

ENCUESTA SOBRE DISCRIMINACIÓN EN LA CIUDAD DE MÉXICO 2026 (EDIS-CDMX, 2026)

Informe metodológico del procedimiento de ponderación

Contenido

1. Introducción	3
2. Objetivos de la ponderación.....	3
3. Fundamentos metodológicos de la ponderación.....	4
3.1. Población objeto de estudio y muestra final.....	4
3.2. Diseño de muestreo.....	5
3.3. Probabilidad de selección	6
3.3.1. Probabilidad de selección para las UPM.....	7
3.3.2. Probabilidad de selección para los individuos	7
4. Ponderador muestral.....	9
5. Calibración mediante <i>raking</i>	11
5.1. Variables auxiliares y recodificación.....	11
5.2. Distribuciones marginales de referencia de edad-sexo	13
5.3. Distribuciones marginales de referencia de escolaridad.....	15
5.4. Distribuciones marginales de referencia de alcaldía.....	15
5.5. Construcción del ponderador final	17
6. Efecto de diseño y margen de error de las estimaciones.....	17
7. Tasas de respuesta.....	19
7.1. Tasa de respuesta.....	21
7.2. Tasa de cooperación.....	21
7.3. Tasa de rechazo.....	22
7.4. Tasa de contacto	22
8. Limitaciones y alcances.....	22
9. Entregables técnicos asociados	23
ANEXOS.....	24
Anexo 1. Código de programación en SPSS para el cálculo del ponderador final	25

1. Introducción

El levantamiento de la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026) concluyó con un total de 4,410 entrevistas completas, realizadas durante los meses de junio y julio de 2026 en 441 Unidades Primarias de Muestreo conforme al diseño muestral previamente establecido, abarcando las 16 alcaldías de la Ciudad de México y garantizando la cobertura geográfica definida para la población objetivo del estudio. Una vez finalizada la recolección de información, se llevó a cabo el cierre metodológico, que comprendió la revisión de los resultados de contacto para el cálculo de tasas de respuesta, la clasificación de las disposiciones de elegibilidad de los casos y la construcción de los factores de expansión necesarios para el análisis de la encuesta.

El presente informe documenta los procedimientos aplicados para transformar la muestra obtenida en estimaciones representativas de la población de 15 años y más residente en viviendas particulares de la Ciudad de México. Para ello, se describe la construcción del ponderador muestral y su posterior calibración mediante un procedimiento iterativo de *raking*, utilizando como variables de ajuste los grupos de edad y sexo y el nivel de escolaridad del entrevistado. Asimismo, se presentan las tasas de respuesta conforme a las definiciones de la Asociación Americana de Estudios de Opinión Pública (AAPOR) y el cálculo del error muestral para una estimación de la encuesta.

2. Objetivos de la ponderación

La ponderación constituye una etapa fundamental del procesamiento estadístico de la encuesta, cuyo propósito es garantizar que las estimaciones derivadas de la muestra sean representativas de la población objetivo. Mediante la asignación de factores de expansión y su posterior calibración, cada entrevista aporta a las estimaciones en función de la proporción de la población que representa, considerando las características del diseño muestral y la distribución de variables sociodemográficas seleccionadas.

De esta manera, la ponderación cumple las siguientes funciones:

- Ajustar las estimaciones para compensar diferencias observadas entre la composición de la muestra final y la estructura poblacional para personas de 15 años o más en la Ciudad de México.

- Reducir los posibles efectos derivados de la no respuesta y de las sustituciones realizadas durante el operativo de campo, preservando la representatividad de las estimaciones.
- Generar una base de datos calibrada para la producción de indicadores estadísticamente consistentes con la población de 15 años o más en la Ciudad de México.

Es importante señalar que la ponderación no sustituye los controles de calidad implementados durante el levantamiento, ni elimina por sí misma todas las fuentes de error inherentes a una encuesta. Su función consiste en asignar a cada entrevista un peso acorde con la población que representa y corregir diferencias conocidas entre la muestra final y la estructura poblacional utilizada para la calibración, contribuyendo así a mejorar la calidad y representatividad de las estimaciones.

3. Fundamentos metodológicos de la ponderación

La construcción del ponderador requiere partir de los elementos que sustentan el diseño estadístico de la encuesta. En este apartado se describen los principales fundamentos metodológicos considerados para el cálculo de los factores de expansión y la posterior calibración de la muestra, incluyendo la población objeto de estudio, el diseño de muestreo y las probabilidades de selección.

3.1. Población objeto de estudio y muestra final

La población objeto de estudio está integrada por personas de 15 años cumplidos o más, residentes permanentes en viviendas particulares localizadas dentro del territorio de la Ciudad de México. La base final de análisis contiene 4,410 entrevistas completas, equivalentes a la meta prevista para el estudio.

Como resultado del operativo de campo, las entrevistas se distribuyeron territorialmente en las 16 alcaldías de la Ciudad de México, manteniendo la cobertura prevista en las Áreas Geoestadísticas Básicas (AGEB) seleccionadas en el diseño muestral. La Figura 1 muestra la distribución territorial de las Áreas Geoestadísticas Básicas (AGEB) en las que se realizaron entrevistas. Se observa una amplia dispersión geográfica de la muestra en las 16 alcaldías de la Ciudad de México, lo que evidencia la cobertura territorial alcanzada durante el levantamiento y la adecuada implementación del diseño muestral.

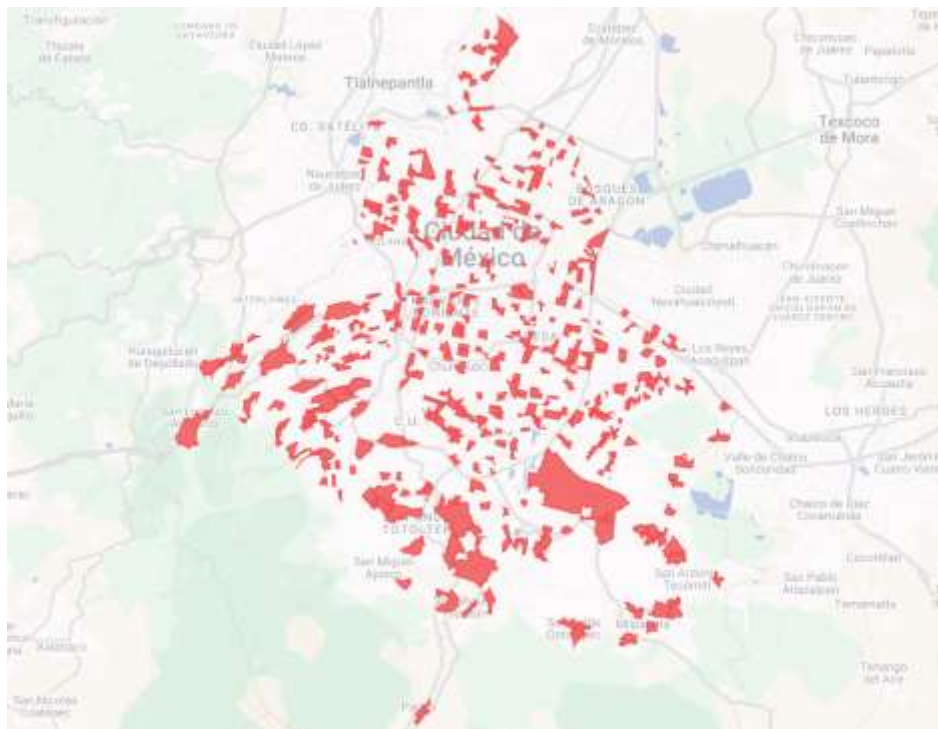


Imagen 1: Distribución territorial de las Áreas Geoestadísticas Básicas (AGEB) de la muestra (en rojo).

Fuente: Elaboración propia con base en información cartográfica del INEGI y Google My Maps.

Descripción: Mapa que muestra la ubicación de las Áreas Geoadministrativas Básicas (AGEB) en las que se realizaron entrevistas durante el levantamiento de la encuesta EDIS 2026. La distribución espacial evidencia la amplia cobertura territorial y la adecuada dispersión geográfica de la muestra en las 16 alcaldías de la Ciudad de México.

3.2. Diseño de muestreo

El diseño fue polietápico y va desde unidades geográficas grandes, como lo son las AGEB, hasta la selección del individuo. Las cuatro etapas de selección fueron:

- **Unidades Primarias de Muestreo (UPM):** Selección de AGEB urbanas y localidades rurales mediante un muestreo estratificado y sistemático con probabilidad proporcional al tamaño (PPT), asegurando distribución territorial y representatividad poblacional.
- **Unidades Secundarias de Muestreo (USM):** En cada UPM seleccionada, se determinan dos puntos de arranque para el levantamiento. En zonas urbanas corresponden a manzanas elegidas aleatoriamente; en áreas rurales se utilizan segmentos de viviendas.

- Unidades Terciarias de Muestreo (UTM): A partir de cada punto de arranque, se seleccionan viviendas mediante un recorrido sistemático establecido hasta completar las entrevistas asignadas.
- Unidades Cuaternarias de Muestreo (UQM): En cada vivienda elegible, se selecciona aleatoriamente a una persona de 15 años o más para ser entrevistada.

Tabla 1: Etapas, unidades y métodos del diseño muestral

Etapas	Unidad de muestreo	Método de selección	Cantidad seleccionada
Unidades Primarias de Muestreo (UPM)	AGEBS urbanas o localidades rurales	Muestreo estratificado y proporcional al tamaño	441 en total
Unidades Secundarias de Muestreo (USM)	Manzanas de punto de arranque	Aleatorio simple sin reemplazo	2 por UPM (882 en total)
Unidades Terciarias de Muestreo (UTM)	Hogares	Selección sistemática con método de espiral	5 por USM (4410 en total)
Unidades Cuaternarias de Muestreo (UQM)	Personas (15 años o más)	Aleatorio simple (igual probabilidad entre integrantes del hogar)	1 por UTM (4410 en total)

Fuente: Elaboración propia, diseño muestral EDIS CDMX 2026.

Descripción: La tabla resume las cuatro etapas del diseño muestral polietápico implementado para la EDIS CDMX 2026, detallando las unidades de muestreo, el método de selección y el número de unidades previstas en cada nivel. El esquema asegura representatividad y control de la variabilidad en cada etapa, desde la selección de UPM hasta la del individuo dentro del hogar.

3.3. Probabilidad de selección

La probabilidad de selección permite cuantificar la oportunidad que tiene cada unidad de ser incluida en la muestra en cada etapa del proceso de muestreo. En un diseño probabilístico, conocer estas probabilidades es indispensable para garantizar la representatividad de los resultados, así como para la construcción de los factores de expansión y los pesos muestrales utilizados en la estimación de indicadores. En particular, dado que la primera etapa del diseño se basa en un muestreo estratificado con probabilidad proporcional al tamaño (PPT), las UPM con mayor población de 15 años o más tienen una mayor probabilidad de ser seleccionadas, manteniendo al mismo tiempo la igualdad de probabilidad final a nivel individuo.

3.3.1. Probabilidad de selección para las UPM

Bajo el esquema descrito, la probabilidad de selección de la i -ésima Unidad Primaria de Muestreo en el estrato h se define como:

$$P_{ih} = \frac{M_{ih}}{\sum_{j=1}^{N_h} M_{jh}} \times n_h$$

Donde:

M_{ih} = Población de 15 años o más en la UPM i del estrato h .

N_h = Número total de UPM en el estrato h .

n_h = Número de UPM a seleccionar en el estrato h .

3.3.2. Probabilidad de selección para los individuos

La probabilidad de selección de una persona j corresponde al producto de la probabilidad de seleccionar la UPM y la probabilidad de seleccionar dicha persona dentro de la UPM correspondiente. Debido a que el diseño muestral contempló un número constante de **10 entrevistas por UPM**, la probabilidad de selección de un individuo puede expresarse como:

$$P_{ihj} = P_{ih} * \frac{10}{M_{ih}}$$

Sustituyendo la expresión de la probabilidad de selección de la UPM se obtiene:

$$P_{ihj} = \left(\frac{M_{ih}}{\sum_{j=1}^{N_h} M_{jh}} \times n_h \right) * \frac{10}{M_{ih}} = \frac{10n_h}{\sum_{j=1}^{N_h} M_{jh}}$$

Como puede observarse, el término M_{ih} se cancela algebraicamente, por lo que la probabilidad final de selección deja de depender del tamaño poblacional de la UPM. Es decir, una vez incorporada la selección de personas dentro de cada UPM, las diferencias de tamaño entre las unidades primarias dejan de influir en la probabilidad final de selección.

Si además se consideran las siguientes características del diseño muestral:

- 1) El tamaño de muestra para cada estrato (n_h) tuvo una asignación proporcional al tamaño de su población de 15 años y más. En términos prácticos, la proporción que representa cada alcaldía respecto de la población total de la Ciudad de México es aproximadamente la misma proporción de las 441 UPM que le fueron asignadas. Es decir,

$$n_h \approx \frac{\text{Población en la alcaldía } h}{\text{Población en la CDMX}} * 441$$

- 2) El término

$$\sum_{j=1}^{N_h} M_{jh}$$

corresponde a la población total de 15 años y más del estrato h , que en este estudio equivale a la población de cada alcaldía.

Sustituyendo esta relación en la expresión anterior, la probabilidad de selección puede aproximarse como:

$$P_{ihj} = \frac{10n_h}{\sum_{j=1}^{N_h} M_{jh}} \approx \frac{10 * \frac{\text{Población en la alcaldía } h}{\text{Población en la CDMX}} * 441}{\text{Población en la alcaldía } h} \approx \frac{4410}{\text{Población en la CDMX}}$$

La expresión anterior muestra que la probabilidad final de selección es prácticamente constante para todas las personas de la población objetivo, independientemente de la alcaldía o del tamaño de la UPM en la que residen. Esto es consecuencia de la asignación proporcional de la muestra entre estratos, del esquema de selección de UPM con probabilidad proporcional al tamaño (PPT) y del número fijo de entrevistas por UPM, características que, en conjunto, producen un **diseño prácticamente autoponderado**.

De acuerdo con el Censo de Población y Vivienda 2020 del INEGI, la población de 15 años y más residente en la Ciudad de México asciende a 7,547,545 personas. En consecuencia, la probabilidad aproximada de selección de una persona es:

$$P_{ihj} \approx \frac{4410}{7,547,545} = 0.0005843$$

Lo anterior significa que, bajo el diseño muestral implementado, todas las personas de 15 años y más residentes en la Ciudad de México tuvieron una probabilidad de selección prácticamente uniforme, cercana al 0.0584%.

4. Ponderador muestral

El ponderador muestral es el factor que se asigna a cada unidad seleccionada con el propósito de representar el número de personas de la población objetivo que dicha unidad representa dentro del universo de estudio. Su cálculo se basa en las probabilidades de selección definidas por el diseño muestral y constituye el mecanismo mediante el cual la información obtenida en la muestra se expande para producir estimaciones representativas de la población objetivo. Posteriormente, este ponderador puede ser ajustado mediante técnicas de calibración para mejorar la correspondencia entre la muestra y la estructura poblacional de referencia.

Como se mostró en la sección anterior, la probabilidad aproximada de selección de una persona en la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026) es aproximadamente de 0.0005843. Y dado que el ponderador muestral corresponde al inverso de la probabilidad de selección, el ponderador de diseño se obtiene como:

$$P_{thj} \approx \frac{1}{P_{thj}} = \frac{1}{0.0005843} = 1,711.45$$

Este valor representa el ponderador de diseño, es decir, el número promedio de personas de 15 años y más de la Ciudad de México que representa cada entrevista antes de aplicar cualquier ajuste derivado de la distribución final de la muestra o de los procesos posteriores de calibración.

Sin embargo, debido a que la asignación de UPM por estrato se realizó de manera proporcional al tamaño de la población, el número teórico de UPM que correspondía a cada alcaldía debía redondearse a valores enteros, ya que no es posible seleccionar fracciones de UPM ni realizar un número fraccionario de entrevistas. Como consecuencia, el número efectivo de entrevistas por estrato difiere ligeramente del valor teórico esperado, lo que genera pequeñas variaciones en el ponderador básico entre alcaldías.

Análogamente, los factores de expansión por estrato se obtienen al dividir la población de 15 años y más de cada estrato entre el número de entrevistas efectivamente realizadas

en dicho estrato. La Tabla 2 presenta, para cada estrato (alcaldía), la población de referencia de 15 años y más, su participación porcentual respecto del total de la Ciudad de México, el número de entrevistas efectivamente realizadas y el factor de expansión correspondiente.

Tabla 2: Distribución por estrato de UPM en la muestra y ponderador muestral

Alcaldía	Población de 15 años y más	% de Población de 15 años y más	Número de entrevistas	Factor de expansión
Álvaro Obregón	622,262	8.2%	360	1728.51
Azcapotzalco	361,474	4.8%	210	1721.30
Benito Juárez	382,872	5.1%	220	1740.33
Coyoacán	522,992	6.9%	310	1687.07
Cuajimalpa de Morelos	172,263	2.3%	100	1722.63
Cuauhtémoc	462,349	6.1%	270	1712.40
Gustavo A. Madero	960,772	12.7%	560	1715.66
Iztacalco	335,205	4.4%	200	1676.03
Iztapalapa	1,468,256	19.5%	860	1707.27
La Magdalena Contreras	200,201	2.7%	120	1668.34
Miguel Hidalgo	346,049	4.6%	200	1730.25
Milpa Alta	116,360	1.5%	70	1662.29
Tláhuac	308,117	4.1%	180	1711.76
Tlalpan	570,245	7.6%	330	1728.02
Venustiano Carranza	365,208	4.8%	210	1739.09
Xochimilco	352,920	4.7%	210	1680.57
Ciudad de México	7,547,545	100%	4,410	

Fuente: Elaboración propia con información del INEGI (Censo de Población y Vivienda 2020, bases ITER_NALCSV20 y RESAGEBURB_09_2020).

Descripción: La tabla muestra el número de entrevistas y el factor de expansión por alcaldía.

Como puede observarse, los ponderadores presentan valores muy similares entre alcaldías, cercanos a 1,700 personas por entrevista. Las diferencias observadas obedecen exclusivamente al redondeo realizado durante la asignación del número de UPM y, en consecuencia, del número de entrevistas por estrato, por lo que no modifican la característica fundamental del diseño muestral, consistente en mantener probabilidades de selección y factores de expansión prácticamente uniformes en toda la población objetivo.

5. Calibración mediante *raking*

Una vez calculados los factores de expansión derivados del diseño muestral, éstos fueron ajustados mediante un procedimiento de calibración por *raking* (iterative proportional fitting), con el propósito de mejorar la correspondencia entre la distribución de la muestra y la estructura sociodemográfica conocida de la población en la Ciudad de México.

El procedimiento de *raking* consiste en realizar ajustes iterativos sobre los factores de expansión hasta que las distribuciones marginales de determinadas variables de control en la muestra ponderada coincidan con las distribuciones observadas en la población. En este estudio, las variables utilizadas para la calibración fueron sexo, grupos de edad y nivel educativo, utilizando como referencia los resultados del Censo de Población y Vivienda 2020 del Instituto Nacional de Estadística y Geografía (INEGI).

El proceso de ajuste se realiza de manera secuencial para cada una de las variables de control. En cada iteración se calcula, para cada categoría de la variable correspondiente (por ejemplo, hombres y mujeres en el caso del sexo), un factor de ajuste obtenido como la razón entre el total poblacional conocido y el total estimado a partir de la muestra. Posteriormente, los factores de expansión de las observaciones pertenecientes a cada categoría se multiplican por dicho factor de ajuste. Este procedimiento se repite sucesivamente para todas las variables de control hasta alcanzar la convergencia, es decir, hasta que las diferencias entre las distribuciones ponderadas de la muestra y los totales poblacionales sean prácticamente nulas.

Como resultado de este proceso se obtiene un ponderador calibrado, el cual conserva las propiedades del diseño muestral y, al mismo tiempo, reproduce la estructura poblacional conocida para las variables de calibración, mejorando la representatividad de las estimaciones obtenidas a partir de la encuesta.

5.1. Variables auxiliares y recodificación

Previo a la aplicación del procedimiento de calibración mediante *raking*, se verificó que todas las variables de control utilizadas en el ajuste de los factores de expansión contaran con categorías válidas y completas para la totalidad de las entrevistas.

En la variable sexo asignado al nacer se identificaron 26 casos en los que la pregunta "¿Cuál es el sexo que le asignaron al nacer (el que aparece en su acta de nacimiento original)?" fue respondida como "No estoy segura(o)" o "Prefiere no contestar". Sin embargo, el procedimiento de calibración requiere que todas las unidades de análisis dispongan de una categoría válida en cada una de las variables utilizadas como controles poblacionales.

En este estudio, los controles de referencia provinieron del Censo de Población y Vivienda 2020 del Instituto Nacional de Estadística y Geografía (INEGI), cuyas distribuciones para el sexo asignado al nacer únicamente contemplan las categorías hombre y mujer, sin considerar categorías de no respuesta. En consecuencia, fue necesario asignar una categoría válida a dichos registros para garantizar la correcta ejecución del algoritmo de calibración.

Para ello, se utilizó la información registrada en el censo inicial de contacto, instrumento empleado durante el trabajo de campo para identificar a la persona seleccionada y verificar el cumplimiento de las características de selección por sexo y grupo de edad, así como para efectuar las sustituciones autorizadas cuando la persona originalmente seleccionada no aceptaba participar. Dado que esta información corresponde al sexo asignado al nacer utilizado durante el proceso de selección muestral, se consideró la fuente más consistente para completar los 26 registros con información faltante en esta variable.

Una vez realizada esta imputación, todas las entrevistas contaban con información completa en las variables de control utilizadas para la calibración de los factores de expansión.

Además, las variables de edad y escolaridad fueron recodificadas en un número reducido de categorías para el proceso de calibración mediante *raking*. Esta recodificación permite mejorar la estabilidad del algoritmo y favorecer su convergencia, evitando una excesiva fragmentación de la muestra en categorías con pocos casos, lo que podría generar factores de expansión inestables o extremos y aumentar la variabilidad de las estimaciones. Las categorías utilizadas durante el procedimiento de calibración se presentan en la Tabla 3.

Tabla 3: Variables utilizadas en la calibración

Variable	Origen o cálculo	Categorías / tipo	Función
sexo	p25, con imputación puntual para 26 casos	1 = Hombre; 2 = Mujer	Componente del cruce edad-sexo
edad	S0, recodificación en 5 grupos de edad	15 a 29 años 30 a 39 años 40 a 49 años 50 a 59 años 60 años o mayor	Componente del cruce edad-sexo
escolaridad	p17, recodificación en 4 grupos de nivel educativo	Hasta primaria: 0=Ninguno; 1=preescolar; 2=primaria; 99=prefiere no contestar. Hasta secundaria: 3=secundaria Educación media superior: 4= Preparatoria o bachillerato Educación superior 5=Carrera técnica con preparatoria terminada; 6=Licenciatura o profesional; 7=Especialidad; 8=Maestría o doctorado	Ajustar nivel educativo
estrato	Variable territorial de la muestra	16 alcaldías	Preservar la estructura territorial

Fuente: Elaboración propia, sintaxis de ponderación para EDIS CDMX 2026.

Descripción: La tabla identifica las variables auxiliares, su recodificación y su función dentro del ajuste iterativo.

5.2. Distribuciones marginales de referencia de edad-sexo

La distribución conjunta de sexo y grupos de edad constituye uno de los principales componentes de la estructura demográfica de la población, ya que ambas variables se encuentran estrechamente asociadas con múltiples características sociales, económicas y culturales de las personas. En consecuencia, su incorporación como variables de calibración contribuye a reducir posibles sesgos derivados de diferencias en la composición

de la muestra respecto de la población objetivo y mejora la representatividad de las estimaciones obtenidas a partir de la encuesta.

Antes de la calibración se observan diferencias entre la composición de la muestra y la distribución poblacional de referencia. En particular, la muestra presenta una sobrerrepresentación de personas de 50 años y más, especialmente mujeres, y una subrepresentación de la población de 15 a 49 años, principalmente hombres. Estas diferencias son corregidas mediante el procedimiento de *raking*, el cual ajusta los factores de expansión para que la distribución ponderada de la muestra reproduzca las proporciones poblacionales observadas en el Censo de Población y Vivienda 2020 del INEGI. La Tabla 4 presenta las distribuciones utilizadas como referencia para este proceso de calibración.

Tabla 4: Distribución de la muestra y distribución poblacional de referencia de edad-sexo

Grupo	Entrevistas	Porcentaje sin ponderar	Total de referencia	Proporción objetivo
1 Hombres 15 a 29 años	468	10.6%	1,062,239	14.1%
2 Mujeres 15 a 29 años	494	11.2%	1,055,044	14.0%
3 Hombres 30 a 39 años	258	5.9%	703,993	9.3%
4 Mujeres 30 a 39 años	383	8.7%	745,468	9.9%
5 Hombres 40 a 49 años	284	6.4%	631,158	8.4%
6 Mujeres 40 a 49 años	364	8.3%	708,094	9.4%
7 Hombres 50 a 59 años	339	7.7%	526,045	7.0%
8 Mujeres 50 a 59 años	503	11.4%	623,885	8.3%
9 Hombres más de 60 años	620	14.1%	638,412	8.5%
10 Mujeres más de 60 años	697	15.8%	853,207	11.3%
Total	4,410	100%	7,547,545	100%

Fuente: Elaboración propia con información de la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026) y datos del CENSO de población 2020 del INEGI.

Descripción: La tabla presenta la distribución de las entrevistas sin ponderar y la distribución poblacional de referencia por sexo y grupo de edad, empleada como control para la calibración de los factores de expansión mediante el procedimiento de *raking*.

5.3. Distribuciones marginales de referencia de escolaridad

El nivel educativo constituye otra de las variables de calibración, debido a su estrecha relación con las condiciones socioeconómicas, el acceso a oportunidades y diversos patrones de comportamiento de la población. Aunque la distribución de la muestra presenta diferencias relativamente pequeñas respecto de la población de referencia, se observa una ligera subrepresentación de las personas con primaria y una sobrerrepresentación de quienes cuentan con licenciatura o estudios superiores. Estas diferencias son corregidas mediante el procedimiento de *raking*. La Tabla 5 presenta las distribuciones utilizadas como referencia para la calibración.

Tabla 5: Distribución de la muestra y distribución poblacional de referencia de escolaridad

Grupo	Entrevistas	Porcentaje sin ponderar	Total de referencia	Proporción objetivo
1 Primaria	575	13.0%	1,141,284	15.1%
2 Secundaria	976	22.1%	1,705,811	22.6%
3 Preparatoria	1,226	27.8%	2,086,124	27.6%
4 Licenciatura o más	1,633	37.0%	2,614,326	34.6%
Total	4,410	100%	7,547,545	100%

Fuente: Elaboración propia con información de la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026) y datos del CENSO de población 2020 del INEGI.

Descripción: La tabla presenta la distribución de las entrevistas sin ponderar y la distribución poblacional de referencia por grupos de nivel educativo, empleada como control para la calibración de los factores de expansión mediante el procedimiento de *raking*.

5.4. Distribuciones marginales de referencia de alcaldía

Finalmente, se incorporó la alcaldía de residencia como variable de calibración con el propósito de garantizar la correspondencia exacta entre la distribución territorial de la muestra y la población de referencia y para evitar un desajuste importante en esta variable al momento de buscar la convergencia de los porcentajes objetivo de las otras variables de control (edad-sexo y escolaridad). Como se mencionó en secciones anteriores, debido a que el diseño muestral contempló una asignación proporcional de las UPM por alcaldía, la distribución observada en la muestra ya presentaba una elevada concordancia con los valores poblacionales. Las pequeñas diferencias existentes obedecen únicamente al

redondeo realizado durante la asignación del número de UPM y entrevistas por estrato, por lo que el ajuste mediante *raking* constituye un refinamiento final para asegurar la coincidencia exacta con los controles poblacionales. La Tabla 6 presenta las distribuciones utilizadas durante el proceso de calibración.

Tabla 6: Distribución de la muestra y distribución poblacional de referencia de alcaldía

Grupo	Entrevistas	Porcentaje sin ponderar	Total de referencia	Proporción objetivo
Álvaro Obregón	360	8.2%	622,262	8.2%
Azcapotzalco	210	4.8%	361,474	4.8%
Benito Juárez	220	5.0%	382,872	5.1%
Coyoacán	310	7.0%	522,992	6.9%
Cuajimalpa de Morelos	100	2.3%	172,263	2.3%
Cuauhtémoc	270	6.1%	462,349	6.1%
Gustavo A. Madero	560	12.7%	960,772	12.7%
Iztacalco	200	4.5%	335,205	4.4%
Iztapalapa	860	19.5%	1,468,256	19.5%
La Magdalena Contreras	120	2.7%	200,201	2.7%
Miguel Hidalgo	200	4.5%	346,049	4.6%
Milpa Alta	70	1.6%	116,360	1.5%
Tláhuac	180	4.1%	308,117	4.1%
Tlalpan	330	7.5%	570,245	7.6%
Venustiano Carranza	210	4.8%	365,208	4.8%
Xochimilco	210	4.8%	352,920	4.7%
Total	4,410	100%	7,547,545	100%

Fuente: Elaboración propia con información de la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026) y datos del CENSO de población 2020 del INEGI.

Descripción: La tabla presenta la distribución de las entrevistas sin ponderar y la distribución poblacional de referencia por alcaldía, empleada como control para la calibración de los factores de expansión mediante el procedimiento de *raking*.

5.5. Construcción del ponderador final

La base de datos de la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026) incorpora la variable **“ponderador”**, la cual constituye el ponderador final que debe utilizarse en la obtención de todas las estimaciones estadísticas de la encuesta. Este ponderador integra el factor de expansión derivado del diseño muestral con los ajustes de calibración efectuados mediante el procedimiento de *raking*, por lo que refleja tanto las probabilidades de selección como la estructura demográfica de la población de referencia.

Una vez concluido el proceso de calibración, el ponderador final satisface los criterios metodológicos establecidos para su utilización. En particular, la suma de los ponderadores reproduce el total de la población objetivo de la Ciudad de México (7,547,545 personas de 15 años y más); las distribuciones ponderadas de las variables utilizadas en la calibración (sexo y grupos de edad, escolaridad y alcaldía) coinciden con los controles poblacionales provenientes del Censo de Población y Vivienda 2020 del INEGI; y todas las observaciones cuentan con un ponderador válido, sin valores faltantes, iguales a cero o negativos.

En consecuencia, la variable **“ponderador”** constituye el insumo que debe emplearse en todos los análisis descriptivos e inferenciales derivados de la encuesta, garantizando que las estimaciones obtenidas sean representativas de la población de 15 años y más residente en la Ciudad de México.

6. Efecto de diseño y margen de error de las estimaciones

Una vez construido el ponderador final, las estimaciones derivadas de la encuesta incorporan tanto las probabilidades de selección definidas por el diseño muestral como los ajustes de calibración realizados durante el proceso de ponderación, garantizando su representatividad respecto de la población objetivo. Como complemento de este proceso, el presente apartado describe la precisión de las estimaciones mediante el cálculo del margen de error, el cual incorpora las características del diseño muestral y el ponderador utilizado en el análisis.

La precisión de las estimaciones de una encuesta depende tanto del tamaño de la muestra como de las características de su diseño. En los diseños por conglomerados, las personas seleccionadas dentro de una misma UPM pueden presentar características más similares entre sí que las personas seleccionadas en distintas zonas, debido a que

comparten determinadas condiciones territoriales, sociales y económicas. Esta semejanza dentro de los conglomerados puede incrementar la varianza de las estimaciones respecto de la que se obtendría mediante un muestreo aleatorio simple del mismo tamaño.

El efecto de este tipo de diseño sobre la precisión se resume mediante el efecto de diseño, conocido como *DEFF*, que se define como la razón entre la varianza de una estimación obtenida bajo el diseño muestral complejo y la varianza que tendría la misma estimación bajo un muestreo aleatorio simple de igual tamaño:

$$DEFF = \frac{VAR_{Diseño\ complejo}(\hat{p})}{VAR_{MAS}(\hat{p})}$$

De manera aproximada, el efecto de diseño asociado al muestreo por conglomerados puede expresarse como:

$$DEFF = 1 + \rho(K - 1)$$

donde K representa el número promedio de personas entrevistadas dentro de cada conglomerado y ρ es el coeficiente de correlación intraconglomerado. Un *DEFF* igual a 1 indica una precisión equivalente a la de un muestreo aleatorio simple; valores superiores a 1 reflejan un incremento de la varianza ocasionado por el diseño muestral.

Considerando el efecto de diseño, el margen de error de una proporción estimada puede calcularse mediante la siguiente expresión:

$$ME = Z_{\alpha/2} \sqrt{DEFF \left(\frac{\hat{p}(1 - \hat{p})}{n} \right)}$$

donde $\hat{p}$ es la proporción estimada, n es el número de entrevistas efectivas y $Z_{\alpha/2}$ corresponde al valor crítico asociado con el nivel de confianza seleccionado. Para un nivel de confianza de 95%, se utiliza un valor de $Z = 1.96$.

Como ejemplo de la precisión de las estimaciones obtenidas mediante la EDIS CDMX 2026, la Tabla 7 presenta los principales indicadores de precisión para la pregunta: "En los últimos 4 años, de junio/julio de 2022 a la fecha, ¿ha sido discriminada(o) o menospreciada(o)

en la Ciudad de México, por su forma de vestir o arreglo personal (tatuajes, ropa, arreglo de su cabello o perforaciones)". Para cada categoría de respuesta se reporta el punto estimado, el error estándar, el margen de error, el intervalo de confianza al 95%, el efecto de diseño (*DEFF*) y el coeficiente de variación (*CV*), calculados considerando el diseño muestral complejo y el ponderador final de la encuesta.

Tabla 7: Indicadores de precisión de las estimaciones para la pregunta sobre discriminación por la forma de vestir o arreglo personal

Respuesta	Punto Estimado	Error estándar	Margen de error	Límite inferior	Límite superior	<i>DEFF</i>
Sí	19.9	0.6	+/- 1.3	18.6	21.1	1.1
No	79.7	0.6	+/- 1.3	78.5	81	1.2
Prefiere no responder	0.4	0.1	+/- 0.2	0.2	0.6	0.9

Fuente: Elaboración propia con información de la Encuesta sobre Discriminación en la Ciudad de México (EDIS CDMX, 2026).

Descripción: La tabla presenta, a manera de ejemplo, los principales indicadores de precisión para las estimaciones correspondientes a la pregunta sobre discriminación por la forma de vestir o arreglo personal. Para cada categoría de respuesta se reporta el punto estimado, el error estándar, el margen de error, los límites del intervalo de confianza al 95% y el efecto de diseño (*DEFF*), calculados considerando el diseño muestral complejo y el ponderador final de la EDIS CDMX 2026. Los valores de estos indicadores varían entre preguntas, ya que dependen de la distribución de las respuestas y de las características de cada estimación.

7. Tasas de respuesta

Una vez descrito el procedimiento de construcción y calibración de los factores de expansión, resulta pertinente evaluar el desempeño del proceso de levantamiento que dio origen a la muestra analizada. Para ello, se presentan las tasas de respuesta calculadas conforme a los lineamientos de la Asociación Americana para la Investigación de la Opinión Pública (AAPOR), las cuales constituyen un indicador fundamental de la calidad del trabajo de campo y permiten cuantificar la proporción de unidades seleccionadas que concluyeron exitosamente el proceso de entrevista.

La tasa de respuesta de la encuesta fue del **15.3%**, calculada según la **tasa de respuesta RR1** de los Standard Definitions de AAPOR ([AAPOR, 2024](#)). A continuación se presenta la tabla de disposiciones de contacto, así como el detalle del cálculo de las tasas

de respuesta, siguiendo los lineamientos de AAPOR para garantizar la transparencia y comparabilidad de los resultados.

Tabla 8: Resultados del contacto y elegibilidad de los casos

1	Entrevistas	4,410
1.1	Entrevistas completas	4,410
1.2	Entrevistas parciales	0
2	Elegible, no entrevistado	23,852
2.1	Rechazos	9,538
2.11	Rechazo en vivienda	7,515
2.12	Rechazo del respondiente	2,023
2.2	Interrupción de la entrevista	0
2.3	No contacto	13,896
2.31	Nadie en casa	11,910
2.32	Respondente no disponible	1,986
2.4	Otro	418
2.41	Respondente muerto	0
2.42	Respondente física o mentalmente incapacitado	16
2.43	Problema de lenguaje	12
2.44	Otro	390
3	Elegibilidad desconocida, no entrevista	647
3.1	No se sabe si es un hogar	0
3.2	Inseguro/ no se puede llegar a la vivienda	555
3.3	No se sabe si el respondiente es elegible	0
3.4	Otro	92
4	No elegible	2,280
4.1	No es una vivienda	795
4.2	Vivienda no ocupada	155
4.3	Otro	1,330
4.3.1	Respondente fuera del rango de edad elegible	1,330

Fuente: Elaboración propia con información de la EDIS CDMX, 2026.

Descripción: La tabla presenta los resultados de los intentos de contacto con los entrevistados y la elegibilidad de los casos, según los estándares de la American Association for Public Opinion Research (AAPOR).

Con el fin de transparentar el proceso de campo y dar cumplimiento a las definiciones establecidas por AAPOR, se presenta a continuación la distribución de los códigos de disposición de casos. Esta clasificación reúne el total de resultados obtenidos durante el levantamiento, y constituye la base para el cálculo de las tasas de respuesta, cooperación,

contacto y rechazo. Los códigos se presentan conforme a los Standard Definitions de AAPOR.

Tabla 9: Códigos de Disposición Final de Casos (AAPOR)

Disposición	Código	Casos
Entrevistas completas	I	4,410
Entrevistas parciales	P	0
Interrupción de la entrevista y rechazos	R	9,538
No contacto	NC	13,896
Otro	O	418
No se sabe si es un hogar / Elegibilidad desconocida, no entrevista	UH	555
Desconocido si la vivienda contiene un respondiente elegible	UR	0
Desconocido, otro	UO	92

Fuente: Elaboración propia con información de la EDIS CDMX, 2026.

Descripción: La tabla presenta los resultados los códigos de disposición final de casos que se utilizan para el cálculo de tasas, según los estándares de la American Association for Public Opinion Research (AAPOR).

7.1. Tasa de respuesta

$$RR1 = \frac{I}{(I + P) + (R + NC + O) + (UH + UR + UO)}$$

$$RR1 = \frac{4,410}{(4,410 + 0) + (9,538 + 13,896 + 418) + (555 + 0 + 92)} = 15.3\%$$

7.2. Tasa de cooperación

$$COOP1 = \frac{I}{(I + P) + R + O} = \frac{4,410}{(4,410 + 0) + 9,538 + 418} = 30.7\%$$

7.3. Tasa de rechazo

$$REF1 = \frac{R}{(I + P) + (R + NC + O) + (UH + UR + UO)}$$

$$REF1 = \frac{9,538}{(4,410 + 0) + (9,538 + 13,896 + 418) + (555 + 0 + 92)} = 33.0\%$$

7.4. Tasa de contacto

$$CON1 = \frac{(I + P) + R + O}{(I + P) + (R + NC + O) + (UH + UR + UO)}$$

$$CON1 = \frac{(4,410 + 0) + 9,538 + 418}{(4,410 + 0) + (9,538 + 13,896 + 418) + (555 + 0 + 92)} = 49.7\%$$

8. Limitaciones y alcances

La ponderación mejora la correspondencia de la muestra con las distribuciones poblacionales incluidas en el ajuste, pero no corrige errores de medición, problemas de cobertura del marco, respuestas inexactas ni diferencias entre respondentes y no respondentes en variables no observadas. La calidad final de las estimaciones depende conjuntamente del diseño, la ejecución del levantamiento, la clasificación de casos, la limpieza de la base y el uso apropiado de los pesos.

El procedimiento de calibración mediante *raking* puede incrementar la varianza de las estimaciones cuando requiere realizar ajustes muy pronunciados en los factores de expansión. Sin embargo, en la EDIS CDMX 2026 las diferencias entre la distribución de la muestra y la población de referencia fueron, en general, mínimas y se concentraron principalmente en la composición por sexo y grupos de edad, mientras que la distribución territorial por alcaldía presentó únicamente variaciones menores derivadas de la asignación entera de las unidades primarias de muestreo. En consecuencia, el proceso de calibración requirió ajustes moderados, lo que contribuyó a mantener una dispersión controlada de los ponderadores. Esta situación se refleja en los efectos de diseño obtenidos para las estimaciones, los cuales indican que la calibración no produjo incrementos importantes en su variabilidad.

El procedimiento de ponderación de la EDIS CDMX 2026 integra de forma coherente el diseño territorial y la composición sociodemográfica de la población objetivo. El primer componente asigna factores de expansión por alcaldía mediante la relación entre la población de control y el número de entrevistas completas. El segundo componente calibra dichos factores mediante *raking* para reproducir márgenes de edad-sexo, alcaldía y escolaridad.

La base final de 4,410 entrevistas dispone de dos variables diferenciadas: “**ponde_muestral**”, que conserva el factor inicial por estrato, y “**poderador**”, que constituye el peso final recomendado para la estimación de resultados. La suma de los pesos finales debe corresponder al total de control de 7,547,545 personas de 15 años o más.

9. Entregables técnicos asociados

Tabla 10: Productos técnicos de la ponderación

No.	Producto	Contenido
I	Informe sobre los procedimientos y objetivos de la ponderación	El presente documento metodológico que describe el cálculo de los factores de expansión, la construcción del ponderador final, el proceso de calibración mediante <i>raking</i> y la evaluación de la precisión de las estimaciones.
II	Código SPSS para el cálculo de factores de expansión	Sintaxis reproducible incluida como Anexo 1
III	Variable ponderador	Ponderador final calibrado para la estimación de resultados
IV	Base de datos	Base de datos final de la EDIS CDMX 2026 que incorpora el ponderador calibrado y las variables necesarias para realizar estimaciones representativas de la población de 15 años y más residente en la Ciudad de México.

Fuente: Elaboración propia.

Descripción: La tabla resume los entregables técnicos derivados del proceso de ponderación de la EDIS CDMX 2026, así como el contenido de cada uno de ellos.

ANEXOS

Anexo 1. Código de programación en SPSS para el cálculo del ponderador final

```
*****.
** PONDERACION MUESTRAL.
*****.

compute factor_m= 0.
VARIABLE LABELS factor_m 'Factor de ponderacion muestral'.

IF (estrato=1 ) factor_m= 622262.
IF (estrato=2 ) factor_m= 361474.
IF (estrato=3 ) factor_m= 382872.
IF (estrato=4 ) factor_m= 522992.
IF (estrato=5 ) factor_m= 172263.
IF (estrato=6 ) factor_m= 462349.
IF (estrato=7 ) factor_m= 960772.
IF (estrato=8 ) factor_m= 335205.
IF (estrato=9 ) factor_m= 1468256.
IF (estrato=10 ) factor_m= 200201.
IF (estrato=11 ) factor_m= 346049.
IF (estrato=12 ) factor_m= 116360.
IF (estrato=13 ) factor_m= 308117.
IF (estrato=14 ) factor_m= 570245.
IF (estrato=15 ) factor_m= 365208.
IF (estrato=16 ) factor_m= 352920.

compute uno = 1.

* Ponderacion muestral.
  sort cases by estrato.
  aggregate
  /outfile = *
  mode= addvariables
  overwritevars = yes
  /presorted
  /break estrato
  /tmp = sum(unos).

COMPUTE ponde_muestral = factor_m / tmp.
```

freq ponde_muestral.

```
*****.
** POST-ESTRATIFICACION.
*****.
```

```
compute sexo=p25.
IF(id=804732371) sexo=1.
IF(id=796326605) sexo=1.
IF(id=805813523) sexo=1.
IF(id=806900364) sexo=1.
IF(id=802346925) sexo=1.
IF(id=799214532) sexo=2.
IF(id=801773659) sexo=1.
IF(id=819005473) sexo=2.
IF(id=814989936) sexo=2.
IF(id=815560177) sexo=2.
IF(id=815081144) sexo=2.
IF(id=820091285) sexo=1.
IF(id=820095040) sexo=1.
IF(id=820091355) sexo=1.
IF(id=808483622) sexo=2.
IF(id=819059775) sexo=2.
IF(id=807951909) sexo=2.
IF(id=791520796) sexo=2.
IF(id=796318988) sexo=1.
IF(id=806912379) sexo=2.
IF(id=796273123) sexo=2.
IF(id=773399270) sexo=2.
IF(id=791378013) sexo=2.
IF(id=801730974) sexo=2.
IF(id=802342764) sexo=1.
IF(id=819086470) sexo=1.
```

freq sexo p25.

```
ADD VALUE LABELS sexo
1 'Hombre'
2 'Mujer'.
freq sexo.
```

```
*****.
** EDAD Y SEXO.
*****.
compute grupo_edad=s0.
FORMATS grupo_edad (F8.0).
```



```
recode grupo_edad
```

```
(15 thru 29=1)
```

```
(30 thru 39=2)
```

```
(40 thru 49=3)
```

```
(50 thru 59=4)
```

```
(60 thru 999=5).
```

```
compute grupo_edad=grupo_edad*2.
```

```
if (sexo=1) grupo_edad=grupo_edad-1.
```

```
value labels grupo_edad
```

```
1 'Hombres 15 a 29 años'
```

```
2 'Mujeres 15 a 29 años'
```

```
3 'Hombres 30 a 39 años'
```

```
4 'Mujeres 30 a 39 años'
```

```
5 'Hombres 40 a 49 años'
```

```
6 'Mujeres 40 a 49 años'
```

```
7 'Hombres 50 a 59 años'
```

```
8 'Mujeres 50 a 59 años'
```

```
9 'Hombres de 60 años o mas'
```

```
10 'Mujeres de 60 años o mas'.
```

```
CROSSTABS grupo_edad by s0 sexo.
```

```
*****
```

```
** E S C O L A R I D A D.
```

```
*****
```

```
freq p17.
```

```
compute escolaridad=p17.
```

```
recode escolaridad (0,1, 2, 99=1)
```

```
(3=2)
```

```
(4=3)
```

```
(5,6,7,8=4).
```

```
value labels escolaridad
```

```
1 'Hasta primaria'
```

```
2 'Hasta secundaria'
```

```
3 'Educación media superior'
```

```
4 'Educación superior'.
```

```
FORMATS grupo_edad (F8.0).
```

```
VARIABLE LABELS grupo_edad 'Grupo de edad-sexo'.
```

```
FORMATS sexo (F8.0).
```

```
VARIABLE LABELS sexo 'Género de nacimiento'.
```

FORMATS escolaridad (F8.0).

VARIABLE LABELS escolaridad ' Grupo de nivel de escolaridad'.

CROSSTABS escolaridad by p17.

** ITERATIVE FITTING .

** IMPORTANTE: PONDERAR ANTES DE PROCEDIMIENTO.

weight off.

weight by ponde_muestral.

SPSSINC RAKE DIM1= grupo_edad 1 0.14073967 2 0.13978638 3 0.09327444 4
0.09876960 5 0.08362428 6 0.09381779 7 0.06969750 8 0.08266065 9 0.08458539 10
0.11304431

DIM2= estrato 1 0.08244562 2 0.04789292 3 0.05072802 4 0.06929300 5 0.02282371 6
0.06125820 7 0.12729596 8 0.04441245 9 0.19453425 10 0.02652531 11 0.04584921 12
0.01541693 13 0.04082347 14 0.07555371 15 0.04838765 16 0.04675958

DIM3= escolaridad 1 0.15121261 2 0.22600872 3 0.27639769 4 0.34638098

POPTOTAL= 7547545

FINALWEIGHT=ponderador

/OPTIONS ITERATIONS=50 CONVERGENCE=.0001 DELTA=.5 SHOW=NO.

FRECUENCIES ponderador.

frequencies estrato alc sexo escolaridad grupo_edad .

** TRIMMING .

** _____ .

EXAMINE

VARIABLES=ponderador

/PLOT BOXPLOT

/COMPARE GROUP

/PERCENTILES(5,10,25,50,75,90,95) HAVERAGE

/STATISTICS EXTREME

/MISSING LISTWISE

/NOTOTAL.

** CHECK.

weight off.

frequencies estrato nom_alc grupo_edad escolaridad.

weight by ponderador.

frequencies estrato nom_alc grupo_edad escolaridad.

DELETE VARIABLES uno factor_m tmp.

VARIABLE LABELS ponde_muestral ' Ponderador muestral. Inverso de la probabilidad de selección de la UPM'.