Documento Técnico: Diseño Muestral y Tamaño de Muestra - Encuesta de Satisfacción SDS 2025

Introducción y Objetivos

El presente documento expone la propuesta metodológica para el diseño del muestreo de la **Encuesta de Satisfacción de Usuarios del Sistema Distrital de Salud de Bogotá** correspondiente al año 2025. El propósito es estructurar un esquema de muestreo **probabilístico, estratificado y robusto**, que garantice la **representatividad estadística** de los resultados en los principales dominios de análisis y permita realizar **inferencias válidas** sobre la población usuaria del sistema.

El estudio se enfocará en caracterizar la experiencia de atención en la ciudad de Bogotá, considerando como dominios clave las **Subredes Integradas de Servicios de Salud** (Norte, Sur, Centro Oriente y Sur Occidente) y **Capital Salud**, cruzadas por los **canales de atención presencial, telefónico y virtual**, dentro de las 20 localidades del Distrito Capital.

El análisis contemplará diferentes tipos de servicios, incluyendo tanto **trámites específicos** (como agendamiento de citas y entrega de medicamentos), como la **gestión de Peticiones, Quejas, Reclamos y Sugerencias (PQRS)**. Además, se tendrá en cuenta el **perfil sociodemográfico** de los usuarios, segmentado por **grupo etario y sexo**.

Este diseño se apoya en criterios técnicos establecidos y utiliza como insumo principal la información recolectada en la **Encuesta de Satisfacción** 2023, lo cual permite optimizar la eficiencia muestral y asegurar la **comparabilidad histórica** de los resultados.

Diseño Muestral

Se propone la implementación de un **diseño de muestreo probabilístico, estratificado y polietápico**, con aplicación diferenciada según el canal de atención. Esta estructura permite **optimizar la distribución muestral**, **garantizar la representatividad de los subgrupos de análisis** y **controlar la precisión de las estimaciones** dentro de cada dominio definido.

Marco Muestral

El **marco muestral** estará conformado por los **registros administrativos de atenciones y trámites** realizados en un periodo de referencia reciente —preferiblemente entre los últimos **3 a 6 meses**— por todas las entidades participantes. Este marco estará basado en **eventos de servicio o atención**, no en usuarios individuales, lo cual implica que una persona

puede estar representada en múltiples registros si accedió a más de un servicio durante el periodo analizado.

Unidades de Muestreo

Las unidades muestrales se definen de acuerdo con el canal de atención y se estructuran en tres niveles:

- **Unidad Primaria de Muestreo (UPM):**
  - *Canal presencial:* Corresponde a los **puntos de atención física**, como sedes de Subredes, CADES y otros centros de servicio.
  - *Canal telefónico y virtual:* Se refiere al **evento de atención o trámite registrado** en las bases administrativas.
- **Unidad Secundaria de Muestreo (USM):**
  - *Canal presencial:* Son los **usuarios que acuden físicamente** a los puntos de atención seleccionados, en días y franjas horarias específicas. La selección se realizará mediante **muestreo sistemático en sitio**.
  - *Canal telefónico y virtual:* Se seleccionará aleatoriamente **el registro de atención o trámite**, que representa una interacción con el sistema.
- **Unidad de Observación:**
  Es el **usuario final** que recibió el servicio o ejecutó el trámite, independientemente del canal por el cual accedió.

Estratificación

La **estratificación** consiste en dividir el marco muestral en **subgrupos homogéneos y mutuamente excluyentes**, denominados **estratos**, con el fin de reducir el error de muestreo y asegurar una adecuada cobertura de todos los dominios analíticos relevantes.

Se propone una **estratificación primaria cruzada**, combinando la **Entidad/Subred** con el **Canal de Atención**, lo cual da lugar a estratos como:

- *"Subred Norte - Canal Presencial"*,
- *"Capital Salud - Canal Telefónico"*, entre otros.

Este enfoque permite un mayor control sobre los principales dominios de interés y garantiza una asignación eficiente de la muestra.

Adicionalmente, variables como **localidad**, **tipo de servicio** y **grupo etario** serán consideradas como **dominios secundarios** o empleadas en una etapa posterior de **post-estratificación**. Es decir, la muestra se asignará inicialmente a los estratos principales, y posteriormente se verificará el cumplimiento de cuotas mínimas dentro de los dominios secundarios, ajustando el diseño si fuese necesario para mantener su representatividad.

Estimación del Tamaño de Muestra

Fórmula de Cálculo

Para estimar el tamaño de muestra requerido en cada dominio o estrato de interés, se utiliza la **fórmula clásica para estimación de proporciones** con corrección por población finita. Esta permite garantizar un nivel de precisión estadística deseado, considerando las características del diseño y la estructura poblacional.

1. Tamaño de muestra base:

$$n_0 = \frac{z^2 \cdot p \cdot q}{e^2}$$

Donde:

- $Z = 1.96$ $(correspondiente\ a\ un\ nivel\ de\ confianza\ del\ 95\%)$
- P = 0.5 (proporción esperada de satisfacción; valor conservador que maximiza la varianza).
- $q = 1 - p = 0.5$
- $e = Margen\ de\ error\ aceptable$ $(usualmente\ entre\ 0.04\ y\ 0.05)$

2. Corrección por población finita:

Donde (N_h), representa el tamaño poblacional del estrato h, el tamaño de muestra corregido (n'_h) es:

$$n'_h = \frac{n_0}{1 + \frac{n_0 - 1}{N_h}}$$

3. Ajuste final por efecto de diseño (DEFF) y tasa de no respuesta (TNR):

$$n_h = \frac{n'_h \cdot DEFF}{1 - TNR}$$

Parámetros Asumidos

- $\boldsymbol{Nivel\ de\ Confianza}$: 95% $(Z = 1.96)$
- **Margen de Error (e):** Parametrizable. Se recomienda 4% para dominios principales (como Subred × Canal), y 5% para dominios secundarios o estimaciones globales.
- **Efecto de Diseño (DEFF):** 1.2. Este valor incorpora el aumento de varianza derivado de la estratificación y el muestreo polietápico, especialmente en el canal presencial donde hay conglomeración y ponderaciones diferenciadas.
- **Tasa de No Respuesta (TNR):** 15% (0.15). Este valor se deriva de experiencias previas; puede ser más bajo en canal presencial y más alto en remoto. Se recomienda validarlo en fase piloto y ajustar si es necesario.

Cálculo y Distribución de la Muestra

Una vez definidos los estratos, se procede a aplicar la fórmula de estimación de muestra en cada uno de ellos, siguiendo los parámetros establecidos.

Los pasos son los siguientes:

- **1. Estimar el tamaño de muestra base** para cada estrato principal (Entidad × Canal), utilizando un margen de error estricto, preferiblemente del 4%, para garantizar mayor precisión en los dominios clave.
- **2. Sumar los tamaños muestrales estimados** por estrato, lo que dará como resultado el tamaño de muestra total requerido para la encuesta.
- **3. Distribuir (afijar) la muestra de manera proporcional** al tamaño poblacional de cada estrato, utilizando los universos estimados a partir de registros históricos. Para asegurar la representatividad estadística en estratos de menor tamaño, se establece un **piso mínimo técnico**, recomendado en 100 **encuestas por estrato**, lo cual garantiza que ningún dominio quede subrepresentado.

La estimación del tamaño poblacional de cada estrato (N_h), se basa en la **información recolectada en la Encuesta SDS 2023**, específicamente a partir del **factor de expansión (**f_{exp}**)**. Esta metodología se detalla en la sección siguiente.

Uso de Información Histórica (Encuesta 2023)

La base de datos ***Infometrika_EncuestaSDS2023.xlsx*** constituye un insumo clave para el diseño muestral propuesto, por dos razones fundamentales:

1. Estimación de la Distribución Poblacional

La variable f_{exp} $(factor\ de\ expansión)$ incluida en la encuesta 2023 indica el número de personas de la población total que representa cada encuestado. Al **sumar los valores de f_{exp} dentro de cada estrato**, se obtiene una estimación robusta del tamaño poblacional de ese estrato (N_h),

Ejemplo: Para estimar la población de usuarios del canal presencial en la Subred Norte en 2023, se realiza el siguiente cálculo:

$$N_{Norte,\ Presencial} = \sum f_{exp}$$

$$donde\ SUBRED = \text{Norte}\ y\ \ CANAL = "Presencial"$$

2. Comparabilidad Temporal

Fundamentar el diseño de la muestra 2025 en la estructura poblacional observada en 2023 permite asegurar la **comparabilidad histórica** de los resultados. Esta decisión metodológica garantiza que los indicadores de satisfacción puedan ser evaluados de forma consistente a lo largo del tiempo, fortaleciendo el análisis longitudinal del sistema distrital de salud.

Entidad	*Canal*	*Universo_Encuesta*	*n_base*	*n_corr*	*n_deff*	*n_final*
Capital Salud	Presencial	1.393.725	601	600	720	1.848
Capital Salud	Virtual	3.734	601	518	621	731
Capital Salud	Telefónico	2.669	601	491	589	693
Subred Centro Oriente	Presencial	1.405.784	601	600	720	1.936
Subred Centro Oriente	Virtual	1.916	601	458	549	246
Subred Centro Oriente	telefónico	33.401	601	590	708	1.002
Subred Norte	Presencial	1.356.010	601	600	720	1.601
Subred Norte	Virtual	35.794	601	591	709	834
Subred Norte	telefónico	464.507	601	600	720	1.203
Subred Sur	Presencial	485.862	601	600	720	947
Subred Sur	Virtual	37.523	601	591	709	835
Subred Sur	telefónico	303.342	601	600	719	1.111
Subred Suroccidente	Presencial	8.202	601	560	672	132
Subred Suroccidente	Virtual	31.631	601	590	707	832
Subred Suroccidente	telefónico	144.195	601	598	718	992
Secretaría de Salud	Presencial	112.541	601	598	717	902
Secretaría de Salud	Virtual	94.378	601	597	716	843
Secretaría de Salud	telefónico	97.114	601	597	716	798

Diseño Muestral y Supuestos

Esquema Propuesto

Se adopta un **diseño de muestreo probabilístico, estratificado y polietápico**, estructurado con base en los siguientes lineamientos:

- **Unidad Primaria de Muestreo (UPM):** Definida según el canal de atención — puntos físicos en el canal presencial o eventos de servicio en los canales telefónico y virtual.
- **Estratificación primaria:** Cruzando **Entidad/Subred × Canal de atención**, lo que permite un control operativo y analítico en los principales dominios.
- **Afijación proporcional:** La muestra se distribuye de manera **proporcional al tamaño poblacional estimado** por subred y canal, con un **piso mínimo técnico por estrato** (ej. 100 encuestas), garantizando la **estimabilidad en dominios de menor tamaño**.
- **Post-estratificación:** Por **grupo etario y sexo**, para asegurar un enfoque diferencial y la representatividad de poblaciones clave.

Supuestos Técnicos del Diseño

- **Nivel de confianza:** 95% $(Z = 1.96)$
- **Margen de error:**
  - 4% para dominios principales (Subred × Canal)
  - Hasta 5% para estimaciones en dominios secundarios o globales
- **Efecto de diseño (DEFF):** 1.2
  Justificado por la varianza adicional introducida por la estratificación, el muestreo por conglomerados y la ponderación diferencial.
- **Tasa de no respuesta (TNR):** 15%
  Estimación basada en experiencias anteriores; debe ser validada y ajustada mediante monitoreo en un piloto 2025 si se considera necesario.