

Nota Metodológica Cuantitativa

1. Introducción

El proyecto de “Evaluación de las Políticas Públicas en la Ciudad de México” tiene por objetivo conocer la percepción de la población en torno a las políticas públicas de vivienda, empleo, ingreso, educación, salud, cultura, medio ambiente, seguridad y movilidad, con la finalidad de producir información estadística representativa de la Ciudad de México sobre las condiciones de vida y el bienestar de los hogares. El interés de conocer la percepción de bienestar, las desigualdades sociales y el desempeño institucional, parte a través de la recuperación económica y social que se desarrolla a partir del contexto de la pandemia por COVID-19 en la Ciudad de México.

Fundamentalmente, el desarrollo de la presente investigación tiene el propósito de generar un diagnóstico preciso que permita identificar los principales problemas sobre el bienestar en los hogares capitalinos y contribuya a la generación de mejores programas y políticas públicas futuras.

2. Metodología cuantitativa

Para llevar a cabo la investigación, diversas metodologías cualitativas y cuantitativas han sido propuestas tanto para la recopilación de la información como para el análisis de los datos recabados. En cuanto a metodología cuantitativa, se propone como primer paso el levantamiento de una encuesta en vivienda a 2,400 individuos residentes en la Ciudad de México, mismos que serán seleccionados mediante una muestra aleatoria probabilística. Dicho muestreo tiene un diseño estratificado con conglomerados, selección multietápica y con probabilidad proporcional al tamaño de una característica de interés.

Una vez que la muestra sea recopilada y almacenada en una base de datos, se realizará un análisis de frecuencia de las respuestas obtenidas. Este análisis se realiza tanto de manera general como con agregaciones demográficas en tablas cruzadas. Para ambos tipos de análisis se construirán diversos estadísticos (media, mediana, coeficiente de variación, intervalos de confianza, etc.), a partir de los factores de expansión diseñados para cada una de las preguntas del cuestionario.

Una vez que todas las preguntas sean analizadas, marginal y conjuntamente, mediante las características sociales y demográficas seleccionadas, se procederá con la construcción del índice de percepción de bienestar social. El índice que este proyecto propone parte de la aplicación de técnicas multivariadas que permitan asociar características geoespaciales con las respuestas vertidas por cada uno de los participantes de la encuesta, referentes a los temas de vivienda, empleo, ingreso, educación, salud, cultura, medio ambiente, seguridad y movilidad.

En los siguientes pasos, se detalla el proceso metodológico que se llevará a cabo en cada una de las etapas mencionadas anteriormente.

3. Población objetivo

El levantamiento de la encuesta está dirigido a la muestra representativa de la población de 15 años y más que resida en la Ciudad de México en las primeras dos semanas del mes de octubre de 2021, misma temporalidad comprendida para la aplicación de los cuestionarios.

Debido a que el diseño de muestreo es probabilístico, la cantidad de individuos de cada sexo y grupo de edad en la muestra es aleatorio, de manera que la muestra representa a la población en general y no representa a un grupo en específico.

4. Marco muestral

El marco de muestreo se construye a través del Censo de Población y Vivienda 2020, realizado por el Instituto Nacional de Estadística y Geografía (2020), y con base en la georreferenciación de los polígonos de las manzanas del INEGI, tanto urbanas como rurales.

5. Tamaño de muestra

El tamaño de muestra calculado fue de 2,400 casos. Este tamaño de muestra se distribuye entre los distintos estratos a fin de generar una disminución en los márgenes de error estadísticos. Para determinar el tamaño de muestra se consideró la siguiente fórmula estadística:

$$n_0 = \frac{z_{1-\alpha/2}^2 p(1-p)}{r^2} ; \quad n = \frac{n_0}{1 + \frac{n_0}{N}} \frac{Def f}{(1-TNR)}$$

Donde:

n = Es el tamaño de muestra final

$z_{1-\alpha/2}^2$ = Es el percentil $\alpha/2$ de la distribución normal estándar que permite significancia de α en los resultados.

p = Probabilidad o proporción esperada de observar una categoría particular. El escenario más conservador para el cálculo de muestra es fijar este parámetro en $p = 0.5$.

N = Es el tamaño de la población objetivo.

r = Es el error muestral máximo por tolerar.

TNR = Se refiere a la tasa de no respuesta esperada

$Def f$ = Es el efecto de diseño al utilizar uno distinto al aleatorio simple.

Considerando un nivel de confianza $(1-\alpha)$ de 95%, error máximo a tolerar (r) de 2.5%, proporción (p) de 0.5, efecto de diseño ($Def f$) de 1.4 y tasa de no respuesta (TNR) del 10%, se obtiene un tamaño de muestra de 2 400 unidades muestrales.

6. Estratificación

Para obtener resultados con alta precisión y baja variabilidad, se propone un diseño de muestreo estratificado. La variable de estratificación propuesta corresponde al grado de Desarrollo Social calculado y entregado por EVALUA. En cada uno de los estratos propuestos se hará la extracción de una muestra parcial, la cual será proporcional al número de viviendas particulares habitadas en cada estrato. Tanto la selección como el análisis de resultados de la muestra en cada uno de los estratos se realizan de manera independiente.

El método de distribución de la muestra se realiza de acuerdo con la afijación proporcional, la cual se muestra en la siguiente expresión matemática:

$$n_h = \frac{N_h}{N}n$$

Donde:

n_h = Tamaño de muestra en el h -ésimo estrato.

N_h = Total de viviendas particulares habitadas en el h -ésimo estrato.

N = Total de viviendas particulares habitadas.

n = Tamaño de muestra.

7. Esquema de selección

El esquema de muestreo propuesto corresponde a un multietápico, estratificado y por conglomerados. La Ciudad de México será dividida en estratos geográficos. Posteriormente, la selección de las unidades de muestreo se realizará a través de etapas sucesivas y de manera independiente para cada uno de los estratos generados:

- Unidades primarias de muestreo (Manzanas): En cada uno de los h estratos se realiza la extracción de m manzanas. La selección se realiza sin reemplazo y con probabilidad proporcional al número de viviendas.
- Unidades secundarias de muestreo (Viviendas): En cada una de las manzanas seleccionadas en la etapa anterior, se realiza un muestreo sistemático con arranque aleatorio para elegir 5 viviendas.
- Unidades últimas de muestreo (Individuos): En cada vivienda se selecciona 1 individuo de entre el total de personas que pertenezcan a la población objetivo. El método de selección corresponde a la selección de la persona con el cumpleaños más próximo.

8. Cálculo de factores de expansión

Los factores de expansión permiten conocer el número de elementos a los que representa cada una de las unidades seleccionadas en las etapas que forman parte del esquema de

muestreo. Para obtener los factores finales con los que se realiza la ponderación para el cálculo de estimadores, se usa el inverso de la probabilidad de selección de un individuo, la cual puede estimarse con la siguiente fórmula estadística:

$$P[X_{hij}] = \left(\frac{N_{hi} m_h}{N_h} \right) \left(\frac{5}{N_{hi}} \right) \left(\frac{1}{p_{hij}} \right) = \frac{5m_h}{N_h p_{hij}}$$

Donde:

$P[X_{hij}]$ = Es la probabilidad de selección de un individuo perteneciente a la población objetivo en el *h-ésimo* estrato, la *i-ésima* manzana y *j-ésima* vivienda seleccionada.

N_{hi} = Número de viviendas en la *i-ésima* manzana del *h-ésimo* estrato.

N_h = Número de viviendas en el *h-ésimo* estrato.

m_h = Número de manzanas a seleccionar en el *h-ésimo* estrato

p_{hij} = Población objetivo en la *j-ésima* vivienda seleccionada dentro de la *i-ésima* manzana del *h-ésimo* estrato.

A través del cálculo de la probabilidad de selección del individuo en muestra, se calcula el factor de expansión, mismo que se calcula con el inverso de probabilidad de selección:

$$F[X_{hij}] = \frac{1}{P[X_{hij}]}$$

Donde:

$F[X_{hij}]$ = Es el factor de expansión asociado al individuo seleccionado en la *j-ésima* vivienda, *i-ésima* manzana y *h-ésimo* estrato.

Con este factor de expansión se realiza el cálculo ponderado de las estadísticas de interés para cada una de las preguntas del cuestionario.

9. Cálculo de estimadores

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

a) Estimaciones puntuales

Para la estimación puntual de la población total perteneciente a cada uno de los reactivos del cuestionario se hizo uso de los siguientes cálculos estadísticos:

$$\hat{T} = \sum_{h=1}^H \sum_{i=1}^I \sum_{j=1}^5 X_{hij} F_{hij}$$

Donde:

\hat{T} = Es la estimación de la población total

X_{hij} = Es la característica observada del individuo seleccionado en la j-ésima vivienda de la i-ésima manzana del h-ésimo estrato.

F_{hij} = Es el factor de expansión corregido correspondiente al individuo seleccionado en la j-ésima vivienda de la i-ésima manzana en el h-ésimo estrato.

Para el cálculo de proporciones y promedios, se utilizó el estimador de razón:

$$\hat{R} = \frac{\hat{T}}{\hat{Y}}$$

Donde:

\hat{R} = Es la estimación de la proporción o estimador de razón de la población total perteneciente a alguna categoría.

$$\hat{Y} = \sum_{h=1}^H \sum_{i=1}^I \sum_{j=1}^5 F_{hij}$$

b) Medidas de dispersión

Las estimaciones del error estándar (E.E.), coeficiente de variación o error relativo del estimador (C.V.) y el efecto de diseño (DEFF) se calculan mediante las siguientes expresiones:

$$E.E. = \sqrt{\hat{V}(\hat{e})} \quad C.V. = \frac{E.E.}{\hat{e}} \quad DEFF = \frac{\hat{V}(\hat{e})}{\hat{V}(\hat{e})_{mas}}$$

Donde:

\hat{e} = estimador de parámetro poblacional

\hat{V} = estimador de la varianza bajo repeticiones Bootstrap

$\hat{V}(\hat{e})_{mas}$ = estimador de varianza, bajo un muestreo aleatorio simple

10. Análisis socioespacial de las condiciones de vida y bienestar de los hogares de la CDMX

Como parte de la estrategia metodológica, este proyecto recurrirá al pensamiento espacial para analizar las condiciones de vida y bienestar de los hogares de la Ciudad de México. En términos generales, el análisis espacial consiste en examinar la información de los componentes de los subsistemas del espacio geográfico (social, económico, cultural, político,

ambiental) a partir de su localización, distribución, distancia y situación en relación e interacción con otros elementos que, en su conjunto, constituyen un sistema complejo —construido en el tiempo— de acciones y objetos (Gómez y Gómez, 2013; Massiris, 2002, 2012; Santos, 2000).

Acorde con lo anterior y con base en el índice que resulte de este proyecto, en este trabajo el análisis socioespacial tiene como objetivo tipificar unidades territoriales de la Ciudad de México a nivel de Área Geoestadística Básica (AGEB) para la identificación de áreas críticas o vulnerables, cuya atención resulta estratégica para implementar mecanismos de apoyo para contrarrestar la marginación, la pobreza y la desigualdad de la población, recientemente acentuadas por la pandemia de la COVID 19.

Los Sistemas de Información Geográfica (SIG) han contribuido al análisis espacial con la automatización de la información georreferenciada, a través de softwares especializados que posibilitan su procesamiento, visualización y análisis (Buzai y Baxendale, 2015). Si bien, los SIG se caracterizan por la elaboración de mapas temáticos, esta investigación aprovechará su potencialidad como una herramienta de cálculos espaciales (Buzai y Baxendale, 2015; Quintero, 2009), que permitirá considerar las características de las condiciones de vivienda y bienestar de los hogares capitalinos a partir de su situación geográfica. Para complementar lo anterior, también se incluirá la perspectiva de los Sistemas de Información Geográfica Cualitativos (SIG Cualitativos) para incorporar la experiencia, opinión y percepción de los habitantes de los territorios al análisis socioespacial, a partir del diálogo con herramientas metodológicas de orden cualitativo, como es el caso de los grupos de enfoque (Teixeira, 2018).

El análisis espacial requiere de la elaboración de capas de información geográfica (*shapefiles*), las cuales contienen la georreferenciación y los datos —cualitativos y cuantitativos— de cada atributo en cuestión (Buzai y Baxendale, 2015; Quintero, 2009; Teixeira, 2018).

Para esta investigación, el análisis socioespacial consistirá, en una primera fase, en la construcción de capas de información geográfica de cada uno de los indicadores (vivienda, empleo, ingreso, educación, salud, cultura, medio ambiente y movilidad), así como del índice propuesto, a partir de los resultados, cuantitativos y cualitativos, de la “Encuesta sobre bienestar de la Ciudad de México 2021”. Este paso permitirá conocer la accesibilidad o distribución de cada uno de los componentes del índice en cuestión. Por ejemplo, midiendo la distancia de la población o de los hogares a los espacios públicos, estaciones de transporte público, hospitales o escuelas, o bien, con la visualización de la densidad de delitos o de la disparidad espacial de ingresos y niveles de educación.

En una segunda etapa, con base en el tal índice, se constituirá una capa de información geográfica que nos permita analizar el comportamiento territorial de los procesos de marginación social y habitacional en la ciudad. De acuerdo con la espacialidad de nuestro índice será posible conocer y delimitar zonas con mayor, mediana y menor accesibilidad de los hogares a condiciones de vida dignas y bienestar, a fin de caracterizar y diagnosticar unidades territoriales estratégicas de atención.

En un tercer momento, las experiencias, reflexiones y propuestas de los habitantes que se obtendrán de los 27 grupos de enfoque, nos ofrecerán elementos para profundizar en las características y en la situación de las unidades territoriales. En función de la convergencia espacializada de la información cuantitativa y cualitativa (también representada en mapas temáticos), será posible realizar el análisis estratégico para la priorización y atención de la población más vulnerable y sin acceso a los servicios sociales básicos de las unidades territoriales; tales insumos servirán como una herramienta útil para proponer políticas públicas y programas diseñados con un enfoque de derechos, acorde con las dinámicas socioeconómicas y espaciales de los hogares de la Ciudad de México.