

# ANEXO

## METODOLOGÍA

### **Población objetivo**

Población de 12 años o más que habita en viviendas particulares habitadas durante el periodo de levantamiento, dentro del territorio mexicano.

### **Tamaño de muestra nacional**

El diseño empleado para calcular el número de entrevistados para este proyecto se realizó a partir de la fórmula muestral para poblaciones infinitas. La fórmula de muestreo para calcular el número de población a encuestar toma en cuenta tanto la tasa de no respuesta como el efecto de diseño.

Si se considera un nivel de confianza del 95%, un error de estimación no mayor a 5%, probabilidades  $p=0.5$  y  $q=0.5$ , un efecto de diseño de 1.3 se calcula un total de al menos 4 mil 039 entrevistas. Sin embargo, al contemplar una tasa de no respuesta esperada de 15%, y así evitar sustituciones, el tamaño de muestra se incrementa a 4,752 contactos. De ésta forma, la muestra tendrá representación a nivel nacional, por estratos poblaciones o regiones del país.

## Diseño de la muestra nacional

Para conformar el marco de muestreo de la Encuesta Nacional de Lectura y Escritura 2015 se utilizó la información disponible en el Sistema para la Consulta de Información Censal correspondiente al año 2010 generada por el Instituto Nacional de Estadística y Geografía (INEGI), utilizando el listado de los municipios y la información de la población total y de la población de 12 y más años, por entidad federativa y municipio.

Para la realización del análisis de las prácticas y hábitos de lectura de los mexicanos de 12 años y más se diseñó una muestra representativa. El diseño sigue un procedimiento azaroso, estratificado, polietápico y por conglomerados, con selecciones distintas en sus unidades de muestreo (probabilidades proporcionales al tamaño de la población y selección aleatoria sistemática), donde la última unidad de selección es el individuo de 12 años y más.

Para mantener una consistencia con los resultados de la Encuesta Nacional de Lectura 2006, se consideraron estratos basados en seis Regiones geográficas y 6 estratos de municipios por tamaño de población, como se muestra a continuación:

ESTRATIFICACIÓN POR REGIÓN	
	CATEGORÍAS
Región	Sur
	Noroeste
	Centro
	Centro Occidente
	Noreste
	Distrito Federal

ESTRATIFICACIÓN POR TAMAÑO DE POBLACIÓN EN EL MUNICIPIO	
	CATEGORÍAS
Municipios por estrato de población	2,500 hab. o menos
	2,501 a 15,000 hab.
	15,001 a 50,000 hab.
	50,0001 a 100,000 hab.
	100,001 a 500,000 hab.
	Más de 500,000 hab.

El motivo fundamental por la cual se realiza la estratificación estriba en dos razones: La primera, porque existe una relación entre el tamaño de la población y los recursos y servicios; y, segundo, para tener una muestra más eficiente.

Para llegar al informante se consideraron cinco etapas de selección en la muestra nacional:

### Primera Etapa.

La Unidad Primaria de Muestreo (UPM) es el municipio. Se realiza la selección de municipios con probabilidad proporcional a sus tamaños (PPT) de población de 12 años o más. Es decir, el municipio que tiene el doble de tamaño tiene el doble de probabilidades de ser seleccionado. Para hacer la selección sistemática por PPT se hizo una estratificación implícita por el grado de marginación que reporta CONAPO. La distribución de municipios en muestra en la siguiente tabla:

Estratos de municipios por tamaño de población	Regiones						Total
	Sur	Noroeste	Centro	Centro-occidente	Noreste	DF	
2,500 hab. o menos	6	7	6	7	7		33
2,501 a 15,000 hab.	7	6	7	6	7		33
15,001 a 50,000 hab.	7	7	6	7	6		33
50,001 a 100,000 hab.	6	6	7	7	7		33
100,001 a 500,000 hab.	5	6	5	5	5	7	33
Más de 500,000 hab.	6	5	6	5	5	6	33
<b>Total</b>	<b>37</b>	<b>37</b>	<b>37</b>	<b>37</b>	<b>37</b>	<b>13</b>	<b>198</b>

### **Segunda etapa**

La unidad secundaria de muestreo es el área geoestadística básica, AGEB. Al interior de cada uno de los municipios seleccionado se realiza la selección de dos AGEB con probabilidad proporcional a su tamaño (PPT) de población de 12 años o más.

### **Tercera etapa**

La unidad terciaria de muestreo son las manzanas de las AGEB, una vez seleccionadas aleatoriamente dos manzanas con igual probabilidad.

### **Cuarta etapa**

La Unidad Cuaternaria de muestreo fueron las viviendas de las manzanas en las AGEB. Una vez seleccionada una manzana se seleccionan seis viviendas de manera sistemática.

### **Quinta etapa**

Una vez seleccionada una vivienda, se seleccionó al azar a un individuo de 12 años o más, residente habitual de la vivienda, para contestar el cuestionario individual. La selección del individuo se hizo con base al cumpleaños más próximo al momento de la realización de la encuesta y que se identificó en el cuestionario de hogar.

De esa manera, los contactos a realizar por región y estrato quedaron distribuidos de la siguiente manera:

Estratos de municipios por tamaño de población	Regiones						Total
	Sur	Noroeste	Centro	Centro- occidente	Noreste	DF	
<b>2,500 hab. o menos</b>	144	168	144	168	168		<b>792</b>
<b>2,501 a 15,000 hab.</b>	168	144	168	144	168		<b>792</b>
<b>15,001 a 50,000 hab.</b>	168	168	144	168	144		<b>792</b>
<b>50,0001 a 100,000 hab.</b>	144	144	168	168	168		<b>792</b>
<b>100,001 a 500,000 hab.</b>	120	144	120	120	120	168	<b>792</b>
<b>Más de 500,000 hab.</b>	144	120	144	120	120	144	<b>792</b>
<b>Total</b>	<b>888</b>	<b>888</b>	<b>888</b>	<b>888</b>	<b>888</b>	<b>312</b>	<b>198</b>

## Diseño de muestra para zonas metropolitanas

Para mantener consistencia con los resultados generados, se trabajó una segunda muestra, independiente, con las zonas metropolitanas de la Ciudad de México, Guadalajara y Monterrey. Según el INEGI, CONAPO y SEDESOL, la zona metropolitana se define como “el conjunto de dos o más municipios donde se localiza una ciudad de 50,000 o más habitantes, cuya área urbana, funciones y actividades rebasan el límite del municipio que originalmente la contenía, incorporando como parte de sí misma o de su área de influencia directa a municipios vecinos, predominantemente urbanos, con los que mantiene un alto grado de integración socioeconómica; en ésta definición se incluye además a aquellos municipios que por sus características particulares son relevantes para la planeación y política urbanas”<sup>1</sup>. El estudio de estos espacios poblacionales resulta de importancia ya que en éstos hay un desarrollo económico, cultural y social similar, que supondría a su vez una similitud en la infraestructura relacionada con la promoción a la lectura.

En la encuesta de 2006 se realizaron 1,080 cuestionarios adicionales a la población de 12 años y más que residen en las zonas metropolitanas de Ciudad de México, Guadalajara y Monterrey. Para la Encuesta Nacional de Lectura y Escritura 2015 se realizaron al menos 1,200 entrevistas efectivas. Considerando una tasa de no respuesta en del 15%, al igual que en la muestra nacional, se planteó la visita a 1,410 viviendas.

---

<sup>1</sup> INEGI (2005) Delimitación de las Zonas Metropolitanas de México 2005. INEGI: Aguascalientes.



Para la muestra en Zonas Metropolitanas, se recurrió a la estratificación de municipios utilizada en la muestra nacional, con la variante de que los municipios que conforman una zona metropolitana sólo incluyen dos estratos de poblaciones: de 101,000 a 500,000 habitantes y más de 500,000 habitantes. La selección sigue el siguiente procedimiento.

**Primera etapa.** La Unidad Primaria de Muestreo (UPM) es el AGEB. Se realiza la selección de las AGEB con probabilidades proporcionales a su tamaño (PPT) de población de 12 años o más, seleccionando un total de 141 AGEB, de acuerdo a la distribución que se muestra en la tabla:

Estratos de municipios por tamaño de población	ZM Ciudad de México				Total
	Distrito Federal	Estado de México	ZM Monterrey	ZM Guadalajara	
100,001 a 500,000 hab.	13	10	23	23	69
Más de 500,000 hab.	11	13	24	24	72
<b>Total</b>	<b>24</b>	<b>23</b>	<b>47</b>	<b>47</b>	<b>141</b>

**Segunda etapa.** La Unidad Secundaria de muestreo fueron las manzanas dentro de cada AGEB. Una vez seleccionado un AGEB se seleccionarán dos manzanas.

**Tercera etapa.** La Unidad Terciaria de Muestreo serán las viviendas de las manzanas en las AGEB. Una vez seleccionada una manzana se seleccionaron cinco viviendas.

**Cuarta etapa.** Una vez seleccionada la vivienda, el informante fue el individuo mayor de 12 años con el cumpleaños más próximo. La distribución de entrevistas quedó conformada de acuerdo a la tabla:

Estratos de municipios por tamaño de población	ZM Ciudad de México		ZM Monterrey	ZM Guadalajara	Total
	Distrito Federal	Estado de México			
100,001 a 500,000 hab.	130	100	230	230	690
Más de 500,000 hab.	110	130	240	240	720
<b>Total</b>	<b>240</b>	<b>230</b>	<b>470</b>	<b>470</b>	<b>1410</b>

### Diseño total de la muestra

Al sumar la muestra nacional y la muestra en Zonas Metropolitanas, se contempló un total de 6,162 contactos realizados en el levantamiento. La tabla siguiente presenta la distribución total de contactos de acuerdo a la región y el estrato de tamaño de Municipio.

Estratos de municipios por tamaño de población	Regiones						Total
	Sur	Noroeste	Centro	Centro-occidente	Noreste	DF	
2,500 hab. o menos	144	168	144	168	168		792
2,501 a 15,000 hab.	168	144	168	144	168		792
15,001 a 50,000 hab.	168	168	144	168	144		792
50,0001 a 100,000 hab.	144	144	168	168	168		792
100,001 a 500,000 hab.	120	144	220	350	350	298	1,482
Más de 500,000 hab.	144	120	274	360	360	254	1,512
<b>Total</b>	<b>888</b>	<b>888</b>	<b>1,118</b>	<b>1,358</b>	<b>1,358</b>	<b>552</b>	<b>6,162</b>

## Proceso de estimación

Dado que el estudio utilizó una selección de naturaleza azarosa, el procesamiento se realiza asumiendo un Muestreo Aleatorio Simple (MAS). De esa manera, las estimaciones presentadas no están ponderadas por ningún factor de expansión, debido a que el procedimiento de selección utilizado imposibilita el cálculo de las probabilidades de selección y, por ende, de dichos ponderadores. Sin embargo, considerando que la selección se hizo con un esquema probabilístico en las primeras etapas de muestreo, que se tiene un tamaño de muestra amplio y que se distribuyeron las entrevistas por todo el territorio nacional, tocando distintas regiones y poblaciones de diverso tamaño, tanto del ámbito urbano como el rural, se asume que las estimaciones globales brindan una buena aproximación de los parámetros poblacionales, por lo que los resultados del presente reporte muestran estimaciones a nivel nacional. Para todos los porcentajes estimados se utiliza la base de casos válidos, por lo que se excluyen aquellos casos con No Respuesta o en los que en el informante no haya contestado. Por este motivo, el número de casos reportado en algunas tablas varía del total de observaciones de la encuesta.

De igual manera, para el cálculo de los errores estándar de los estimadores y sus correspondientes intervalos de confianza, se hace el cálculo asumiendo un muestreo aleatorio simple, pero en este caso corrigiendo el error estándar con base al efecto del uso de conglomerados en la primera etapa de la selección.

## Factor de ajuste

Se utilizaron factores de ajuste en el proceso de estimación, pensando en ajustar las distribuciones marginales de algunas variables geográficas y demográficas (de principal interés para el estudio) a los porcentajes totales nacionales. Para el cálculo del factor de ajuste se consideraron las siguientes variables:

VARIABLES DEMOGRÁFICAS	
VARIABLE	CATEGORÍAS
Sexo	Hombre
	Mujer
Edad	12-17 años
	18-22 años
	23-30 años
	31-45 años
	46-55 años
	56 años o más
Escolaridad	Ninguna
	Primaria
	Secundaria
	Bachillerato
	Universidad o más

VARIABLES GEOGRÁFICAS	
VARIABLE	CATEGORÍAS
Región	Sur
	Noroeste
	Centro
	Centro Occidente
	Noreste
	Distrito Federal
Municipios por estrato de población	2,500 hab. O menos
	2,501 a 15,000 hab.
	15,001 a 50,000 hab.
	50,0001 a 100,000 hab.
	100,001 a 500,000 hab.
	Más de 500,000 hab.

El factor de ajuste se generó de acuerdo a los porcentajes nacionales marginales de cada uno de los grupos de las variables anteriores. Cabe agregar que todos los porcentajes nacionales se tomaron de datos oficiales del Censo de Población y vivienda 2010, considerando como base el número de habitantes de 12 años o más, población objetivo del estudio.

El algoritmo utilizado para generar los factores fue el *Ajuste Proporcional Iterativo*, el cual permite corregir los marginales de la distribución de la muestra, ajustándolo a los porcentajes nacionales. Dado el número de variables que se consideraron, el factor de ajuste para algunas observaciones resultó demasiado alto o bajo en algunos casos (en alrededor del 19% de las observaciones). Por ello, se decidió limitar los factores de ajuste calculados a un intervalo, de manera que ninguna observación recibiera un peso excesivamente alto o bajo. Por tal motivo, los factores de ajuste calculados, se delimitaron al intervalo [0.3 , 4.0]. El peso “excesivo” de las observaciones que se limitaron fue dividido de forma equitativa entre las observaciones restantes, aunque manteniendo una asignación estricta para que el ajuste no excediera en ningún caso los límites establecidos. Debido a este ajuste, las distribuciones marginales ajustadas de la muestra se aproximan a los porcentajes nacionales, aunque sin ser exactamente iguales.

Todos los cálculos y estimaciones se hicieron utilizando alguna combinación de los siguiente software de análisis:

- SPSS Inc. Released 2008. SPSS Statistics for Windows, Version 17.0. Chicago: SPSS Inc.
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

## Margen de error

En la siguiente tabla se muestra el nivel de precisión aproximado para una proporción estimada en la Encuesta Nacional de Lectura y Escritura 2015, manteniendo fija una confiabilidad del 95% y bajo el supuesto de un Muestreo Aleatorio Simple y máxima varianza. Se muestra también el nivel de precisión aproximado, corrigiendo por un valor de efecto de diseño conservador (DEFF=4).

		NIVEL DE CONFIANZA 95%	
		Margen de error (M.A.S.)	Margen de error (DEFF=4)
Total nacional		<b>±1.3%</b>	<b>±2.6%</b>
Región geográfica	Sur	±3.4%	±6.7%
	Noroeste	±3.5%	±6.9%
	Centro	±3.1%	±6.1%
	Centro Occidente	±2.7%	±5.4%
	Noreste	±2.7%	±5.5%
	Distrito Federal	±4.5%	±8.9%
Zona metropolitana	Ciudad de México	±3.4%	±6.7%
	Monterrey	±4.2%	±8.3%
	Guadalajara	±4.0%	±8.0%