En concordancia con la regulación establecida en el* Reglamento de Elecciones del Instituto Nacional Electoral*, aprobado por el Consejo General de este organismo y publicado en el* Diario Oficial de la Federación *el 13 de septiembre de 2016, se detalla lo siguiente:

1. Objetivos del estudio:

El objetivo central del estudio fue dar seguimiento al estado de la opinión pública de los ciudadanos mexicanos residentes en el territorio nacional.

Este estudio forma parte de una serie de mediciones mediante encuestas nacionales por muestreo realizadas por las empresas investigadoras responsables desde junio de 1998.

Se atendieron los siguientes objetivos particulares:

- Seguir las opiniones de la ciudadanía en torno a la situación prevaleciente en el país.
- Establecer el nivel de aprobación final del Ejecutivo federal y la evaluación de sus principales acciones y programas.
- Conocer las expectativas ante el nuevo gobierno y la actitud frente a las propuestas de reformas constitucionales.
- Medir aspectos relacionados con la cultura política, la percepción de las instituciones democráticas del país y la evaluación del proceso electoral reciente.
- Determinar el conocimiento e imagen general entre la ciudadanía hacia los partidos políticos y las preferencias electorales actuales.
- *De manera complementaria, se añadieron al final de las entrevistas reactivos sobre integridad electoral incorporados por la Facultad Latinoamericana de Ciencias Sociales (México) con el Fondo de Apoyo para la Observación Electoral (FAOE).*

GEA
Grupo de Economistas y Asociados

ISA® Investigaciones Sociales Aplicadas

2. Marco muestral:

Listado de secciones electorales en el territorio nacional con sus datos de identificación geográfica y numérica, el número de electores registrados en el Listado Nominal al corte más próximo a la fecha de toma de datos, del 19 de septiembre de 2024, y los resultados oficiales de las elecciones federales del pasado 2 de junio de 2024 en las mismas secciones electorales, conforme la información disponible en el sitio oficial del Instituto Nacional Electoral.

3. Diseño muestral:

a) Población objetivo:

Personas de nacionalidad mexicana residentes en viviendas particulares ubicadas dentro del territorio nacional que declararon tener18 o más años de edad al momento de la entrevista.

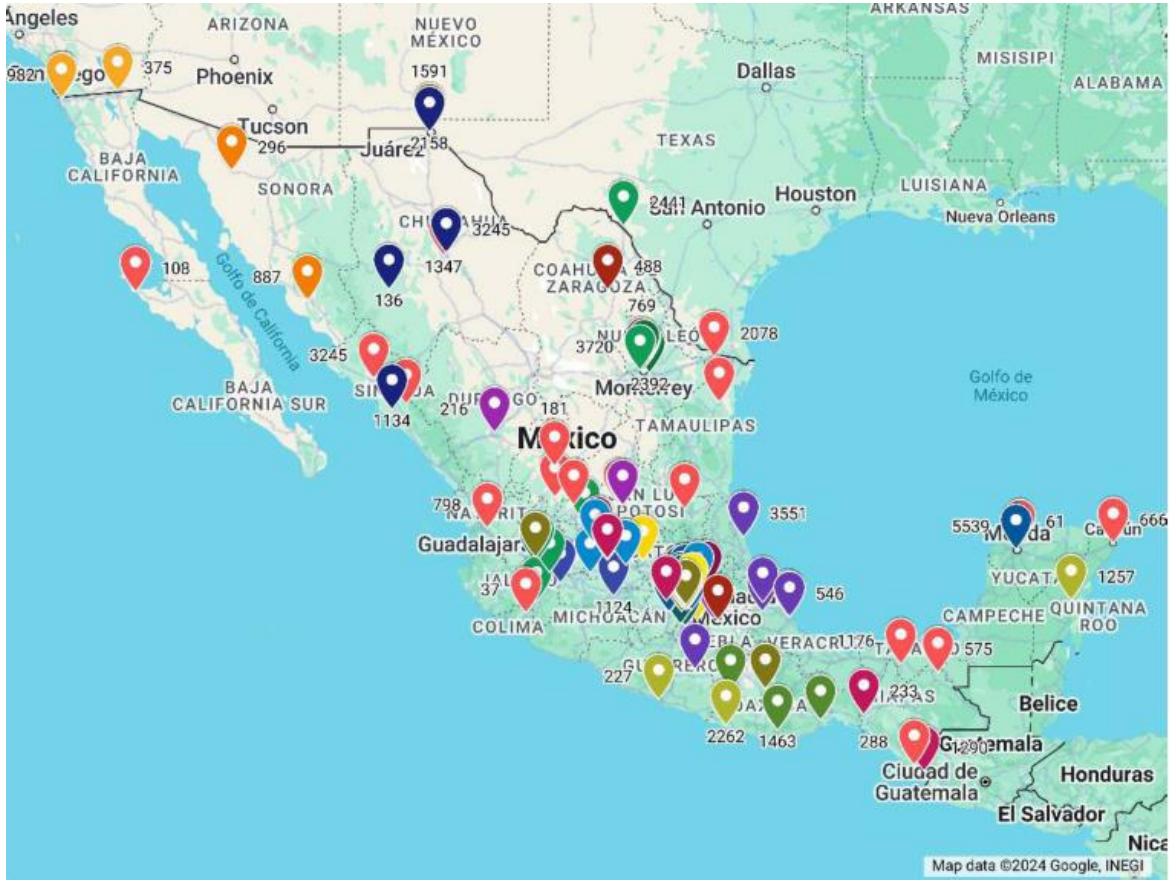
Los resultados de esta encuesta refieren exclusivamente a la población estudiada y sólo tienen validez para expresar la opinión y preferencias electorales de la población objetivo en las fechas específicas del levantamiento de los datos, los que no permiten ni pretenden ser un pronóstico respecto de sucesos posteriores al momento de la medición.

b) Selección de unidades:

Para la encuesta nacional se realizó una selección aleatoria de 107 secciones electorales (unidades primarias de muestreo) con probabilidad proporcional al tamaño (definido por el número de electores registrados en el listado nominal) y con reemplazo.

Se corroboró que las proporciones de voto por opción electoral en las secciones en muestra no se desviaran más de 0.5 por ciento respecto al resultado en las elecciones federales más reciente (en caso contrario, se descartó la muestra y se eligió un nuevo número aleatorio de arranque).

Ubicación geográfica de las secciones electorales en muestra



GEA
Grupo de Economistas y Asociados

ISA® Investigaciones Sociales Aplicadas

Por un desastre natural, tuvieron que ser reemplazadas tres secciones electorales de áreas inundadas, tomándose unidades en la misma entidad federativa y estrato geográfico, pero con acceso viable para poder realizar las entrevistas.

Segmentación de unidades:

Para fines de análisis se registran y emplean dos criterios de post-estratificación de las unidades primarias en muestra:

- **Circunscripción electoral de las unidades.**

Acorde con la división oficial en cinco circunscripciones electorales del territorio nacional, que asume divisiones geográficamente continuas y con una población similar:

1ª Circunscripción: Baja California, Baja California Sur, Chihuahua, Durango, Jalisco, Nayarit, Sinaloa y Sonora (ocho entidades).

2ª Circunscripción: Aguascalientes, Coahuila, Guanajuato, Nuevo León, San Luis Potosí, Tamaulipas y Zacatecas (siete entidades).

3ª Circunscripción: Campeche, Chiapas, Oaxaca, Quintana Roo, Tabasco, Veracruz y Yucatán (siete entidades).

4ª Circunscripción: Ciudad de México, Guerrero, Hidalgo, Morelos, Puebla y Tlaxcala (seis entidades).

5ª Circunscripción: Colima, Estado de México, Michoacán y Querétaro (cuatro entidades).

- **Región en que se ubicaron las unidades.**

La división de unidades primarias de muestreo conforme a este criterio divide a las entidades en cuatro grupos territorialmente continuos y con población similar (cada uno con cerda de una cuarta parte de los electores registrados, dentro de un margen menor al uno por ciento), fue la siguiente:

Región Norte: Baja California, Baja California Sur, Chihuahua, Coahuila, Durango, Nuevo León, Sinaloa, Sonora, Tamaulipas y Zacatecas (diez entidades).

Región Occidente: Aguascalientes, Colima, Guanajuato, Guerrero, Jalisco, Michoacán, Nayarit, Querétaro y San Luis Potosí (nueve entidades).

Región Centro: Ciudad de México, Hidalgo, Estado de México y Morelos (cuatro entidades).

Región Sureste: Campeche, Chiapas, Oaxaca, Puebla, Quintana Roo, Tabasco, Tlaxcala, Veracruz y Yucatán (nueve entidades).

Selección de viviendas:

Selección de diez viviendas por sección, mediante recorrido aleatorio (con arranque a partir del domicilio de la casilla básica en la elección federal pasada e intervalo constante en cada unidad primaria en muestra).

Se intentaron entrevistas completas y válidas en tantas viviendas como se requirió hasta completar los casos efectivos programados para cada sección (es decir: los casos de viviendas deshabitadas, ausentes, donde no residían personas que formaran parte de la población objetivo o donde no se dio respuesta completa a la entrevista fueron consideradas como observadas, pero no como casos válidos, continuando el procedimiento de selección en la siguiente vivienda conforme al intervalo previamente establecido, llevando control de los encuentros en que no se logró la entrevista completa, lo que se reporta para el cálculo de la tasa de rechazo).

Selección del informante: Selección como informante en cada vivienda en muestra de la persona de origen mexicano con 18 o más años de edad que celebrara el próximo cumpleaños de entre las presentes al momento de la entrevista.

Perfil de la población entrevistada (casos ponderados):

DISTRIBUCIONES PARAMÉTRICAS		*DISTRIBUCIONES ESTOCÁSTICAS*	
SEXO		**ESCOLARIDAD**	
Masculino (o no binario)	48.0%	Hasta primaria	8%
Femenino	52.0%	Media básica	38%
EDAD		Media superior	42%
De 18 a 24 años	15.3%	Superior	12%
De 25 a 39 años	32.0%	**OCUPACIÓN**	
De 40 a 54 años	26.5%	Sector público	16%
De 55 y más años	26.2%	Sector privado	17%
REGION		Por su cuenta y agropecuario	13%
Norte	24.4%	Desempleo	4%
Occidente	25.0%	Estudia	11%
Centro	25.3%	Quehaceres domésticos	29%
Sureste	25.3%	Jubilado u otra inactiva	10%

c) Procedimiento de estimación:

Para eliminar deformaciones por la toma de muestra conforme a criterios demográficos, se efectuó un ajuste (post estratificación) para que la muestra reprodujera la distribución por sexo y edad de la ciudadanía conforme lo reportado en el Listado Nominal de Electores al corte más reciente disponible en el sitio oficial del INE.

Por ende, los resultados presentados no son en ningún caso frecuencias simples, sino estimaciones basadas en el empleo de factores de expansión muestral, calculados como el inverso de la probabilidad de selección de cada individuo en muestra, ajustados para hacer acorde la distribución de casos en muestra conforme a sexo y edad con la oficialmente reportada en el Listado Nominal de Electores.

Siguiendo el diseño adoptado, se puede establecer que la probabilidad de selección de un individuo i residente de la sección electoral k es:

$$P_{ki} = \frac{mN_k}{N}\frac{n_k}{N_k} = \frac{mn_k}{N}$$

Dónde:

"P_{ki}" es la probabilidad de selección del individuo $i$de la vivienda j de la sección electoral k;

"N_k" es la población residente en la sección electoral k;

"N" es la población total bajo observación (electores);

"m" es el número de unidades primarias de muestreo (secciones electorales) que se seleccionan para el estudio;

"n_k" es el número de individuos seleccionados en cada sección electoral k.

Entonces, el diseño adoptado correspondería a un diseño de los llamados auto-ponderados, donde la probabilidad de selección de cada individuo en muestra es igual a la de cualquier otro y equivalente al inverso del número de casos efectivos en muestra entre la población total bajo estudio (electores).

Empero, la toma de muestra pudo arrojar diferencias entre la distribución de casos en muestra y la distribución real de la población según variables demográficas de las que se dispone de información. Esto pudiera deberse a varios factores: variaciones en la distribución producto de factores aleatorios; desiguales probabilidades de encuentro de la población según su pertenencia a un determinados segmento demográfico; propensiones distintas a responder la entrevista según segmento demográfico; entre otros.

Por ello es pertinente efectuar un ajuste en el peso de los casos en muestra que permita que la distribución conforme a variables demográficas conocidas y con información actualizada públicamente disponible (región, sexo y edad) reproduzca exactamente la distribución del universo, compensando con ello las diferencias en la selección de unidades en muestra, las tasas de encuentro y propensión a responder de los distintos segmentos demográficos.

Así, cada caso tuvo un peso (w) igual al cociente del número de electores estimados para un estrato de sexo y edad (rse) entre el número de casos observados en la encuesta en el mismo estrato. El factor de expansión muestral para cada caso en un estrato vendría luego dado por:

$$w_{rsg} = \frac{N_{rsg}}{n_{rsg}}$$

Donde

"w" corresponde al valor de expansión de cada caso en muestra;

"r" corresponde a una región geográfica generada para el estudio;

"s" corresponde a un sexo determinado dicotomizado (masculino o no binario, o femenino); y

"g" corresponde a un grupo de edad determinado (agrupado en cuatro estratos: de 18 a 24 años, de 25 a 39 años, de 40 a 54 años y de 55 y más años de edad);

"N"es el número de electores oficialmente reportado; y

"n" al número de casos observados en la muestra tomada.

Con este procedimiento, se dispuso de un total de 32 factores distintos de selección para los 1,070 casos en muestra.

d) Tamaño de la muestra:

1,070 entrevistas efectivas.

La variable relevante a estimar para este ejercicio de medición es la proporción de población con la intención manifiesta de sufragar a favor de alguno de los contendientes en unas próximas elecciones.

El principal componente del error estadístico en una estimación es el producido por la elección de unidades primarias. Siendo así, el tamaño deseable de la muestra de unidades primarias, sin incluir un corrector finito, por intrascendente, se calcula como:

$$m = \frac{{Z_{a/2}}^2 \max {\sigma_i}^2}{{e_m}^2}$$

Donde:

"m" son las unidades primarias en muestra (secciones electorales) a elegir;

"$\boldsymbol{Z_{a/2}}$" corresponde al número de unidades de desviación estándar de la media a la que se estima el error muestral, que en este caso asume un valor de 1.96, para el que el área bajo la curva normal ubicada entre la media y las ordenadas al valor referido a ambos lados es de 0.95, que es el nivel de confianza definido para la estimación; y

"max $({\sigma_i}^2)$" corresponde al valor máximo observado de la varianza por sección electoral de la votación relativa por un partido o coalición "i" para la elección de diputaciones federales más reciente, del 2 de junio de 2024, variable que asume empíricamente un valor de 0.033.

"e_m" es el error estadístico máximo tolerado al nivel del 0.95 de confianza para la muestra de unidades a elegirse en la primera etapa, que se fijó en 0.0345, con lo que rebasa ligeramente la varianza de las proporciones de voto por partido o coalición en las unidades en muestra.

Luego de realizar los cálculos correspondientes y redondeando al alza, se estima que se requiere una muestra de 107 unidades primarias para disponer de estimadores dentro de un margen de error de 0.0345 al 95 por ciento de confianza, que es el nivel de error tolerado para el diseño propuesto. A ello debe sumarse en su momento el error estadístico producido en etapas posteriores por la selección de viviendas y de personas entre los ocupantes en las posteriores etapas del estudio.

Para calcular el tamaño de muestra requerido relativo al número de casos a observar, resulta pertinente considerar la varianza máxima posible e incorporar además el efecto por el diseño polietápico que se esperaría. No se incluiría un corrector finito, por ser irrelevante en el caso. Por lo anterior, se utiliza la siguiente ecuación convencional:

$$n = \frac{{Z_{a/2}}^2 pq\ \delta}{{e_n}^2}$$

Donde

"$Z_{a/2}$" corresponde, como ya se dijo, al número de unidades de desviación estándar de la media a la que se estima el error muestral, que en este caso vuelve a asumir un valor de 1.96 (donde el área bajo la curva normal ubicada entre la media y las ordenadas al valor referido a ambos lados es de 0.95, que es el nivel de confianza definido para la estimación);

"*p*" es la prevalencia esperada del parámetro a evaluar, que en este caso es desconocido y que, por consecuencia, asume el valor que hace mayor el tamaño muestral (0.5);

"*q*" el complemento de "*p*" y, por ende, tiene un valor de 0.5, lo que supone una varianza máxima esperada, "*pq*", de 0.25;

"δ" corresponde al efecto de diseño, estimado en 1.6 como resultado de la elevación del error esperado por la primera etapa para considerar el efecto de etapas subsecuentes de selección;

"e_n" es el error estadístico máximo tolerado al nivel del 0.95 de confianza para la muestra total, el que en este caso se estimó en 0.038, considerando que el efecto de diseño estimado para la primera etapa constituye cerca de noventa por ciento del error estadístico del estudio.

Efectuando los cálculos correspondientes y redondeando al alza, se tiene un requerimiento de 1,070 casos efectivos en muestra para, con el diseño polietápico adoptado, disponer de estimadores dentro de un margen de error tolerado de 0.038 al 95 por ciento de confianza, lo que supone disponer de diez casos efectivos en cada unidad primaria en muestra.

e) Calidad de la estimación:

Error máximo implícito: ±3.8%.

Nivel de confianza: 95%.

En estudios mediante encuesta por muestreo existen errores no muestrales como el fraseo de preguntas, efectos no anticipados de preguntas previas o posteriores a otras preguntas en el cuestionario, y contingencias que se presentan durante el levantamiento de campo que pueden generar sesgos de carácter no sistemático que afecten la calidad de la estimación.

GEA
Grupo de Economistas y Asociados

ISA® Investigaciones Sociales Aplicadas

f) **Frecuencia de no-respuesta:**

En este estudio se indica en cada reactivo aplicado y difundido en que sea relevante las proporciones de no respuesta, diferenciando en cada caso los porcentajes que manifestaron que no privilegiaban alguna opción de respuesta, los que dijeron no saber o se manifestaron indecisos respecto de su opinión o elección y los que se negaron a responder al cuestionamiento específico.

Tratamiento de no-respuesta:

Para este estudio, se consideraron dos tipos de no-respuesta: cuando el entrevistado no contestó el cuestionario (no-respuesta total) y cuando el entrevistado no respondió a alguna de las preguntas del cuestionario (no-respuesta parcial).

Cuando el entrevistado no contestó el cuestionario, se procedió a su sustitución, entrevistándose a una persona en tantas viviendas como se requirió para completar los casos efectivos programados para cada sección. Es decir: los casos de viviendas deshabitadas, ausentes, donde no residían personas que formaran parte de la población objetivo o donde no se dio respuesta completa a la entrevista fueron consideradas como observadas, pero no validadas, por lo que se continuó el recorrido en la sección, seleccionado la siguiente vivienda conforme al intervalo previamente establecido.

Cuando el entrevistado no respondió con alguna de las opciones de respuesta establecidas en alguno de los reactivos del cuestionario, diciendo que no tenía opinión o elección, declarando no saber o negándose a responder, se codificó la respuesta en alguno de dichos rubros, no efectuando imputación alguna de estos casos, sino contabilizándolos por separado en opciones diferenciadas, salvo en el caso de la pregunta sobre preferencia electoral mediante boleta, en que se efectuó el cálculo simple de las proporciones de casos válidos por contendiente, excluyendo los casos en que se anulará y los no válidos.

4. Fecha de recolección: **26 al 29 de septiembre de 2024.**

Duración de la entrevista: **29 minutos.**

Tasa de rechazo general:

Total de intentos de entrevista	=	3,308
Contactos no exitosos	=	976
Contactos no exitosos / intentos	=	30%
Total de personas contactadas	=	2,332
Entrevistas completas	=	1,070
Negativas o abandonos	=	1,262
Negativas o abandonos / contactos	=	54%
Entrevistas no logradas / intentos	=	68%

5. Método de recopilación:

Entrevistas presenciales (cara a cara) en viviendas particulares ubicadas dentro del territorio nacional, con cuestionario en español, estructurado con preguntas primordialmente cerradas que se plantean al informante de manera ordenada, aplicado directamente por entrevistadores asistidos por computadora (CAPI).

Las preguntas sobre intenciones de voto fueron aplicadas recurriendo a imágenes en los dispositivos que simularon boletas impresas a color, las cuales fueron respondidas por los entrevistados de manera directa, para respetar el derecho al secreto del voto y preservar el anonimato y confidencialidad de la respuesta.

Personal involucrado:

Veintisiete entrevistadores con nueve supervisores. La carga de trabajo por persona fue de diez entrevistas cada jornada. Para el control de calidad del trabajo de campo se emplearon diversas técnicas: la supervisión *in situ* de las entrevistas, con un tramo de control de un tercio de los casos en muestra; y el monitoreo en tiempo real de la toma efectiva de muestra en los lugares indicados, acopiándose información sobre las coordenadas geográficas de los puntos específicos de entrevista, dato que guarda la empresa responsable del estudio.

6. Forma de procesamiento:

Las respuestas a los cuestionarios recabados fueron concentradas en el sistema telemático de la empresa encargada de la operación, donde se efectuó un proceso de validación y crítica del correcto llenado de los instrumentos para posibles correcciones inmediatas o posteriores. Las preguntas cerradas fueron directamente codificadas por el personal operativo encargado del levantamiento. Para la codificación de las preguntas abiertas se recurrió a catálogos generados previamente si estaban disponibles o que fueron generados por especialistas cuando así se requirió. Al final la base de datos generada en formato SAV fue revisada y depurada hasta disponer de datos consolidados y consistentes.

A partir de la base de datos definitiva y mediante el uso de un programa estadístico, se generaron las salidas (tablas de frecuencias y cruces) que permitieron el graficado de las respuestas obtenidas en todos y cada uno de los reactivos destinados a su difusión inmediata.

Estimadores:

El estimador para el cálculo de la proporción de cada opción de respuesta dicotómica en cada reactivo aplicado ($\hat{p}_i$) estuvo dado por:

$$\hat{p}_i = \frac{\sum_\theta w_i I_i}{\sum_\theta w_i}$$

Donde

"$\hat{p}_i$" es la proporción estimada de individuos con una característica dada:

"θ" es la población para la que se estima la proporción;

"w_i" son los valores de expansión de los casos en muestra;

"I_i" es la ocurrencia del suceso, que adquiere un valor de 1 si el individuo tiene la característica dada y de 0 si no la tiene.

Intervalos de confianza:

Los intervalos de confianza de cada estimador se calcularon con base en:

$$I_{1-\propto} = \left(\hat{p}_i - Z_{a/2} \sqrt{\frac{\sigma_{\hat{p}}{}^2}{m}} \, , \, \hat{p}_i + Z_{a/2} \sqrt{\frac{\sigma_{\hat{p}}{}^2}{m}} \right)$$

Donde

"$I_{1-\alpha}$" corresponde al intervalo de confianza al 100(1-α) por ciento, siendo en este caso "α" igual a 0.05;

"$Z_{a/2}$" corresponde al número de unidades de desviación estándar de la media a la que se estima el error muestral, que en este caso asume un valor de 1.96;

"$\sigma_{\hat{p}}{}^2$" es la varianza de "$\hat{p}_i$", que cuando las unidades primarias de muestreo son elegidas con probabilidad proporcional a tamaño y las secundarias bien de manera aleatoria o sistemática, puede aproximarse por:

$$\sigma_{\hat{p}}{}^2 = \frac{\sum_{i=1}^{m}(\hat{p}_i - \hat{p}_U)^2}{m-1}$$

Siendo

"m" el número de secciones electorales en muestra, como se indicó anteriormente; y

"$\hat{p}_U$" la proporción para el total de la muestra de una opción de respuesta dicotómica en un reactivo aplicado, estimada como:

$$\hat{p}_U = \sum_{i=1}^{m} \frac{\hat{p}_i}{m}$$