

METODOLOGÍA Percepciones  
Sociales del INEGI 2017  
Encuesta Nacional en  
Unidades Económicas



# Percepciones Sociales del INEGI 2017

## Índice

DISEÑO MUESTRAL.....	2
Población objetivo.....	2
Marco de la encuesta.....	3
Esquema de muestreo .....	4
Tamaño de muestra .....	5
Afijación de la muestra.....	6
Cálculo de los factores de expansión .....	8
Ajuste de los factores de expansión.....	9
Ajuste por no respuesta .....	9
Cálculo de los estimadores.....	10
1. Estimaciones puntuales.....	10
2. Medidas de dispersión .....	11
ANEXO 1 .....	14

## DISEÑO MUESTRAL

### Percepciones Sociales del INEGI 2017

### Encuesta Nacional en Unidades Económicas

La encuesta nacional en unidades económicas tiene como objetivo obtener información acerca de los niveles de conocimiento, utilización y confianza que tienen las unidades económicas del país, conformado por las micro, pequeñas, medianas y grandes empresas, así como las instituciones de servicios financieros y de seguros y de servicios educativos, sobre la información que recopila y produce el **Instituto Nacional de Estadística y Geografía (INEGI)**. Para ello, se diseñó una encuesta que permite captar información a nivel nacional.

#### Población objetivo

Para esta encuesta, la población objetivo la constituyen las unidades económicas del país, conformadas por micro, pequeñas, medianas y grandes empresas que pertenecen a los siguientes sectores económicos, según el Sistema de Clasificación Industrial de América del Norte 2013 (SCIAN 2013):

- 1) Industrial (minería, electricidad y agua, construcción y manufacturas).
- 2) Comercio.
- 3) Algunos sectores de servicios clasificados por el SCIAN 2013.

#### Dominios de estudio

Las Unidades Económicas que conforman la población objetivo están conformadas por los siguientes dominios de estudio, y para cada uno de ellos se pueden obtener estimaciones a nivel nacional:

- 4) Micro, pequeñas y medianas empresas, en conjunto (MIPYMES).
- 5) Grandes empresas.
- 6) Instituciones de servicios financieros y de seguros.
- 7) Instituciones de servicios educativos.

## Marco de la encuesta

Como marco de muestreo se utilizaron las unidades económicas con tipo de asentamiento fijo que integran el Directorio Estadístico Nacional de Unidades Económicas (DENUE) del INEGI, actualizado a Octubre del 2016, del cual se excluyeron a aquellas unidades económicas cuya clasificación económica según el SCIAN 2013 fuera:

- Agricultura, crías y explotación de animales, aprovechamiento forestal, pesca y caza (Sector 11).
- Agencias noticiosas (Clase 519110).
- Otros servicios de suministro de información (Clase 519190).
- Actividades legislativas, gubernamentales, de impartición de justicia y de organismos internacionales y extraterritoriales (Sector 93).

El marco de la encuesta quedó integrado por 4,276,132 unidades, distribuidas por dominio de estudio de la siguiente manera:

Dominio de estudio	Tamaño de la población
MIPYMES	4,075,635
Grandes empresas	18,209
Servicios financieros y de seguros	44,147
Servicios educativos	138,141
<b>Total</b>	<b>4,276,132</b>

### Formación de las Unidades primarias de muestreo (UPM)

Para los dominios de MIPYMES, Instituciones de Servicios Educativos e Instituciones de Servicios Financieros y de Seguros, se construyeron Unidades Primarias de Muestreo (UPM), las cuales estuvieron constituidas por aquellas AGEB1 que tienen la característica de contener al menos 7 unidades económicas del mismo dominio de estudio (14 para el

---

<sup>1</sup>Área Geoestadística Básica: extensión territorial delimitada por el INEGI cuyos habitantes comparten características socioeconómicas parecidas.

caso de MIPYMES). Dicha selección fue realizada con probabilidad proporcional al personal ocupado dentro de las unidades económicas de la AGEB en el dominio de estudio correspondiente<sup>2</sup> (véase Anexo I).

Las UPM fueron utilizadas como muestra maestra para, en una segunda etapa, seleccionar un número determinado de AGEB, de las cuales se seleccionaron las Unidades Económicas pertenecientes a los dominios mencionados.

## Estratificación

Para los dominios de MIPYMES, Instituciones de Servicios Educativos e Instituciones de Servicios Financieros y de Seguros, se agruparon todas las UPM que formaron la muestra maestra, por entidad federativa y para cada dominio de estudio.

Para el dominio de Grandes Empresas, cada entidad federativa formó de manera natural un estrato geográfico.

## Esquema de muestreo

La selección de las unidades económicas para los dominios de MIPYMES, Servicios financieros y de seguros y de Servicios educativos se realizó bajo un diseño probabilístico, polietápico, estratificado y por conglomerados, donde cada unidad de muestreo tuvo una probabilidad conocida y diferente de cero de ser seleccionada. Por otra parte, la selección de grandes empresas llevó un diseño probabilístico de una sola etapa y estratificado, donde cada unidad tuvo una probabilidad conocida y diferente de cero de ser seleccionada.

## Selección de la muestra

Para los dominios de MIPYMES, Instituciones de Servicios Educativos e Instituciones de Servicios Financieros y de Seguros; la selección de unidades se llevó a cabo en dos etapas de manera independiente para cada dominio.

---

<sup>2</sup> La formación de UPM fue elaborada por el INEGI a petición del Departamento de Investigación Aplicada y Opinión, IJ-UNAM. Se presenta la Nota técnica *Selección de UPM* como Anexo 1.

- a) *Unidades Primarias de Muestreo (UPM)*: De las  $kh$  AGEBS seleccionadas de la muestra maestra en la  $h$ -ésima entidad, se seleccionaron  $kh^*$  de manera aleatoria (MAS).
- b) *Unidades Últimas de Muestreo (UUM)*: Se seleccionaron 7 unidades económicas para cada AGEB de la etapa anterior, de manera aleatoria (MAS).

Para el dominio de grandes empresas, la selección de unidades se realizó de manera aleatoria e independiente para cada estrato con un diseño de muestreo aleatorio simple.

## Tamaño de muestra

Para calcular el tamaño de la muestra de cada dominio de estudio se utilizó la siguiente fórmula:

$$n_D = \frac{N_D P_D (1 - P_D) z_{\alpha/2}^2}{N_D r^2 P_D^2 + P_D (1 - P_D) z_{\alpha/2}^2} \frac{1}{(1 - TNR)}$$

donde:

$n_D$  = Tamaño de muestra para el dominio  $D$ .

$N_D$  = Total de unidades en el dominio  $D$ .

$P_D$  = Proporción supuesta en el dominio  $D$ .

$z_{\alpha/2}^2$  = valor en tablas de valores de probabilidad acumulada para la distribución normal estándar. Este valor depende del nivel de confianza asignado  $(1-\alpha)*100$  por ciento.

$r$  = Error relativo.

$TNR$  = Tasa de no respuesta esperada.

El tamaño de la muestra se obtuvo para estimar una proporción considerando un error relativo del 10%, un nivel de confianza del 95%, una tasa de no respuesta esperada del 35% y una proporción diferenciada por dominio de estudio<sup>3</sup>:

---

<sup>3</sup> Las proporciones para los dominios de estudio se consideraron con base en los resultados de la EMECOI 2014, para la variable de *uso de la información proporcionada por el INEGI*.

Dominio de estudio	Proporción
MIPYMES	14%
Grandes Empresas	28%
Servicios Financieros y de Seguros	29%
Servicios Educativos	14%

El tamaño de muestra obtenido fue de 10,040 unidades económicas, distribuido por dominio de estudio de la siguiente manera:

Dominio de estudio	Tamaño de la muestra
MIPYMES	3,633
Grandes empresas	1,430
Servicios financieros y de seguros	1,407
Servicios educativos	3,570
<b>Total</b>	<b>10,040</b>

## Afijación de la muestra

La distribución de la muestra se realizó de forma proporcional al número de unidades económicas dentro de cada estrato para cada dominio, como se muestra en la siguiente expresión:

$$n_i^D = \frac{N_i^D}{N^D} n^D$$

donde:

$n_i^D$  es el número de unidades económicas seleccionadas en la muestra en la  $i$ -ésima entidad, del dominio D.

$n^D$  es el total de unidades económicas en la muestra pertenecientes al dominio D.

$N_i^D$  es el total de unidades económicas en la  $i$ -ésima entidad, del dominio D.

$N^D$  es el total de unidades económicas pertenecientes al Dominio D.

El tamaño de la muestra para cada estrato y para cada dominio de estudio quedó distribuido de la siguiente manera:

Entidad	Grandes Empresas	Servicios Educativos	MIPYMES	Servicios Financieros y de Seguros
Aguascalientes	20	42	49	14
Baja California	79	98	119	56
Baja California Sur	12	21	21	14
Campeche	11	35	28	14
Coahuila de Zaragoza	69	70	105	42
Colima	7	21	28	14
Chiapas	13	98	112	56
Chihuahua	74	63	112	42
Ciudad de México	226	448	364	112
Durango	21	35	56	14
Guanajuato	71	161	175	77
Guerrero	15	77	84	42
Hidalgo	19	63	63	21
Jalisco	96	294	294	112
México	137	553	385	84
Michoacán de Ocampo	19	112	154	63
Morelos	15	77	56	28
Nayarit	9	35	35	14
Nuevo León	125	168	189	91
Oaxaca	11	84	112	42
Puebla	41	210	182	63
Querétaro	44	63	77	35
Quintana Roo	31	35	49	21
San Luis Potosí	31	91	84	28
Sinaloa	30	70	84	49
Sonora	45	56	98	49
Tabasco	16	63	49	28
Tamaulipas	57	70	112	49
Tlaxcala	10	42	42	14
Veracruz de Ignacio de la Llave	47	231	196	70
Yucatán	19	63	77	35
Zacatecas	10	21	42	14
<b>Total</b>	<b>1,430</b>	<b>3,570</b>	<b>3,633</b>	<b>1,407</b>

## Cálculo de los factores de expansión

Para obtener los factores de expansión utilizados en el cálculo de los estimadores de los parámetros en las unidades económicas de cada dominio de estudio, se utilizó la probabilidad de selección de cada etapa de muestreo.

Para los dominios de MIPYMES, Instituciones de Servicios Educativos e Instituciones de Servicios Financieros y de Seguros; la probabilidad de selección de la  $k$ -ésima unidad económica, de la  $j$ -ésima AGEB, de la  $i$ -ésima entidad es:

$$P[x_{i,j,k}] = \frac{n^D P_j^D}{P^D} \frac{n_i^D n_i^{D*}}{n_i^D} \frac{7}{m_{i,j}^D} = \frac{7P_j^D n_i^{D*}}{P^D m_{i,j}^D}$$

donde:

$x_{i,j,k}$  es la  $k$ -ésima unidad económica, en la  $j$ -ésima AGEB, en la  $i$ -ésima entidad.

$n^D$  es el número de AGEB que conforman la muestra maestra para el dominio  $D$ .

$P_j^D$  es la población ocupada del dominio  $D$  en la  $j$ -ésima AGEB.

$P^D$  es la población ocupada total del dominio  $D$ .

$n_i^D$  es el número de AGEB que conforman la muestra maestra en la  $i$ -ésima entidad, para el dominio  $D$ .

$n_i^{D*}$  es el número de AGEB seleccionadas aleatoriamente, a partir de las primeras  $n_i^D$ , en la  $i$ -ésima entidad, para el dominio  $D$ .

$m_{i,j}^D$  es el número de unidades económicas, en la  $j$ -ésima AGEB, en la  $i$ -ésima entidad para el dominio  $D$ .

El recíproco de la probabilidad de selección de la  $k$ -ésima unidad económica es el factor de expansión correspondiente, es decir:

$$F_{i,j,k} = \frac{1}{P[x_{i,j,k}]} = \frac{P^D m_{i,j}^D}{7P_j^D n_i^{D*}}$$

Para el dominio de grandes empresas, la probabilidad de selección de la  $k$ -ésima unidad económica, de la  $i$ -ésima entidad es:

$$P[x_{i,k}] = \frac{n_i}{N_i}$$

donde:

$x_{i,k}$  es la  $k$ -ésima unidad económica, en la  $i$ -ésima entidad.

$n_i$  es el número de unidades económicas seleccionadas en la  $i$ -ésima entidad.

$N_i$  es el total de unidades económicas en la  $i$ -ésima entidad.

Para el dominio de estudio de grandes empresas, el recíproco de la probabilidad de selección de la  $k$ -ésima unidad económica es el factor de expansión correspondiente, es decir:

$$F_{i,k} = \frac{1}{P[x_{i,k}]} = \frac{N_i}{n_i}$$

## Ajuste de los factores de expansión

Los factores de expansión descritos anteriormente para cada dominio de estudio fueron calibrados con el fin de expandir a la población objetivo en las proporciones que se presentan en el país. El procedimiento realizado fue el siguiente:

### Ajuste por no respuesta

El ajuste por no respuesta es un método de calibración que se utiliza cuando no se realizan todas las entrevistas esperadas, y se calculó de la siguiente forma para cada dominio de estudio:

$$F'_i = F_i \frac{x'_i}{x''_i}$$

donde:

$F'_i$  es el factor de expansión corregido por no respuesta en la  $i$ -ésima entidad.

$F_i$  es el factor de expansión natural  $i$ -ésima entidad.

$x'_i$  es el total de unidades económicas seleccionadas en la  $i$ -ésima entidad.

$x''_i$  es el total de unidades económicas seleccionadas con respuesta en la  $i$ -ésima entidad.

## Cálculo de los estimadores

Para el cálculo de las estimaciones puntuales y sus respectivas medidas de dispersión, se utilizaron para cada dominio de estudio las siguientes fórmulas.

### 1. Estimaciones puntuales

Para los dominios de MIPYMES, Instituciones de Servicios Educativos e Instituciones de Servicios Financieros y de Seguros; el cálculo de proporciones, promedios y tasas se hace con la expresión de estimador de razón siguiente:

$$\hat{R} = \frac{\sum_{i=1}^{32} \sum_{j=1}^J \sum_{k=1}^7 X_{ijk} F''_{ijk}}{\sum_{i=1}^{32} \sum_{j=1}^J \sum_{k=1}^7 Y_{ijk} F''_{ijk}}$$

Para el dominio de estudio de Grandes Empresas, el cálculo del estimador de razón se hizo de la siguiente manera:

$$\hat{R} = \frac{\sum_{i=1}^{32} \sum_{k=1}^K X_{ik} F''_{ik}}{\sum_{i=1}^{32} \sum_{k=1}^K Y_{ik} F''_{ik}}$$

donde:

$X_{ijk}$  es el valor observado de la característica de interés  $X$  en la  $k$ -ésima unidad económica, de la  $j$ -ésima UPM, en la  $i$ -ésima entidad.

$Y_{ijk}$  es el valor observado de la característica de interés  $Y$  en la  $k$ -ésima unidad económica, de la  $j$ -ésima UPM, en la  $i$ -ésima entidad.

$F''_{ijk}$  es el factor de expansión de la k-ésima unidad económica, de la j-ésima UPM, en la i-ésima entidad.

$X_{ik}$  es el valor observado de la característica de interés  $X$  en la k-ésima unidad económica, en la i-ésima entidad.

$Y_{ik}$  es el valor observado de la característica de interés  $Y$  en la k-ésima unidad económica, en la i-ésima entidad.

$F''_{ik}$  es el factor de expansión de la k-ésima unidad económica, en la i-ésima entidad.

## 2. Medidas de dispersión

Para el estimador de la Varianza en los dominios de estudio MIPYMES, Instituciones de Servicios Educativos e Instituciones de Servicios Financieros y de Seguros; se utilizó el método de Conglomerados Últimos<sup>4</sup>, al que se le aplicó el método de series de Taylor, obteniéndose la siguiente fórmula para estimar la precisión del estimador de razón  $\hat{R}$ .

$$\hat{V}(\hat{R}) = \frac{1}{\hat{Y}^2} \sum_{h=1}^L \left\{ \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[ \left( \hat{X}_{hi} - \frac{\hat{X}_h}{n_h} \right) - \hat{R} \left( \hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right) \right]^2 \right\}$$

donde:

$\hat{X}_{hi}$  es el valor total observado y ponderado de la característica de interés  $X$  en la i-ésima UPM, en la h-ésima entidad.

$\hat{Y}_{hi}$  es el valor total observado y ponderado de la característica de interés  $Y$  en la i-ésima UPM, en la h-ésima entidad.

$\hat{X}_h$  es el valor total observado y ponderado de la característica de interés  $X$  en la h-ésima entidad.

$\hat{Y}_h$  es el valor total observado y ponderado de la característica de interés  $Y$  en la h-ésima entidad.

---

<sup>4</sup> Hansen, M.H. Horwitz, W.N. y Madow, W.G, Sample Survey Methods and Theory, (1953), Vol. 1 página 242.

$\hat{Y}$  es el valor total observado y ponderado de la característica de interés Y a nivel nacional.

$n_h$  es el número de UPM en la h-ésima entidad.

Para el dominio de Grandes Empresas, el estimador de la varianza para el estimador de razón se hizo de la siguiente forma:

$$\hat{V}(\hat{R}) = \sum_{h=1}^L \frac{N_h^2}{N^2} \left(1 - \frac{n_h}{N_h}\right) \frac{1}{n_h} \frac{1}{\hat{Y}_h^2} \sum_{i=1}^{n_h} \frac{(\hat{X}_{hi} - \hat{R}_h \hat{Y}_{hi})^2}{N_h - 1}$$

donde:

$N_h$  es el total de unidades económicas en el h-ésimo estrato.

$N$  es el total de unidades económicas en el país.

$n_h$  es el total de unidades económicas muestreadas en el h-ésimo estrato.

$\hat{Y}_h$  es el promedio estimado de la característica de interés Y en el h-ésimo estrato.

$\hat{X}_{hi}$  es el total estimado de la característica de interés X.

$\hat{Y}_{hi}$  es el valor observado de la característica de interés Y de la i-ésima unidad económicas en el h-ésimo estrato.

La estimación de la desviación estándar (D.E.), intervalo de confianza al 100(1- $\alpha$ )% (I.C.), el coeficiente de variación (C.V.) y efecto de diseño (DEFF) se hace a través de las siguientes expresiones:

$$D.E. = \sqrt{\hat{V}(\hat{R})} \qquad C.V. = \frac{\sqrt{\hat{V}(\hat{R})}}{\hat{R}} \qquad DEFF = \frac{\hat{V}(\hat{R})}{\hat{V}(\hat{R})_{MAS}}$$

$$I.C._{1-\alpha} = \hat{R} \pm Z_{\alpha/2} \sqrt{\hat{V}(\hat{R})}$$

donde:

$\hat{R}$  es el estimador de razón.

$\hat{V}(\hat{R})$  es la varianza estimada del estimador de razón.

$\hat{V}(\hat{R})_{MAS}$  es la varianza estimada del estimador de razón.

$Z_{\alpha/2}$  es el valor en tablas de valores de probabilidad acumulada para la distribución normal estándar. Este valor depende del nivel de confianza asignado  $(1-\alpha)*100$  por ciento.

## ANEXO 1

### Nota técnica sobre la selección de UPM



#### INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA Estudio Percepciones sociales del INEGI

#### Nota Técnica sobre Selección de UPM

#### Antecedentes

Para llevar a efecto el Estudio sobre Percepciones Sociales del INEGI enfocado en unidades económicas, el Departamento de Investigación Aplicada y Opinión del Instituto de Investigaciones Jurídicas de la UNAM solicitó para cada dominio de estudio, la formación y selección de unidades primarias de muestreo (UPM) que contaran con un determinado número de unidades económicas como mínimo.

Se convino con la UNAM que la UPM a considerar fuera el AGEB, el detalle de la solicitud se especifica en el cuadro siguiente; el cual se complementa con el detalle de las UPM finalmente seleccionadas, el total de unidades económicas que éstas incluyen y la participación porcentual alcanzada para cada dominio.

Dominio de estudio	Solicitud		UPM en marco que cubren mínimo solicitado	UPM Seleccionadas		
	UPM (AGEB)	Mínimo de unidades económicas por UPM		Total UPM	Total Unidades económicas	Participación porcentual alcanzada en unidades económicas
<b>Total</b>	<b>11,474</b>		<b>42,230</b>	<b>17,924</b>	<b>1,126,724</b>	
Servicios Financieros y de Seguros	1,632	7	1,835	1,632	23,089	52.3
Grandes empresas*	1,610	7	438	8,060	18,209	100
MIPyMES	4,152	14	34,975	4,152	1,039,287	25.5
Servicios Educativos	4,080	7	5,420	4,080	46,139	33.4

En el dominio correspondiente a Grandes empresas, se determinó entregar todo el directorio debido a que no fue posible cubrir el total de UPM solicitadas según el requerimiento.

### **Método de selección de las UPM**

La selección de las UPM en cada dominio se realizó en forma proporcional al tamaño según la variable de personal ocupado. Para ello se realizó lo siguiente:

- a) Se identificaron e integraron las UPM (AGEB) que cubrieran con el mínimo de unidades económicas solicitadas.
- b) Para cada UPM se obtuvo el total de personal ocupado y se calculó la probabilidad de selección para esta variable.
- c) Se realizó la selección para cada UPM aplicando el método del tamaño acumulativo sin reemplazo.

El método de tamaño acumulativo consiste en acumular los valores de la variable de interés, crear intervalos de selección a partir de esta distribución acumulada y después generar  $n$  números aleatorios entre uno y el valor máximo de la distribución y finalmente seleccionar los elementos en cuyos intervalos de selección se ubiquen los números aleatorios generados.