



# Curaduría de datos abiertos del INEGI

Guerrero, Sepúlveda Claudia  
Universidad Iberoamericana Torreón

## Resumen

*Los datos abiertos que proporciona el INEGI requieren de tratamiento y análisis preliminares derivados del volumen, de la complejidad de las variables, de las consideraciones del muestreo, de la interpretación de los indicadores reportados en los tabulados básicos y de la necesidad de generar información de zonas localizadas dentro de territorios particulares. Se presenta el tratamiento de los datos censales en un conjunto de manzanas y AGEB seleccionadas; el proceso de variables calculadas en los tabulados básicos de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2016 —ENDIREH—, el análisis de independencia de dos de ellas, así como del uso de los factores de expansión para las estimaciones poblacionales de las unidades de muestreo. Se realizó en el ambiente de programación del lenguaje R y la paquetería necesaria los paquetes utilizados en el procesamiento de los datos.*

**Palabras clave:** Censos en AGEBs, análisis de independencia, factores de expansión, datos abiertos, INEGI, ENDIREH.

## 1. Introducción

La política de datos abiertos permite el acceso sin restricciones de derechos de autor, a la información pública de las distintas instituciones gubernamentales. [1]

El sitio WEB del Instituto nacional de estadística y geografía INEGI contiene información de México en cuanto al territorio, los recursos, la población y las actividades económicas que ayudan a la toma de decisiones y a la generación, y administración de políticas públicas en los distintos niveles de gobierno. Las fuentes que opera el instituto son censos, encuestas y registros administrativos a partir de los cuales produce indicadores demográficos, sociales y económicos.

El INEGI elabora reportes alrededor de los datos que genera y comparte las bases de datos, estructuras y metodologías para que puedan ser exploradas y analizadas por los ciudadanos.

Los datos censales se registran a nivel manzana y se realizan reportes a nivel nacional, estatal y municipal, en este documento se

presenta la gestión de datos en zonas determinadas por intereses particulares con niveles de agrupamiento distinto a nivel de municipios, localidades o por áreas geográficas básicas (AGEB). Se presentan los casos de los reportes censales de algunas zonas de la ciudad de Torreón a nivel manzana y AGEB.

Por otro lado las encuestas realizadas en hogares o en establecimientos con datos dedicados a temas específicos se reportan en indicadores que son conocidos como tabulados básicos, y éstos resultan no necesariamente de las respuestas directas de las personas sino de alguna combinación de variables que pueden ser analizadas estadísticamente para determinar independencia, correlación o correspondencia, y los estimadores poblacionales determinados en los factores de expansión de las unidades de muestreo que utilizan. En este apartado se trabajó con la encuesta ENDIREH que es una encuesta especial del INEGI que estudia la violencia de género y las relaciones de los hogares y se calculan los tabulados básicos a nivel municipal.

En el desarrollo del tratamiento de los datos se trabaja en el ambiente de



programación del lenguaje R, RStudio, Markdown, y los paquetes del repositorio CRAN [2] que cubren los distintos métodos estadísticos y que son distribuidos gratuitamente bajo los términos de la GNU.

Los datos y las herramientas de acceso abierto plantean oportunidades de generación de conocimiento distintos a los convencionales, que pueden ser utilizados tanto por particulares para el desarrollo de aplicaciones y servicios como por los académicos universitarios para llevar a cabo sus funciones de docencia, investigación y difusión.

## 2. CENSOS

Los datos para la proyección de los mapas se encuentran en el marco geoestadístico que proporciona el INEGI [3] y los datos censales en el sitio del Censo de población y vivienda de 2010 [4]

La Cartografía de Coahuila, tiene el nivel de la manzana y de AGEB en la capa 05M y 05A, se filtra la información para el municipio de Torreón (35) y se obtiene el mapa de los polígonos de manzanas en azul y de AGEB en rojo.

```
# datos censo a nivel manzana, se selecciona la capa de la manzana
# Lectura del archivo del CENSO de 2010 obtenido para Coahuila
# Se unen los datos geográficos con los datos censales
# lectura de archivo con información censal a nivel del AGEB
# Elección del municipio de Torreón (35)
# Se superponen las capas de las manzanas con las de AGEB
mapagebmun<-ggplot(torreon.map.df)+
aes(long,lat,group=group)+geom_path(color="blue") +
  geom_path(data=torreonageb.map.df, color="red", size=0.5)
mapagebmun
```

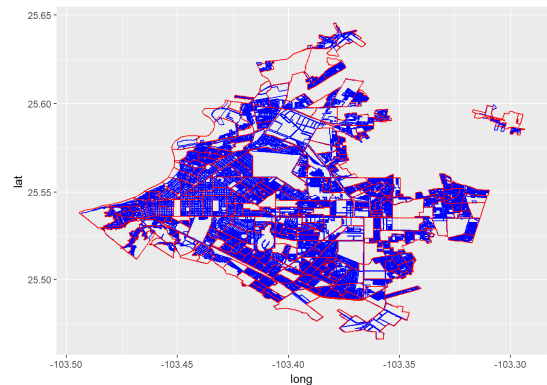


Fig. 1 AGEBs y manzanas de Torreón Coahuila

### 2.1 Mapas censales, manzanas y AGEB

Una vez establecido el polígono de cada una de las secciones geográficas la representación gráfica de los datos numéricos del número de personas que reporta el censo se reporta en rangos que representan distintos colores, cómo se muestra en el siguiente mapa.

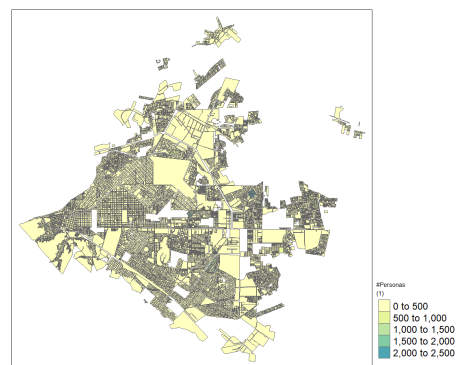
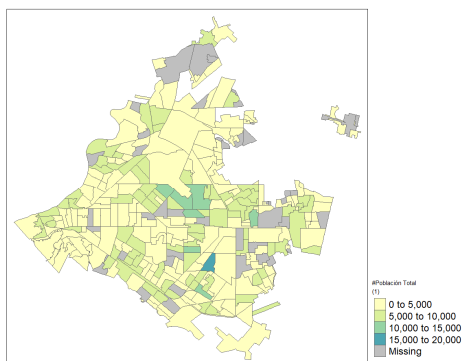


Fig. 2 Población a nivel de manzana Torreón

Los datos que proporciona el INEGI se encuentran a nivel de manzana, por lo que es necesario agrupar y sumar a nivel de AGEB y dar origen a la siguiente gráfica.

```
dfagebcensocom<-aggregate(. ~cvegeoageb, data=dfageb, sum,
na.rm=TRUE)
```

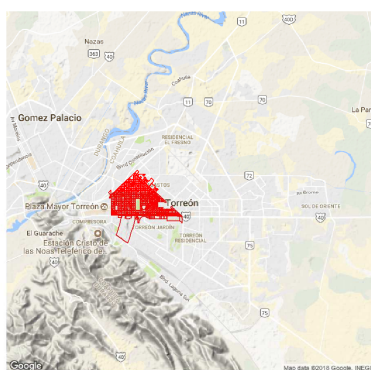


**Fig. 3 Población a nivel AGEB Torreón**

## 2.2 Elección de zonas particulares

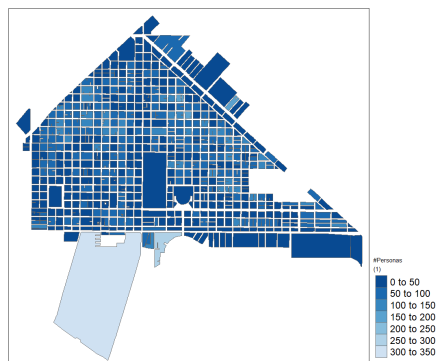
Se pueden determinar datos censales de zonas determinadas por otros límites, para conocer las manzanas que corresponden a la zona que se quiere estudiar el sitio del INEGI proporciona el servicio en la sección llamada Espacio y Datos [5]

La elección se hace mediante el polígono y se solicita en la zona a nivel manzana y se pide la descarga de archivo con los datos de las manzanas correspondiente, con los datos de la manzana y el AGEB se crea la clave geográfica correspondiente utilizada para recuperar los datos del censo.



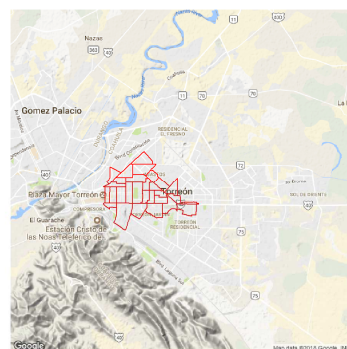
**Fig. 4 Zona seleccionada límites particulares**

Con los datos de las personas que habitan la zona seleccionada se obtienen mapas de la zona elegida.



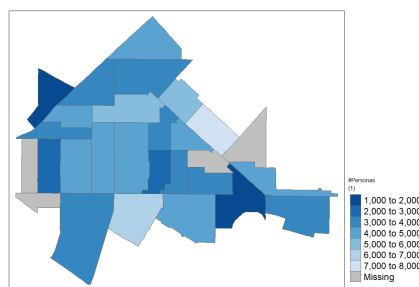
**Fig. 5 Censo de población zona seleccionada**

La selección también puede hacerse a nivel de AGEB



**Fig. 6 Zona seleccionada AGEBS**

Y los datos censales correspondientes



**Fig. 7 Censo por AGEB zona seleccionada**

La información agrupada de acuerdo a las necesidades que plantean los distintos problemas sociales o económicos es posible determinarla una vez que se determinan las

### 3. Encuestas ENDIREH

Se trata de una tabla combinada de la variable del estado conyugal y de los ingresos de acuerdo a la siguiente característica

1. Sin ingresos propios o de otras fuentes
2. Sólo dispone de ingresos propios
3. Dispone de ingresos propios y algún otro apoyo (ayudas familiares, becas escolares o programas gubernamentales)
4. Dispone de ingresos por ayuda de familiares o conocidos y de otros apoyos (programas gubernamentales y/o becas)
5. Solo dispone de ingresos por apoyos de programas gubernamentales y/o becas para ella o para sus hijas/os

La variable relacionada con los ingresos está en considerada en el marco conceptual de acuerdo a la siguiente tabla [7]

Ítem	IV. SITUACIÓN DE LA RELACIÓN DE PAREJA/INGRESOS Y RECURSOS		
Subítem	Ingresos		
Variable	Definición	Reactivos cuestionario	Respuestas
	Obtener información sobre los ingresos económicos de la mujer elegida.	4.1. Actualmente, ¿usted trabaja por un ingreso, salario o ganancia? 4.2. Aproximadamente, ¿cuánto gana o recibe usted por su trabajo?	Ingresos monetarios por trabajo. Ingresos monetarios por remuneración / igual o más de 999.997 / No
Ingresos propios	4.1.; 4.2		
	Identificación de las mujeres que reciben dinero por fuentes diferentes a las del trabajo o algunos distinto al esposo, (pensión, parental (del padre), etc.)	4.8. Otra fuente de ingresos de la mujer elegida Por jubilación o pensión / De familiares o conocidos que viven en Estados Unidos de América (No incluye a su esposo o pareja) / De familiares o conocidos dentro del país (No incluye al esposo o pareja) / Por becas escolares para sus hijos o hijas / Por becas escolares para usted / Del Programa Prospera (antes Oportunidades) / De otro programa del gobierno / Otro tipo de ingresos (recursos familiares, etc.)	4.8.1. 4.8.2(a,b,c); 4.8.4; 4.8.5(a,b,c,d,e,f,g,h,i); 4.8.5(f)/No
Aportaciones de otros	4.8.; 4.9.; 4.10		

**Tabla 2 Variables y reactivos del cuestionario**

Y que pueden observarse en la siguiente imagen del cuestionario.

<p><b>4.1. Actualmente, ¿usted trabaja por un ingreso, salario o ganancia?</b></p> <p style="text-align: center;">REGISTRE SÓLO UN CÓDIGO</p> <p>Si <u>1</u></p> <p>No <u>2</u> → <b>FASE A 4.3</b></p>	<p><b>4.2. Aproximadamente, ¿cuánto gana o recibe usted por su trabajo?</b></p> <p style="text-align: center;">REGISTRE LA CANTIDAD Y EL CÓDIGO CORRESPONDIENTE AL PERÍODO</p> <p style="text-align: center;">\$ <u>1</u> <u>1</u> <u>1</u> <u>1</u></p> <p>             Igual o más de \$999.997 <b>999.997</b>              No recibe ingresos <b>000.000</b> → <b>FASE A 4.3</b>              No sabe <b>999.998</b> </p> <p style="text-align: right;"><b>4.2.1. ¿Cada cuánto?</b></p> <p>             A la semana <u>1</u>              A la quincena <u>2</u>              Al mes <u>3</u>              No sabe <u>6</u> </p>
<p><b>4.3. Actualmente, ¿su esposo o pareja trabaja por un ingreso, salario o ganancia?</b></p> <p style="text-align: center;">REGISTRE SÓLO UN CÓDIGO</p> <p>Si <u>1</u></p> <p>No <u>2</u> → <b>FASE A 4.6B</b></p>	<p><b>4.4. En su trabajo, ¿cuál es la ocupación que desempeña su esposo o pareja (o a qué se dedica principalmente)?</b></p> <p style="font-size: small;">[Ejemplo: electricista, maestro de primaria, vendedor de frutas, médico, cocinero, albañil, mecánico de autos, técnico, policía, militar, militar, etc.]</p> <p style="text-align: center;">REGISTRE LA CANTIDAD Y EL CÓDIGO CORRESPONDIENTE AL PERÍODO</p> <p style="text-align: center;">\$ <u>1</u> <u>1</u> <u>1</u> <u>1</u></p> <p>             Igual o más de \$999.997 <b>999.997</b>              No recibe ingresos <b>000.000</b> → <b>FASE A 4.6B</b>              No sabe <b>999.998</b> </p> <p style="text-align: right;"><b>4.4.1AB ¿Cada cuánto?</b></p> <p>             A la semana <u>1</u>              A la quincena <u>2</u>              Al mes <u>3</u>              No sabe <u>6</u> </p> <p style="text-align: center;"><b>ANOTE COMPLETO EL NOMBRE DE LA OCUPACIÓN</b></p>

[illegible]

## Tabla 1 Tabulado básica sección IV ENDIREH 2016





10  
11  
12  
OCTUBRE

9° Encuentro de Investigadores  
8° Encuentro de jóvenes investigadores  
2° Concurso de Tesis Doctorales



**4.8 Actualmente, ¿usted recibe dinero...**  
REGISTRE EL CÓDIGO CORRESPONDIENTE EN CADA COLUMNA.  
PARA LOS CÓDIGOS 1, PREGUNTE 4.9 Y 4.10; SI RESPONDIÓ EN TODAS, CÓDIGO 2, PASE A 4.11

Si .....1  
No .....2

**4.9 Aproximadamente, ¿cuánto recibe al mes por... (RESPUESTA "SI" (1) en 4.8)?**  
REGISTRE LA CANTIDAD CORRESPONDIENTE  
Igual o más de \$999.997 .....999.997  
No sabe .....999.998

**4.10 ¿Quiénes le ayudan con dinero? (No incluya lo que su esposo o pareja aporta para los gastos del hogar)**  
REGISTRE LOS CÓDIGOS CORRESPONDIENTES  
Padre o madre .....1  
Hija(s)/hijos .....2  
Hermana(s)/hermano(s) .....3  
Otras familiares .....4  
Personas conocidas .....5

**1. por jubilación o pensión?** .....  
**2. de familiares o conocidos que viven en Estados Unidos de América?** (No incluya a su esposo o pareja) .....  
**3. de familiares o conocidos dentro del país?** (No incluya al esposo o pareja) .....  
**4. por becas escolares para sus hijos e hijas?** .....  
**5. por becas escolares para usted?** .....  
**6. del Programa Prospera?** (antes Oportunidades) .....  
**7. de otro programa del gobierno?** .....  
**8. otro tipo** (rentas, intereses bancarios, etc.) .....  
**4.11 ¿Usted cuenta con dinero que puede utilizar como quiera?**  
REGISTRE SÓLO UN CÓDIGO  
Si .....1  
No .....2

Fig. 9 Cuestionario sobre ingresos y recursos

### 3.1 Creación de nuevas variables (Sección IV, ingresos)

Las respuestas no corresponden a la variable que se tabula por lo que es necesario la creación de las variables que corresponden a la combinación de las respuestas a las diferentes preguntas.

Es necesario desde cada pregunta la variable compuesta, para determinar si solo dispone de ingresos propios (salario y pensiones), si no tiene ingresos de ningún tipo y la combinación de si tiene ingreso proveniente de ayudas familiares o del gobierno.

Se consideró el siguiente espacio muestral de las probabilidades de ingreso.

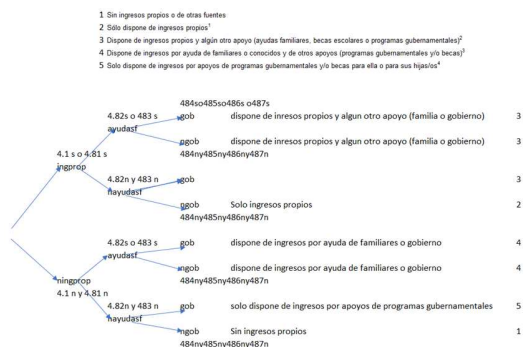


Fig. 10 Muestra de posibles respuestas

Así Sin ingresos propios o de otras fuentes, responde que no a 4.1,4.8.1: 4.8.2;4.8.3; 4.8.4; 4.8.5;4.8.6;4.8.7;4.8.8

Los valores muestrales para los conteos de las distintas respuestas de las mujeres seleccionadas y su estado civil se muestran en la siguiente tabla.

Situación de ingresos transpuesta	Cas o uni	Sep.div,viuda	Soltera
ingayudasfamgob	2652	1764	1391
ingpropyayudasfamgob	4768	3031	1098
siningpropnietrafuente	29550	2798	6266
soloayudagob	12729	2676	1871
soloingprop	23136	9321	8185

Tabla 3 Ingresos y estado civil muestrales

Interpretación de la tabla de contingencia, el tamaño de los puntos refleja la magnitud de los valores de la tabla

### Disponibilidad de ingresos según estado conyugal

	Cas o uni	Sep.div,viuda	Soltera
ingayudasfamgob	●	●	●
ingpropyayudasfamgob	●	●	●
siningpropnietrafuente	●	●	●
soloayudagob	●	●	●
soloingprop	●	●	●

Tabla 4 Tabla gráfica de valores obtenidos

### 3.2 Independencia Chi cuadrado

Cuando las tablas de contingencia no son muy grandes o no es fácil visualizar e interpretar la relación entre las columnas y renglones se utilizan otros métodos estadísticos.

Uno de ellos es la prueba de independencia de Xi cuadrado Es una prueba de hipótesis utilizada en tablas de contingencia para probar la independencia entre variables, esto es si están significativamente asociadas.

La hipótesis nula dice que el ingreso es independiente del estado civil, la alternativa es que están asociados. El resultado del cálculo de la Chi cuadrada da un valor de Xi de 7759.5 con 8 grados de libertad y un valor pvalue de 2.2 e-16



El nivel de significancia (.05) es mayor por lo que se rechaza la hipótesis nula lo que apunta a que hay evidencia para concluir que la alternativa podría ser la correcta, esto es que los ingresos se asocian al estado civil. Para saber cuál es el cruce o la celda que más contribuye al resultado de la Chi cuadrada se utilizan los residuales, esto es la Chi cuadrada de cada celda

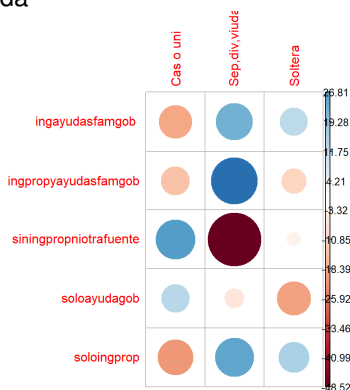


Fig. 11 Valor de residuales de la CHI cuadrada

El tamaño del círculo en cada celda es proporcional a la cantidad de contribución. El signo de cada valor residual también es importante para interpretar la asociación entre las columnas y los renglones de acuerdo a lo siguiente. La contribución en términos de porcentajes es la relación entre cada residual al cuadrado y la xi cuadrada total. Se calcula como contribuye en %

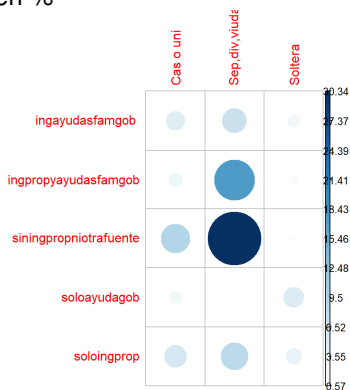


Fig. 12 Contribución del cruce en la Chi cuadrada

### 3.3 Gráfico de mosaico

Se utiliza en las tablas de contingencia para examinar la asociación entre variables categóricas muestra la diferencia gráfica entre lo observado y esperado

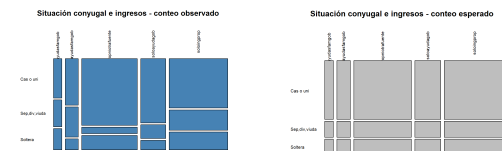


Fig. 13 Mosaico observados y esperados

### 3.4 Uso de factores de expansión (Sección X, violencia en el ámbito familiar)

La violencia en el ámbito familiar se recupera de las preguntas elaboradas en el tema del capítulo X y que se muestran a continuación.

Fig. 14 Reactivos del cuestionario sección X

También es necesario crear nuevas variables que incluya una respuesta positiva a cualquiera de las 18 preguntas de esta primera pregunta y de acuerdo a las variables que las determinan se crea la siguiente clasificación de tipos de violencia.



10  
11  
12  
OCTUBRE

9° Encuentro  
de Investigadores  
8° Encuentro de  
jóvenes investigadores  
2° Concurso  
de Tesis Doctorales



Condición, tipo, clase y situación de violencia	Absolutos	Relativos
Total de mujeres de 15 años y más	46 501 740	100.0
Sin incidentes de violencia familiar	41 706 326	89.7
Con incidentes de violencia familiar	4 795 414	10.3
<b>Violencia emocional</b>	<b>3 782 923</b>	<b>8.1</b>
Degradación emocional	2 185 123	4.7
La han ofendido o humillado por el hecho de ser mujer (la hicieron sentir menos o mal)	1 180 540	2.5
La han ignorado o no la han tomado en cuenta por ser mujer	1 703 572	3.7
Intimidación	1 243 102	2.7
Le han roto o escondido algún objeto personal	1 243 102	2.7
Amenaza	1 540 551	3.3
La han amenazado con hacerle daño a usted o alguien que a usted le importa	564 477	1.2
La han corrido de su casa o amenazado con correrla	1 237 763	2.7
Aislamiento	457 039	1.0
La han encerrado o impedido salir de su casa	457 039	1.0
<b>Violencia económica o patrimonial</b>	<b>1 439 366</b>	<b>3.1</b>
Control o chantaje económico	785 243	1.7
La han impedido o prohibido estudiar o trabajar	785 243	1.7
Despojo y coerción	800 530	1.7
La han obligado a poner a nombre de otra persona alguna propiedad de usted o le han quitado o robado papeles de alguna propiedad	149 434	0.3
La han quitado su dinero o lo han usado sin su consentimiento	585 548	1.3
Le han quitado bienes o propiedades de usted	212 773	0.5
<b>Violencia física</b>	<b>1 312 240</b>	<b>2.8</b>
La han pellizcado, jalado el cabello, empujado, jaloneado, abofeteado o aventado algún objeto	1 028 305	2.2
La han atacado o agredido con un cuchillo, navaja o arma de fuego	102 424	0.2
La han pateado o golpeado con el puño	583 516	1.3
<b>Violencia sexual</b>	<b>526 212</b>	<b>1.1</b>
Abuso sexual	465 850	1.0
Alguna persona le mostró sus partes íntimas o se las manoseó enfrente de usted	170 952	0.4
La han obligado a mirar escenas o actos sexuales o pornográficos (fotos, revistas, videos o películas pornográficas)	37 715	0.1
La han manoseado, tocado, besado o se le han arrimado, recargado o encimado sin su consentimiento	389 108	0.8
Violación e intento de violación	217 289	0.5
Han tratado de obligarla a tener relaciones sexuales en contra de su voluntad	181 510	0.4
La han obligado a tener relaciones sexuales en contra de su voluntad	137 683	0.3

**Tabla 5 Definición de tipos de violencia sección X**

La violencia emocional está compuesta por la degradación emocional, la intimidación, la amenaza y el aislamiento, se trabajó con el complemento de las distintas combinaciones, cuando no hay ninguna de ellas entonces no hay violencia emocional, y el complemento es quienes si la sufren.

La violencia económica está compuesta de la patrimonial, del control o chantaje económico, del despojo y de la coerción, de igual forma cuando no existe ninguna de ellas no hay violencia económica, pero si hubiera al menos una de ellas entonces si existiría la violencia económica.

La violencia física está determinada por las respuestas de agresión solamente, y la violencia sexual, se debe a un abuso sexual, a una violación o intento de, de igual forma si no existe alguno de ellos entonces no hubo violencia sexual

Estos cuatro tipos de violencia dentro del ámbito familiar de igual forma se combinan para

determinar si la persona encuestada ha padecido o no incidentes de violencia.

Los resultados muestrales que se obtienen son:

Sí	No
9.4%	90.6%

Y coinciden con los tabulados básicos que se obtienen para el capítulo X del indicador 1.1

Distribución de mujeres de 15 años y más por condición, tipo, clase y situación de violencia en el ámbito familiar en los últimos 12 meses  
2016  
Observaciones muestrales

Condición, tipo, clase y situaciones de violencia	Absolutos	Relativos
Total de mujeres de 15 años y más	111 256	100.0
Sin incidentes de violencia familiar	100 780	90.6
Con incidentes de violencia familiar	10 476	9.4
<b>Violencia emocional</b>	<b>8 169</b>	<b>7.3</b>
Degradación emocional	4 923	4.4
La han ofendido o humillado por el hecho de ser mujer (la hicieron sentir menos o mal)	2 617	2.4

**Tabla 6 Tabulado básico indicador 1.1**

Las encuestas del INEGI se aplican a muestras complejas en las que se estratifica (se divide a la población en grupos relativamente homogéneos (estratos) y se toma la muestra con un número determinado por cada uno de ellos, y a través de los clústeres, dividiendo a la población y muestreo en grupos lo que disminuye la precisión de un cierto tamaño de muestra, pero a menor costo. Cuando se trabaja con los datos abiertos, los valores de la encuesta resultan de observaciones que no tienen todas las mismas probabilidades de selección ni son independientes, por los que para tomar en cuenta esta probabilidad y poder aplicarlo a los estimadores de la población que resultarían sesgados es necesario tomar en cuenta las estimaciones para muestras complejas. Por esa razón el INEGI presenta en los informes de Tabulados Básicos los resultados para la muestra y además los resultados de los estimadores de población.

*Conglomerado: La variable UPM\_DIS identifica el conglomerado o unidad primaria de muestreo. Estrato: La variable EST\_DIS identifica el estrato al que pertenece la unidad primaria de muestreo. Ponderador a nivel individual: La variable FAC\_ELE contiene el factor de expansión (ponderador) que incorpora la probabilidad de selección del individuo y ajuste por no-respuesta, omisión de delitos y proyección de la población. De acuerdo con la metodología de la encuesta, este ponderador es "requerido para estimar resultados de las preguntas*



de percepción de la seguridad pública y la victimización de la población de 18 años y #más. [8]

Se obtienen los datos de la población de acuerdo a los factores de expansión de dada unidad primaria de muestreo y del estrato, utilizando el paquete survey y la metodología descrita en el documento elaborado por el INEGI “Descripción del cálculo de los principales indicadores con R” [9]

```
design = svydesign(id=~UPM_DIS, strata=~EST_DIS,
weights=~dfm$FAC_MUJ, data=dfm)
svymean(~violenciaF, design)
```

Violencia	Media	Total
Sí	.1031	4,795,414
No	.8968	41,706,326

**Tabla 7 Estimación poblacional calculada**

Estos son los valores estimados para la población, y son diferentes a los de la muestra por lo que es necesario utilizar los factores de expansión que tienen en cuenta el modelo de muestreo.

Y corresponden a los que publica el INEGI

INEGI. Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2016 (ENDIREH). Tabulados básicos.

Estimaciones puntuales  
Observaciones muestrales  
Coeficientes de variación (%)  
Errores estándar

Distribución de mujeres de 15 años y más por condición, tipo, clase y situación de violencia en el ámbito familiar en los últimos 12 meses 2016 Cuadro 10.1

Estimaciones puntuales

Condición, tipo, clase y situación de violencia	Absolutos	Relativos
Total de mujeres de 15 años y más	46 501 740	100.0
Sin incidentes de violencia familiar	41 706 326	89.7
Con incidentes de violencia familiar	4 795 414	10.3

**Tabla 8 Tabulado básico violencia ámbito familiar**

El estimado poblacional de cada estado también se puede obtener utilizando una función de agrupación por estado

```
total.edo = svyby(~violenciaF, by=dfm$ENT, design=design,
svytotal)
total.edo
```

y obtener los siguientes resultados para las 32 entidades de la república, así como el error estándar para cada uno de ellos.

##	by	violenciaFSí	violenciaFNo	se.violenciaFSí	se.violenciaFNo
## 1	1	60427	423747	3430.969	6798.553
## 2	2	101350	1200308	7515.488	21735.792
## 3	3	19262	264204	1654.301	4322.523
## 4	4	31531	320997	2225.483	5127.041
## 5	5	128370	980598	7588.971	15121.262
## 6	6	27049	251018	1886.741	4097.293
## 7	7	140505	1717402	9849.872	26160.632
## 8	8	133590	1229592	8856.877	23265.501
## 9	9	365126	3390998	23620.947	47852.796
## 10	10	80950	571109	4877.552	9997.222
## 11	11	193470	2008590	13456.949	33879.689
## 12	12	119627	1180094	8538.961	20891.386
## 13	13	120776	997152	7935.902	18440.903
## 14	14	387769	2638412	23391.167	45490.092
## 15	15	824995	5860460	47928.805	102977.382
## 16	16	211394	1509545	12790.093	26278.704
## 17	17	76379	696285	5105.051	11984.374
## 18	18	44209	423568	2733.665	6639.695
## 19	19	133955	1813269	10595.299	29342.603
## 20	20	171047	1342669	11291.522	23836.075
## 21	21	261177	2096267	16225.494	34674.340
## 22	22	88764	681460	5068.396	12145.069
## 23	23	58559	524233	3713.304	9719.803
## 24	24	93275	940904	7240.799	14502.302
## 25	25	111489	1031570	7429.641	16695.812
## 26	26	96262	1005744	6264.676	15464.656
## 27	27	72195	822559	5305.809	12668.433
## 28	28	110065	1244495	7611.512	20123.155
## 29	29	49482	444175	3003.183	6861.208
## 30	30	350938	2808720	21580.412	45125.372
## 31	31	75497	752254	4894.115	12379.225
## 32	32	55930	533928	3496.517	8764.784

**Tabla 9 Violencia ámbito familiar por Estado**

La tabla correspondiente de proporciones

##	by	violenciaFSí	violenciaFNo	se.violenciaFSí	se.violenciaFNo
## 1	1	0.12480431	0.8751957	0.006916647	0.006916647
## 2	2	0.07786223	0.9221378	0.005713100	0.005713100
## 3	3	0.06795171	0.9320483	0.005607868	0.005607868
## 4	4	0.08944254	0.9105575	0.006291869	0.006291869
## 5	5	0.11575627	0.8842437	0.006648756	0.006648756
## 6	6	0.09727512	0.9027249	0.006515831	0.006515831
## 7	7	0.07562542	0.9243746	0.005214858	0.005214858
## 8	8	0.09799865	0.9020013	0.006258113	0.006258113
## 9	9	0.09720819	0.9027918	0.006091975	0.006091975
## 10	10	0.12414521	0.8758548	0.006975303	0.006975303
## 11	11	0.08785864	0.9121414	0.005836547	0.005836547
## 12	12	0.09204052	0.9079595	0.006321766	0.006321766
## 13	13	0.10803558	0.8919644	0.006688526	0.006688526
## 14	14	0.12813807	0.8718619	0.007463908	0.007463908
## 15	15	0.12340147	0.8765985	0.006844818	0.006844818
## 16	16	0.12283643	0.8771636	0.007099700	0.007099700
## 17	17	0.09885151	0.9011485	0.006532637	0.006532637
## 18	18	0.09450871	0.9054913	0.005697874	0.005697874
## 19	19	0.06879280	0.9312072	0.005289491	0.005289491
## 20	20	0.11299808	0.8870019	0.007079839	0.007079839
## 21	21	0.11078821	0.8892118	0.006667306	0.006667306
## 22	22	0.11524440	0.8847556	0.006385396	0.006385396
## 23	23	0.10048010	0.8995199	0.006250764	0.006250764
## 24	24	0.09019232	0.9098077	0.006724921	0.006724921
## 25	25	0.09753565	0.9024644	0.006283704	0.006283704
## 26	26	0.08735161	0.9126484	0.005538777	0.005538777
## 27	27	0.08068698	0.9193130	0.005784796	0.005784796
## 28	28	0.08125517	0.9187448	0.005433119	0.005433119
## 29	29	0.10023559	0.8997644	0.005920862	0.005920862
## 30	30	0.11106835	0.8889317	0.006683338	0.006683338
## 31	31	0.09120738	0.9087926	0.005941789	0.005941789
## 32	32	0.09481943	0.9051806	0.005584443	0.005584443

**Tabla 10 Proporción violencia ámbito familiar por entidad**





10  
11  
12  
OCTUBRE

9° Encuentro de Investigadores  
8° Encuentro de jóvenes investigadores  
2° Concurso de Tesis Doctorales



Que corresponden a los valores que reporta el INEGI

Prevalencia de la violencia en el ámbito familiar en los últimos 12 meses, entre las mujeres de 15 años y más por entidad federativa según lugar de residencia 2016

Estimaciones puntuales

Cuadro 10.2

Entidad federativa	Total	En los últimos 12 meses	
		Lugar de residencia	
		Urbano	Rural
Estados Unidos Mexicanos	10.3	10.6	9.4
Aguascalientes	12.5	12.5	12.5
Baja California	7.8	7.8	8.1
Baja California Sur	6.8	7.2	4.8
Campeche	8.9	9.3	7.8
Coahuila de Zaragoza	11.6	11.6	11.0
Colima	9.7	10.2	5.8
Chiapas	7.6	8.6	6.3
Chihuahua	9.8	10.6	4.7

Tabla 11 Tabulado básico ENDIREH INEGI

Y esto mismo se puede hacer para cada una de las variables de violencia

Violencia emocional, resultado de la degradación, intimidación, amenaza y aislamiento

```
svymean(=violenciaemocionalF, design)

##          mean      SE
## violenciaemocionalFSi 0.08135 0.0014
## violenciaemocionalFNo 0.91865 0.0014

svytotal(=violenciaemocionalF, design)

##          total      SE
## violenciaemocionalFSi 3782923 66908
## violenciaemocionalFNo 42718817 163554

total.edo = svyby(=violenciaemocionalF, by=dfmJENT, design=design, svytotal)
total.edo

## by violenciaemocionalFSi violenciaemocionalFNo se.violenciaemocionalFSi
## 1 1 49179 434995 3129.347
## 2 2 76630 1225028 6786.358
## 3 3 15991 267475 1492.076
## 4 4 22810 329718 1957.797
## 5 5 102820 1086148 6917.931
## 6 6 20877 257390 1590.196
## 7 7 183976 1753931 8376.889
```

Tabla 12 Cálculo de la variable emocional

Estimación puntual por municipio.

Los valores correspondientes a los municipios ya no se publica ni se reporta, por lo que la estimación poblacional de cada uno de ellos también puede calcularse con el factor de expansión correspondiente. En la siguiente relación está la estimación de los municipios del estado de Coahuila, —Torreón es el municipio número 35—

##	by	violenciaFSi	violenciaFNo	se.violenciaFSi	se.violenciaFNo
## 2 2	7020	47980	1585.7439	2796.235	
## 3 3	310	9919	310.0000	5526.073	
## 4 4	1140	16941	732.1049	9519.700	
## 6 6	1040	3328	1040.0000	3328.000	
## 7 7	1535	15574	1131.6877	8306.642	
## 9 9	4440	24434	2045.4295	9099.219	
## 10 10	3726	26989	1541.8110	5032.877	
## 17 17	4105	42225	1683.0766	9488.414	
## 18 18	5130	77589	1434.3180	3783.407	
## 20 20	0	8031	0.0000	5734.221	
## 21 21	0	6807	0.0000	6807.000	
## 22 22	797	4675	585.2085	3306.271	
## 23 23	1506	6784	1072.9765	4060.597	
## 24 24	2793	16958	1547.9553	7973.032	
## 25 25	5659	52902	1528.0900	3843.406	
## 26 26	0	4238	0.0000	4238.000	
## 27 27	6586	33562	2189.0461	4323.138	
## 28 28	2799	16341	416.4397	2580.424	
## 30 30	30531	279917	3964.3217	9435.216	
## 31 31	1744	10855	1744.0000	6350.359	
## 32 32	925	10062	739.2598	6708.637	
## 33 33	5870	44418	2073.5905	12577.380	
## 35 35	40236	211023	4222.6513	9102.703	
## 37 37	0	4983	0.0000	4983.000	
## 38 38	478	4063	478.0000	4063.000	

Tabla 13 Violencia ámbito familiar por municipio del Estado de Coahuila

Y la respectiva proporción

##	by	violenciaFSi	violenciaFNo	se.violenciaFSi	se.violenciaFNo
## 2	2	0.12763636	0.8723636	2.751152e-02	2.751152e-02
## 3	3	0.03030599	0.9696940	2.593905e-02	2.593905e-02
## 4	4	0.06304961	0.9369504	3.126204e-02	3.126204e-02
## 6	6	0.23809524	0.7619048	4.28613e-17	5.204170e-17
## 7	7	0.08971886	0.9102811	4.165756e-02	4.165756e-02
## 9	9	0.15377156	0.8462284	3.334585e-02	3.334585e-02
## 10	10	0.12130881	0.8786912	4.894965e-02	4.894965e-02
## 17	17	0.08860350	0.9113965	2.367958e-02	2.367958e-02
## 18	18	0.06201719	0.9379828	1.751279e-02	1.751279e-02
## 20	20	0.00000000	1.0000000	0.000000e+00	0.000000e+00
## 21	21	0.00000000	1.0000000	0.000000e+00	0.000000e+00
## 22	22	0.14565058	0.8543494	2.301969e-02	2.301969e-02
## 23	23	0.18166466	0.8183353	8.553636e-02	8.553636e-02
## 24	24	0.14141056	0.8585894	3.281304e-02	3.281304e-02
## 25	25	0.09663428	0.9033657	2.571152e-02	2.571152e-02
## 26	26	0.00000000	1.0000000	0.000000e+00	0.000000e+00
## 27	27	0.16404304	0.8359570	4.414302e-02	4.414302e-02
## 28	28	0.14623824	0.8537618	1.045954e-02	1.045954e-02
## 30	30	0.09834497	0.9016550	1.264877e-02	1.264877e-02
## 31	31	0.13842368	0.8615763	9.385138e-02	9.385138e-02
## 32	32	0.08419041	0.9158096	3.469243e-02	3.469243e-02
## 33	33	0.11672765	0.8832724	2.701790e-02	2.701790e-02
## 35	35	0.16013755	0.8398625	1.629155e-02	1.629155e-02
## 37	37	0.00000000	1.0000000	0.000000e+00	1.10223e-16
## 38	38	0.10526316	0.8947368	6.938894e-18	1.092876e-16

Tabla 14 Proporción de violencia ámbito familiar por municipio del Estado de Coahuila

Los factores de expansión son importantes porque pueden ser utilizados para el cálculo de las variables en distintos niveles, vemos en la siguiente tabla los distintos valores que se obtienen a nivel muestral y a nivel de las estimaciones poblacionales de acuerdo a las unidades primarias de muestreo.

Violencia ámbito familiar				
Mujeres de 15 años y más				
	Relativos		Absolutos	
	Sí	No	Sí	NO
Muestra				
	9.4%	90.6%	10,476	100,780
Población				



Nacional	10.3%	89.7%	4,795,414	41,706,326
Coahuila	11.6%	88.4%	128,370	980,598
Torreón	16.0%	84.0%	40,236	211,023

**Tabla 15 Valores muestrales y poblacionales**

#### 4. Conclusiones

Los datos abiertos que proporciona el INEGI así como lo que provienen de los distintos órganos e instituciones gubernamentales son una fuente de información invaluable y que responde a las prioridades nacionales y a cuantiosos presupuestos y recursos destinados para los distintos proyectos y que pueden ser aprovechados y utilizados en necesidades más localizadas y acotadas.

El estudio de los reportes que proporciona el INEGI, así como los metadatos, estructuras de las bases de datos y de las metodologías utilizadas son fuente de información que puede ser utilizada por los académicos para sus investigaciones y actividades en su propuesta docente y de difusión.

Las respuestas a los distintos reactivos de las encuestas generan nuevas variables, las respuestas aisladas en sí mismas no completan un diagnóstico cualitativo y de igual forma el cruce de las variables no corresponde a una intuición de independencia, sino que debe ser analizada con los análisis estadísticos correspondientes.

Las estimaciones poblacionales no corresponden a las muestrales por mucho que se cuente con un tamaño de muestra considerable si los individuos o registros no corresponden a una muestra representativa y no se consideran los estratos o conglomerados que tengan en cuenta su distribución proporcional, por lo que es importante recuperar las unidades primarias de muestreo para poder ponderar los cálculos que les corresponden.

Las distintas agrupaciones geográficas pueden ser utilizadas para generar información distinta como la de zonas conurbadas incluso interestatales, o de zonas especiales por alguna característica natural, económica o política o bien zonas particulares dentro de las ciudades. El ambiente del lenguaje R proporciona distintos paquetes que auxilian las operaciones de obtención, limpieza, selección, cálculo, análisis estadístico, así como de representación

y de igual manera los programas pueden ser utilizados para poder realizar investigación reproducible para otros lugares o poblaciones. La Curaduría de los datos es una operación necesaria cuando la complejidad y el volumen de los mismos hace viable su utilización en otros contextos para los que fueron creados y utilizados.

#### Referencias

- [1] CRAN. *The Comprehensive R Archive Network*. Obtenido de <http://cran.r-project.org/src/contrib/PACKAGES.html>
- [2] Secretaría de la función pública *Decreto por el que se establece la regulación en materia de Datos Abiertos (23 de Julio de 2015)* Recuperado el 26 de 07 de 2018, de [https://www.gob.mx/cms/uploads/attachment\\_data/file/9139/decreto-datos-abiertos-151214-ultima-version-cjef.pdf](https://www.gob.mx/cms/uploads/attachment_data/file/9139/decreto-datos-abiertos-151214-ultima-version-cjef.pdf)
- [3] Foundation, F. S. (2004). *The R Project for Statistical Computing*. Obtenido de The R Project for Statistical Computing: <https://www.r-project.org/>
- [4] Instituto Nacional de Estadística y Geografía (INEGI). (2010). *Censo de Población y Vivienda 2010*. Obtenido de Censo de Población y Vivienda 2010: <http://www.beta.inegi.org.mx/proyectos/ccpv/2010/>
- [5] Instituto Nacional de Estadística y Geografía, (INEGI), (26 de 07 de 2018). *Marco Geoestadístico*. Obtenido de Marco Geoestadístico: <http://www.beta.inegi.org.mx/temas/mapas/mg/>
- [6] Instituto Nacional de Estadística y Geografía (INEGI). (26 de 07 de 2018). *Espacio y datos de México, unificado, mapa, mapas, INEGI*. Obtenido de Espacio y datos de México, unificado, mapa, mapas, INEGI: <http://www.beta.inegi.org.mx/app/mapa/espaciodydatos/>



- [7] Instituto Nacional de Estadística y Geografía (INEGI). *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2016*. Obtenido de *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2016*. Recuperado el 26 de 07 de 2018, <http://www.beta.inegi.org.mx/proyectos/enchogares/especiales/endireh/2016/>
  
- [8] Instituto Nacional de Estadística y Geografía (INEGI). (2016). *Marco conceptual Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2016 ENDIREH*. Cd. de México. Recuperado el 26 de 07 de 2018, de [http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod\\_serv/contenidos/espanol/bvinegi/productos/nueva\\_estruc/7028250950555.pdf](http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/7028250950555.pdf)
  
- [9] Márquez, J. (26 de 07 de 2018). *Cómo analizar la ENVIPE en R*. Obtenido de Políticamente Correcto Elecciones y Estadística. Recuperado el 26 de 07 de 2018, <http://javiermarquez.org/2015/10/12/como-analizar-la-envipe-en-r/>
  
- [10] Instituto Nacional de Estadística y Geografía INEGI. (2016). *Descripción del cálculo de los principales indicadores con R. Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública*. Recuperado el 26 de 07 de 2018, de [http://www.beta.inegi.org.mx/contenidos/proyectos/enchogares/regulares/envipe/2016/doc/calculo\\_indicadores\\_r\\_envipe2016.pdf](http://www.beta.inegi.org.mx/contenidos/proyectos/enchogares/regulares/envipe/2016/doc/calculo_indicadores_r_envipe2016.pdf)