

Diseño de Muestra BIC Teusaquillo

Dirección Observatorio de Cultura y Gestión del Conocimiento

2023-05-29

Table of Contents

INTRODUCCIÓN	2
OBJETIVO	2
PROPUESTA TÉCNICA	2
CONSIDERACIONES EN EL DISEÑO METODOLÓGICO	2
UNIVERSO y POBLACIÓN OBJETIVO	2
COBERTURA Y DESAGREGACIÓN GEOGRÁFICA	3
UNIDADES ESTADÍSTICAS	3
DISEÑO MUESTRAL Y TAMAÑO DE MUESTRA	5
Marco muestral	5
Tipo de muestreo	6
Tamaño de muestra	9
Selección de la muestra	10
Sobre muestra	11
Precisión deseada de las estimaciones	12
Estimadores y factores de expansión:	13

INTRODUCCIÓN

La línea de activación 4 sobre el Fomento de la actividad residencial para la sostenibilidad de los entornos patrimoniales, está orientada a fomentar la actividad residencial en los entornos patrimoniales, entendida como un factor de vitalidad en cuanto garantiza no solo la presencia de habitantes en jornadas nocturnas y fines de semana, sino que favorece el desarrollo y permanencia de actividades cotidianas tradicionales y servicios próximos de soporte a la vivienda.

En este contexto del Fomento de la actividad residencial para la sostenibilidad de los entornos patrimoniales, para el IDPC es primordial conocer el tipo de población que habita los diferentes entornos patrimoniales priorizados y poder contar con una instrumento cuantitativo que permita tomar decisiones con respecto a las acciones que se realizan en este entorno, basadas en información rigurosamente recolectada y con metodologías estadísticas robustas.

Los resultados de este instrumento ayudarán a realizar las líneas base para la medición de impactos posteriores y la creación de indicadores que den las pautas necesarias para la constante evaluación de los procesos y para realizar el seguimiento de cada uno de los proyectos que se ejecutan en los diferentes entornos patrimoniales por parte del IDPC.

OBJETIVO

El objetivo de esta encuesta es caracterizar los predios de acuerdo a su uso del suelo en las categorías de comercio, residenciales y servicios en la localidad de Teusaquillo.

La encuesta se llevará a cabo mediante un cuestionario en línea, el cual será distribuido a una muestra representativa de predios seleccionados aleatoriamente de los distintos barrios de la localidad. Se espera que los resultados obtenidos a través de esta encuesta permitan diseñar estrategias o programas que beneficien a la localidad en general.

PROPUESTA TÉCNICA

Para la implementación de esta encuesta, se dispone de aproximadamente XX encuestas presenciales en diversos predios. Los resultados obtenidos a través de la caracterización de estos predios proporcionarán información valiosa a la Secretaría Distrital de Cultura, Recreación y Deporte, con el objetivo de desarrollar estrategias específicas en el territorio.

El formulario se aplicará de forma presencial, y la información recopilada se utilizará para realizar análisis geoespaciales, multivariados y correlaciones. Estos

análisis nos permitirán identificar los aspectos clave que sustentarán la construcción de la estrategia.

El diseño de la muestra y el operativo de campo se establecen de manera espacial y se correlacionan con variables de caracterización, con el propósito de identificar los lugares potenciales para la implementación de la estrategia.

CONSIDERACIONES EN EL DISEÑO METODOLÓGICO

Tal como se especifica en el manual de campo, se deben hacer máximo 3 visitas a los predios para recolectar la totalidad de la información de la misma fuente. Si definitivamente no se logra contactar el informante idóneo del predio seleccionado aleatoriamente, se puede reemplazar dentro de la misma manzana idealmente, en caso de no poder hacerse el reemplazo dentro de la misma manzana, el reemplazamiento se hará seleccionando en una de las manzanas de la muestra adicional que se tiene lista para hacer este proceso.

UNIVERSO y POBLACIÓN OBJETIVO

La Encuesta se aplicará a predios con uso de comercio, residencial y servicios ubicados en la localidad de Teusaquillo

COBERTURA Y DESAGREGACIÓN GEOGRÁFICA

La desagregación geográfica se hace por localidad, en esta se encuentran la localidad de La Candelaria y Santa Fe

Mapa de Teusaquillo

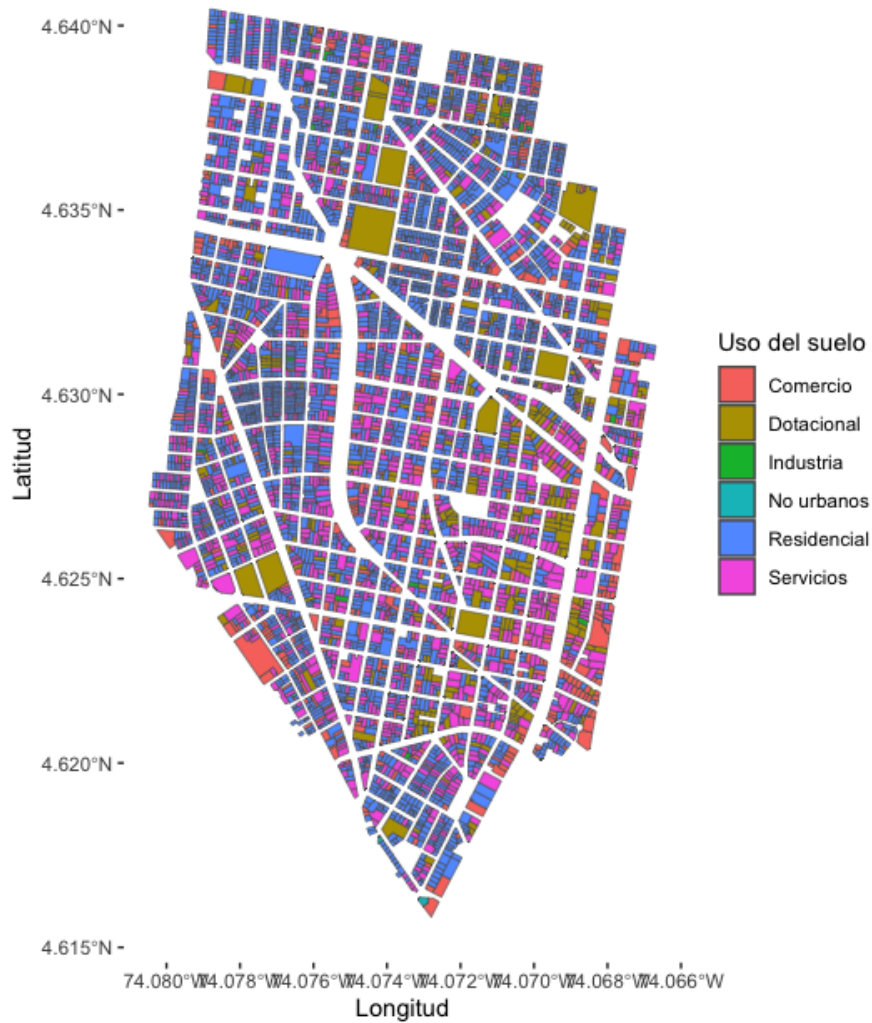


UNIDADES ESTADÍSTICAS

Unidades de Observación y Análisis

Las unidades de observación y análisis son los predios con el uso del suelo como delimitados en la zona de medición anteriormente mencionado

Mapa de predios por uso del suelo



La distribución de la localidad por uso del suelo se tiene por la siguiente tabla:

Distribución de predios por uso del suelo

	Predios	Porcentaje (%)
Comercio	677	11.2
Dotacional	365	6.1
Industria	24	0.4
No urbanos	5	0.1
Residencial	3531	58.5
Servicios	1430	23.7

	Predios	Porcentaje (%)
totales	6032	100.0

El uso del suelo en las categorías de comercio, residencial y servicios constituye el 93.4% de los predios en la localidad. Dentro de estos usos, se observa que los predios residenciales representan la mayor frecuencia, con un 58.5%, seguidos de los predios de servicios con un 23.1%, y los predios de comercio con un 11.2%.

Unidades de Muestreo

El diseño de la muestra probabilístico estratificado de conglomerados multi etapico con selección sin reemplazamiento, por lo cual se tiene una unidad muestral en cada una de las etapas:

- **Primera etapa:** La unidad de muestreo son las manzanas según el uso del suelo.
- **Segunda etapa:** La unidad de muestreo son los predios dentro de las manzanas pre-seleccionadas
- **Tercera etapa:** La unidad de muestreo es el informante idóneo dentro de las según el predio seleccionado:
 - **Hogares:** La unidad de muestreo son los predios con uso del suelo residencial.
 - **Establecimiento comercial:** La unidad de muestreo son los predios con uso del suelo comercial.
 - **Establecimiento servicio:** La unidad de muestreo son los predios con uso del suelo servicios.

DISEÑO MUESTRAL Y TAMAÑO DE MUESTRA

De acuerdo a las características del estudio y las falencias en el marco estadístico se estableció la construcción de un diseño de la muestra probabilístico estratificado de conglomerados multietápico con selección sin reemplazamiento, donde en la primera etapa se hace una selección mediante un diseño MAS estratificado y en la segunda un diseño MAS estratificado según el uso del suelo

Este tipo de muestreo permite ahorrar costos al recoger información básica de la totalidad de elementos de una muestra lo suficientemente grande, recolectando información de una o más variables auxiliares desde la muestra tomada en una primera etapa y estudiar la o las variables de interés en una muestra más pequeña teniendo como marco de selección la obtenida en la primera etapa.

Marco muestral

El marco muestral o estadístico es el mecanismo o material que permite delimitar o identificar en forma apropiada los elementos de una población. Cabe anotar que para la aplicación de las encuestas no se cuenta con un marco actualizado de predios, pues no existe un archivo o listado completo con el nombre y variables de identificación de cada uno.

Por tal razón es necesario contar con un diseño muestral robusto donde se pueda construir un marco estadístico para la selección de predios de acuerdo a los objetivos del estudio y utilizando la información disponible, como es el caso de las bases cartográficas.

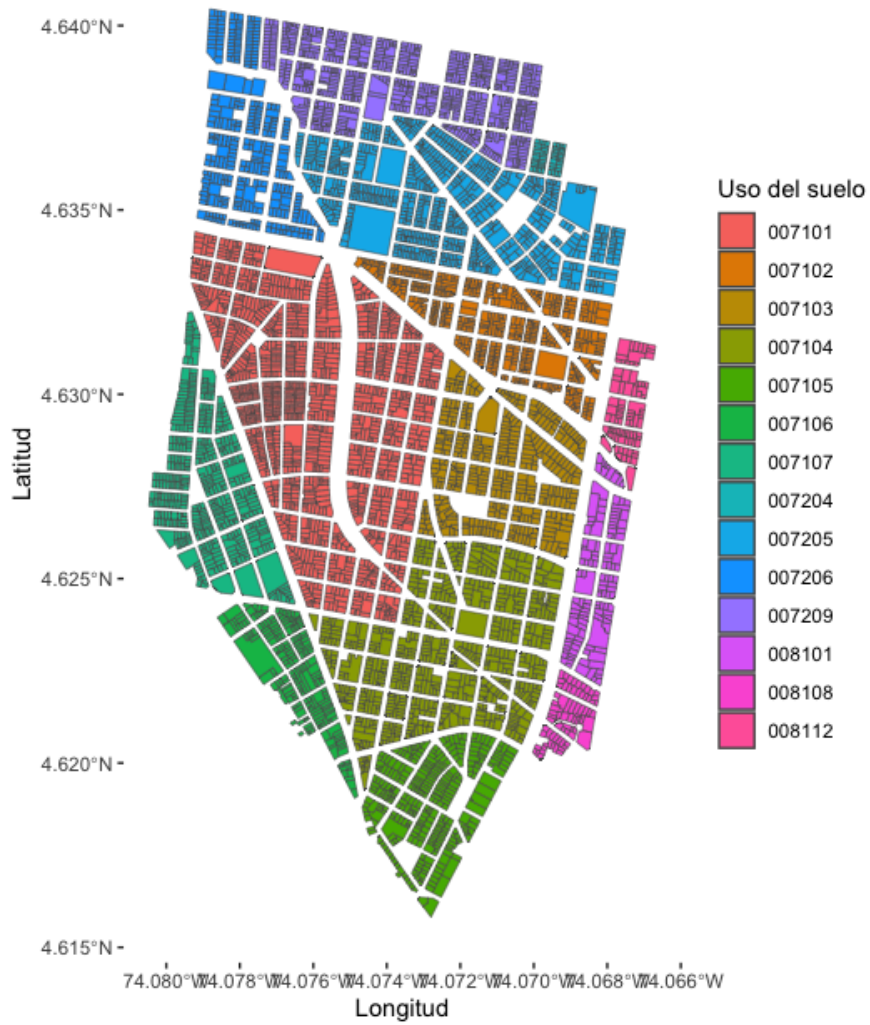
Para la primera etapa el marco muestral que se utiliza es el marco de manzanas de Bogotá, Esta información es suministrada por el Departamento Administrativo de Catastro Distrital - DACD actualizada al año inmediatamente anterior al de aplicación de la encuesta.

Tipo de muestreo

El diseño muestral aplicado en cada etapa de la encuesta, se eligió de acuerdo al marco estadístico con el que se cuenta. Se estructuró un diseño muestral probabilístico en 2 etapas, en la primera etapa se aplica un diseño estratificado aleatorio simple y en la segunda un muestreo sistemático.

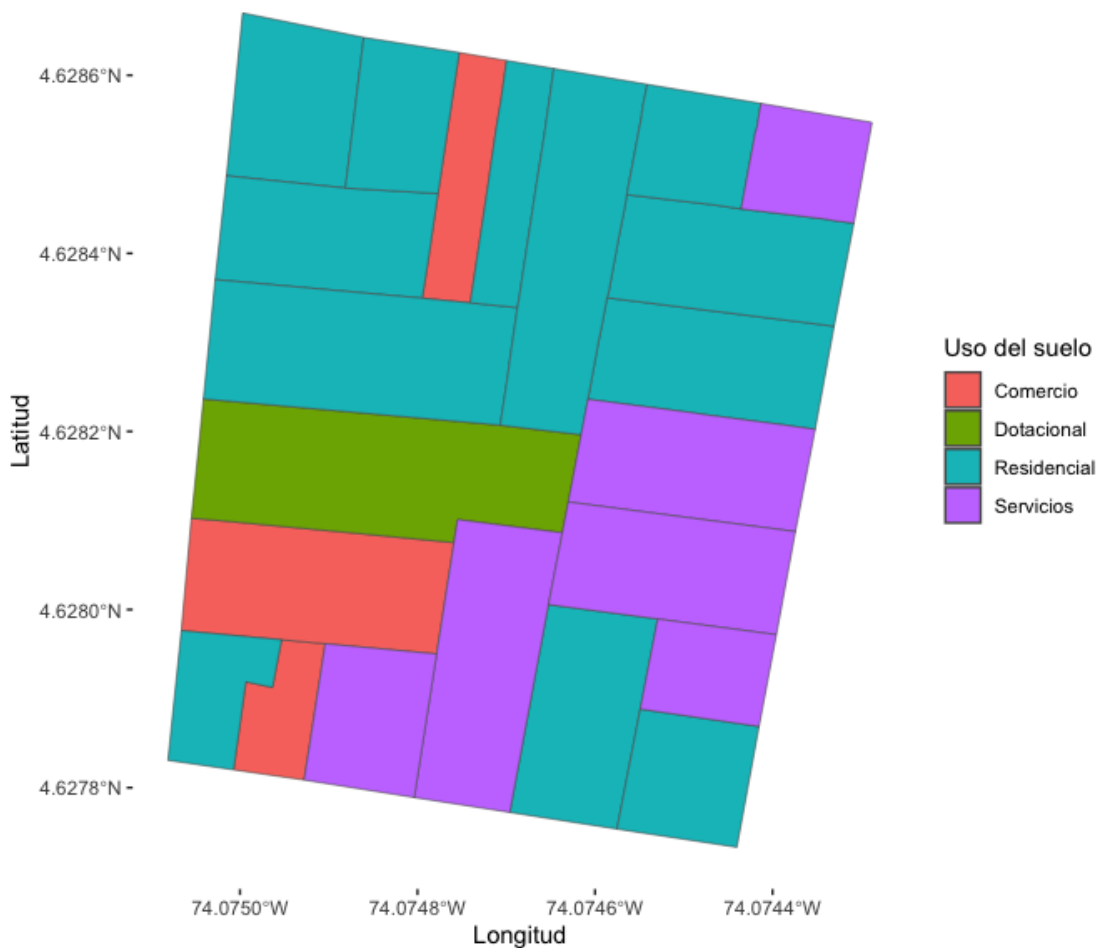
- **Probabilístico:** las unidades a observar se seleccionan mediante un esquema probabilista, lo cual cada unidad de muestreo tiene una probabilidad conocida y mayor a cero, es decir cada unidad de muestreo puede ser seleccionada en la muestra.
- **Conglomerados:** en el contexto de esta encuesta, se definen como los segmentos muestrales conformados por manzanas y predios.
- **Estratificado:** Los estratos se define como la clasificación de información, que permite desagregar la información controlando un margen de error determinado. para esta encuesta se tiene:
 - **Barrio:** las manzanas se segmenta por los barrios catastrales, para mantener una distribución de toda la localidad.

Mapa de Barrios



- ****Uso del suelo:**** las manzanas previamente seleccionadas se estratifica cada manzana por el uso del suelo residenciales, comerciales y de servicios.

Mapa de ejemplo de manzana 007101023



- **Multietapico:** para muestrear los predios se determinan las siguientes etapas
 - **Primera Etapa:** Se seleccionan las manzanas definidas como unidades primarias de muestreo (UPM) en cada barrio, utilizando un diseño π PT que considera la probabilidad de selección basada en la cantidad de predios.
 - **Segunda Etapa :** Se seleccionan los predios definidos como unidades secundarias de muestreo (USM), se selecciona los predios de las manzanas pre-seleccionadas utilizando un muestreo aleatorio simple estratificado por el uso del suelo.
 - **Tercera Etapa:** Las unidades terciarias de muestreo (UTM) se definen según el uso del suelo:

- **Hogares:** La unidad de muestreo son los predios con uso del suelo residencial.
- **Establecimiento comercial:** La unidad de muestreo son los predios con uso del suelo comercial.
- **Establecimiento servicio:** La unidad de muestreo son los predios con uso del suelo servicios.

La selección de los UTM se hace por un muestreo aleatorio simple.

Tamaño de muestra

El cálculo del tamaño de muestra se calcula con la siguiente fórmula:

$$n \geq \frac{Nz_{1-\alpha}^2 P(1-P)Def}{N(ES_p^2) + z_{1-\alpha}^2 P(1-P)}$$

Para,

- N : Número de predios y manzanas total
- n : Tamaño de la muestra de predios y manzanas
- P : Proporción a estimar, para la muestra se especifica $P = 0.5$.
- $z_{1-\alpha}$: Valor de la distribución normal estándar asociado al nivel de confianza $(1 - \alpha)$ deseado, para este estudio es $\alpha = 5$
- ES_p : Error máximo admisible
- Def : efecto de diseño del diseño de conglomerados frente al muestreo aleatorio simple.

La muestra se distribuye de manera proporcional por los barrios. la siguiente tabla muestra los resultados del tamaño de muestra:

Tamaño de muestra de manzanas y predios

Codigo del Barrio	Universo Manzanas	Universo Predios	Muestra de manzanas	Muestra de predios	Predios por manzana
007101	65	1355	15	135	9.0
007102	34	527	8	52	6.5
007103	26	395	6	39	6.5
007104	48	652	11	65	5.9
007105	18	325	4	32	8.0
007106	13	215	3	21	7.0

Codigo del Barrio	Universo Manzanas	Universo Predios	Muestra de manzanas	Muestra de predios	Predios por manzana
007107	24	454	5	45	9.0
007204	2	51	2	5	2.5
007205	45	741	10	74	7.4
007206	18	339	4	34	8.5
007209	28	692	6	69	11.5
008101	9	127	2	13	6.5
008108	4	72	2	7	3.5
008112	8	87	2	9	4.5
Total	342	6032	80	600	7.5

Selección de la muestra

Se selecciona las manzanas de la localidad de teusaquillo por medio de un muestreo π PT, según el tamaño de muestra calculado anteriormente

Muestra de manzanas



Sobre muestra

Se selecciona una sobremuestra del 20%, se puede observar en el siguiente gráfico:

Muestra de manzanas



Precisión deseada de las estimaciones

La precisión de las estimaciones se mide en términos del Coeficiente de Variación (CV). De acuerdo a los procesos de cálculo del tamaño de muestra que se han venido aplicando desde la EBC 2011, se obtienen coeficientes de variación de menos del 2% para la estimación de parámetros, cuya ocurrencia tenga una frecuencia del 15% o más. La precisión de los CV se define de acuerdo a las siguientes consideraciones (Bautista, 1998)¹:

¹ Bautista, L. (1998). Diseños de muestreo estadístico. Bogotá: Universidad Nacional de Colombia

Calificación de la precisión según coeficiente de variación

Valor del CV(R estimado)

%	Calificación de la precisión
Menor a 2%	Excelente
Entre 2% y 4%	Buena
Entre 4% y 6%	Moderada
Entre 6% y 10%	Baja
Entre 10% y 15%	Para usar sólo con mucho cuidado
Superior a 15%	No se puede publicar información

Estimadores y factores de expansión:

El factor de expansión es un valor que multiplica a cada uno de los elementos seleccionados en la muestra con el fin de obtener la estimación del parámetro en el universo. Este factor depende y se elabora a partir del método de selección utilizado.

Los factores de expansión que se necesitan en este caso, se calculan con las estimaciones de personas por manzana y por perfiles, definidos con los muestreos de las etapas anteriores.

Sabemos que para un diseño MAS el factor de expansión de predios $\pi_{Res|Est}$ viene dado por:

$$\pi_{predio|Barrio} = \pi_{Man} \pi_{predio|Man}$$

para:

$$\pi_{Man} = \frac{n_{predio}}{\sum n_{predio}} n_{mzn}$$

y

$$\pi_{predio|Man} = \frac{n_{predio|Man}}{N_{predio|Man}}$$

el estimador está definido como:

$$t_{y,\pi} = \sum_{Est=1}^{Barrios} t_{y,Barrio} = \sum_{k=1}^{n_{predios}} \frac{1}{\pi_{predio|Barrio}} y_k$$

La varianza se define como:

$$\hat{Var}_{t_{y,\pi}} = \sum_{k=1}^{n_{predio}} \frac{N^2}{n_{predio}} \left(1 - \frac{n_{predio}}{N_{predoi}} \right) S_{y,\pi}^2$$