
[22] "Población afrodescendiente sin afiliación a servicio de salud 0 a 17 años"
[23] "Población afrodescendiente sin afiliación a servicio de salud 0 a 4 años"
[24] "Población afrodescendiente sin afiliación a servicio de salud 10 a 14 años"
[25] "Población afrodescendiente sin afiliación a servicio de salud 15 a 17 años"
[26] "Población afrodescendiente sin afiliación a servicio de salud 5 a 9 años"
[27] "Población con condición mental 0 a 17 años"
[28] "Población con condición mental 0 a 4 años"
[29] "Población con condición mental 10 a 14 años"
[30] "Población con condición mental 15 a 17 años"
[31] "Población con condición mental 5 a 9 años"
[32] "Población con condición mental casada o en unión libre 12 a 14 años"
[33] "Población con condición mental casada o en unión libre 12 a 17 años"
[34] "Población con condición mental casada o en unión libre 15 a 17 años"
[35] "Población con condición mental con hijos 12 a 14 años"
```

- [36] "Población con condición mental con hijos 12 a 17 años"
- [37] "Población con condición mental con hijos 15 a 17 años"
- [38] "Población con condición mental que no asiste a la escuela 12 a 14 años"
- [39] "Población con condición mental que no asiste a la escuela 15 a 17 años"
- [40] "Población con condición mental que no asiste a la escuela 3 a 17 años"
- [41] "Población con condición mental que no asiste a la escuela 3 a 5 años"
- [42] "Población con condición mental que no asiste a la escuela 6 a 11 años"
- [43] "Población con condición mental sin afiliación a servicio de salud 0 a 17 años"
- [44] "Población con condición mental sin afiliación a servicio de salud 0 a 4 años"
- [45] "Población con condición mental sin afiliación a servicio de salud 10 a 14 años"
- [46] "Población con condición mental sin afiliación a servicio de salud 15 a 17 años"
- [47] "Población con condición mental sin afiliación a servicio de salud 5 a 9 años"
- [48] "Población hablante de lengua indígena 10 a 14 años"
- [49] "Población hablante de lengua indígena 15 a 17 años"
- [50] "Población hablante de lengua indígena 3 a 17 años"
- [51] "Población hablante de lengua indígena 3 a 4 años"
- [52] "Población hablante de lengua indígena 5 a 17 años"
- [53] "Población hablante de lengua indígena 5 a 9 años"
- [54] "Población indígena casada o en unión libre 12 a 14 años"
- [55] "Población indígena casada o en unión libre 12 a 17 años"
- [56] "Población indígena casada o en unión libre 15 a 17 años"
- [57] "Población indígena con condición mental 10 a 14 años"
- [58] "Población indígena con condición mental 15 a 17 años"
- [59] "Población indígena con condición mental 3 a 17 años"
- [60] "Población indígena con condición mental 3 a 4 años"
- [61] "Población indígena con condición mental 5 a 9 años"
- [62] "Población indígena con hijos 12 a 14 años"
- [63] "Población indígena con hijos 12 a 17 años"
- [64] "Población indígena con hijos 15 a 17 años"
- [65] "Población indígena que no asiste a la escuela 12 a 14 años"
- [66] "Población indígena que no asiste a la escuela 15 a 17 años"
- [67] "Población indígena que no asiste a la escuela 3 a 17 años"
- [68] "Población indígena que no asiste a la escuela 3 a 5 años"
- [69] "Población indígena que no asiste a la escuela 5 a 17 años"
- [70] "Población indígena que no asiste a la escuela 5 años"
- [71] "Población indígena que no asiste a la escuela 6 a 11 años"
- [72] "Población indígena sin afiliación a servicio de salud 10 a 14 años"
- [73] "Población indígena sin afiliación a servicio de salud 15 a 17 años"
- [74] "Población indígena sin afiliación a servicio de salud 3 a 17 años"
- [75] "Población indígena sin afiliación a servicio de salud 3 a 4 años"
- [76] "Población indígena sin afiliación a servicio de salud 5 a 17 años"
- [77] "Población indígena sin afiliación a servicio de salud 5 a 9 años"