

Colombia - Encuesta de Tecnologías de la Información y las Comunicaciones en Hogares - ENTIC Hogares - 2024

Edgar Arlex Garzón Alarcon, Edgar Eduardo Guayazán Sierra

report_generated_on: June 24, 2026

visit_data_catalog_at: <https://microdatos.dane.gov.co/index.php>

Información general

Identificación

NÚMERO DE ID
DANE-DIMPE-ENTIC-H-2024

Información general

RESUMEN

Las encuestas elaboradas a los hogares constituyen una de las principales fuentes de datos socioeconómicos con las que cuentan los países, puesto que a partir de la información obtenida se calculan indicadores para la medición de variados aspectos económicos y sociales. Así mismo, se facilita el conocimiento y la explicación de determinantes o factores causales del comportamiento de la población objeto de estudio, lo cual es importante para el diseño, monitoreo y medición de resultados de políticas públicas.

El uso y la apropiación de las Tecnologías de la Información y las Comunicaciones (TIC) por parte de las personas y los hogares inciden en las preferencias, la productividad y la dinámica interna de la unidad familiar conformada por personas que acceden y se apropian de estas tecnologías. Algunos de los impactos sociales más importantes de acceso y uso de las TIC a través de los hogares se encuentran la oportunidad de acceder a información relevante, la búsqueda de empleo, la generación de ingresos, el intercambio de conocimiento, la creación de habilidades y capacidades y el logro de mayores niveles de integración social. De esta manera, las TIC inciden positivamente en la productividad y la calidad de vida de los hogares y personas.

La necesidad de monitorear las tendencias asociadas a este tipo de desarrollo es reconocida ampliamente por diferentes instituciones del país, pues a nivel de formulación y evaluación de políticas públicas y público-privadas, los planes y programas de estímulo y masificación social de las TIC para Colombia tienden a formularse actualmente con una mejor articulación técnica y mayor congruencia conceptual en relación con los indicadores sociales de acceso y uso de TIC producidos por el Departamento Administrativo Nacional de Estadística (DANE).

Como consecuencia de lo expuesto, a partir del 2007 y hasta 2011, el DANE calcula para Colombia los indicadores básicos de tenencia y uso de las Tecnologías de la Información y las Comunicaciones (TIC) en hogares y por personas, por medio de módulos insertados en la Gran Encuesta Integrada de Hogares (GEIH) y en la Encuesta de Calidad de Vida (ECV) a partir del 2012. Estos indicadores forman parte del programa internacional de medición de la Sociedad de la Información, entendida como el estado del desarrollo económico y social en el cual los individuos y agrupaciones acceden, se apropian, usan y adaptan las TIC de forma frecuente, intensiva, diversificada y significativa para sus vidas.

Para el 2018, Colombia inició el proceso de ingreso a la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y definió un marco de trabajo para medir la transformación digital, por esta razón y con la necesidad de medir información especializada con componentes clave de la economía digital y sus habilitadores (CONPES 3975 de 2019) surgió la Encuesta Tecnologías de la Información y la Comunicación en Hogares (ENTIC) como un instrumento que el DANE desarrolla con el objeto de medir y atender los nuevos requerimientos internacionales sobre medición de la sociedad de la información. Estos indicadores de tenencia y uso de TIC preservan un marco teórico acorde con los acuerdos determinados por expertos nacionales e internacionales acerca del diseño, aplicación e interpretación de encuestas nacionales de TIC y tiene en cuenta los lineamientos del Manual para la medición de acceso y uso de TIC en hogares y personas expedido por la Unión Internacional de Telecomunicaciones (UIT) en el 2020.

Ahora bien, el desarrollo de la ENTIC Hogares utiliza como período de referencia los últimos doce meses y generalmente el periodo de recolección se da entre los meses de agosto y octubre. El formulario de la ENTIC Hogares incluye cinco (5) módulos, con uno inicial de localización y vivienda. Finalmente, este documento contiene los aspectos metodológicos centrales de la ENTIC organizados en tres secciones, la primera sección corresponde a los antecedentes de la encuesta, la segunda sección contiene el diseño de la operación estadística desde la parte temática, estadística, de recolección/acopio, de procesamiento, de análisis, de difusión y declaración, así como la evaluación de las fases del proceso y de los sistemas de producción y flujos de trabajo y para concluir, la tercera sección hace referencia a la documentación relacionada, así mismo, se presenta el glosario, la bibliografía y los anexos.

UNIDAD DE ANÁLISIS
UNIDADES ESTADÍSTICAS

El conjunto de unidades empleadas para el diseño, la medición y la presentación de resultados son:

- a) Unidades de observación: corresponde a las viviendas, los hogares y las personas de 5 o más años.
- b) Unidades de análisis: es representada por las viviendas, los hogares y personas de 5 o más años.
- c) Unidades de muestreo: para esta operación estadística se considera un muestreo en 2 etapas, donde, las unidades primarias de muestreo se relacionan con la primera etapa y las unidades secundarias con la segunda etapa. De tal manera que las unidades de muestreo son las siguientes:
 - Unidades Primarias de Muestreo (UPM): para este estudio las unidades primarias de muestreo corresponden a los municipios del país.
 - Unidades Secundarias de Muestreo (USM): hace referencia a los segmentos o conglomerados de aproximadamente 10 viviendas contiguas, también denominados medidas de tamaño (MT), ubicadas tanto en las cabeceras como en el resto de cada municipio.

Nota: En las USM o MT se mide todas las viviendas, hogares y personas, por lo que no corresponden a unidades muestrales.

Ámbito

NOTAS

CONTENIDO TEMÁTICO

Definición de variables y construcción de indicadores estadísticos

Las principales variables de la ENTIC Hogares son las siguientes:

- Tenencia de bienes y servicios TIC en los hogares: posesión de TIC para el periodo de referencia y características de la conexión.
- Uso de bienes y servicios TIC en los hogares: utilización de modo habitual de TIC. Se mide en términos del principal uso y la frecuencia de uso de TIC.
- Servicios del hogar: conexión a servicios públicos, tales como energía, telefonía e internet, tanto fijo como móvil.
- Seguridad y uso responsable de internet: hábitos y prácticas preventivas y correctivas de las personas para proteger sus conexiones a servicios de telefonía e internet tanto fijo como móvil, redes sociales, datos personales, entre otros.

Por su parte, los indicadores de tenencia y uso de TIC preservan un marco teórico fundamental que corresponde a los principales acuerdos alcanzados por la comunidad de personas expertas, nacionales e internacionales sobre diseño, aplicación e interpretación de encuestas nacionales de TIC, por lo que sigue los lineamientos del Manual para la medición de acceso y uso de TIC en hogares y personas de la ITU (2020). A continuación, se presentan los indicadores principales:

- Indicadores de tenencia de TIC a nivel de hogar.
- Indicadores de uso de TIC a nivel de persona.
- Indicadores de seguridad en TIC.

Resultados estadísticos

Los resultados de la ENTIC Hogares se publican a través de un informe técnico y anexos estadísticos en formato Excel con los principales cuadros de salida. Además, se publican los microdatos anonimizados a través de la Sala de Procesamiento Especializado Externo (SPEE). Los cuadros de salida generados en la operación estadística incluyen las variables representativas de cada módulo de la encuesta a nivel de hogares y personas según temática.

El proceso inicia con el diseño de los cuadros y sus especificaciones, las cuales son elaboradas por el grupo temático y posteriormente enviadas al área de sistemas para su programación y procesamiento. Luego de este paso son enviadas de vuelta al área temática para su revisión y verificación. Para la ENTIC Hogares, los cuadros de salida se presentan con la siguiente desagregación: total nacional, cabecera y centros poblados-rural disperso y total departamental. Los resultados

presentados comprenden tanto valores absolutos como sus respectivos valores porcentuales.

KEYWORDS

Acceso a las TIC.; Banca electrónica y otros servicios financieros en línea.; Brecha digital.; Capacitación en TIC.; Chatbot.; Coeficiente de Variación Estimado (CVE).; Comercio electrónico.; Comprar u ordenar productos o servicios (actividad de uso de Internet).; Consulta de medios de comunicación (actividad de uso de Internet).; Correo y mensajería (actividad de uso de Internet).; Descargar software, imágenes, juegos, música o jugar en línea (actividad de uso de Internet).; Descargar o instalar programas computacionales (software).; Educación y aprendizaje (actividad de uso de Internet).; Formación en habilidades y competencias digitales.; Inteligencia Artificial.; Internet.; Lugar de uso.; Obtener información (actividad de uso de Internet).; Redes sociales.; Suplantación de identidad (incidente digital).; Tableta.; Tecnologías de la Información y Comunicación (TIC).; Teléfono celular convencional.; Teléfono celular inteligente.; Televisión, videos, películas u otro contenido audiovisual para entretenimiento (actividad de uso de Internet).; Televisor.; Trámites con organismos gubernamentales (actividad de uso de Internet).; Entrega de productos en línea en forma digitalizada.; Internet de las cosas: Internet of Things (IoT).; Informante idóneo:

Cobertura

COBERTURA GEOGRÁFICA

COBERTURA GEOGRÁFICA

La ENTIC Hogares cuenta con un cubrimiento nacional y departamental, para el caso de individuos y hogares. Por el diseño muestral no es posible tener desagregación de información a nivel municipal.

Antes de describir la cobertura y desagregación geográfica de la ENTIC Hogares es necesario definir los conceptos de cabeceras, centros poblados y rural disperso

Cabecera: Delimitación geográfica definida por el DANE para fines estadísticos, alusiva al área geográfica delimitada por el perímetro censal. A su interior se localiza la sede administrativa del municipio, es decir la alcaldía. Departamento Administrativo Nacional de Estadística - DANE. (2018). Censo Nacional de Población y Vivienda CNPV, Manual de Conceptos 2018.

Centro poblado: Concepto construido por el DANE con fines estadísticos, para la identificación y localización geográfica de núcleos o asentamientos de población. Se define como una concentración mínima de veinte viviendas contiguas, vecinas o adosadas entre sí, ubicada en el área resto municipal o en un área no municipalizada (corregimiento departamental). Contempla los núcleos de población de los corregimientos municipales, inspecciones de policía y caseríos. Dicha concentración presenta características tales como la delimitación de vías vehiculares y peatonales. Departamento Administrativo Nacional de Estadística - DANE. (2018). Censo Nacional de Población y Vivienda CNPV, Manual de Conceptos 2018.

Este concepto considera:

Caserío: Definición construida por el DANE para fines estadísticos, que corresponde a un centro poblado ubicado generalmente al lado de una vía principal y que no tiene autoridad civil. El límite censal está definido por las mismas viviendas que constituyen el conglomerado. Departamento Administrativo Nacional de Estadística (DANE). Manual Técnico Censo general 2005

Inspección de Policía: El DANE, para fines estadísticos, la define como conglomerado de viviendas ubicadas en el área resto municipal que tiene por autoridad principal al inspector de policía. Instituto Geográfico Agustín Codazzi (IGAC). Glosario Diccionario geográfico de Colombia.

Corregimiento municipal: Tipo de centro poblado, ubicado en el área rural de un municipio, el cual incluye un núcleo de población y está considerado en los Planes de Ordenamiento Territorial (POT). Adaptado de: Ley 136 de 1994. Por la cual se dictan normas tendientes a modernizar la organización y el funcionamiento de los municipios. Diario Oficial No. 41.377 de 2 de junio de 1994. y Ley 1551 de 2012. Por la cual se dictan normas para modernizar la organización y el funcionamiento de los municipios. Diario Oficial No. 48.483 de 6 de julio de 2012.

Rural disperso: Delimitación geográfica definida por el DANE para fines estadísticos, comprendida entre el perímetro censal de las cabeceras municipales y de los centros poblados, y el límite municipal. Se caracteriza por la disposición dispersa de viviendas y de explotaciones agropecuarias existentes en ella. Corresponde al territorio que no forma parte ni de la cabecera municipal (clase 1) ni de los centros poblados (clase 2). Se caracteriza por objetos y elementos relacionados con la agricultura, predios de descanso o recreo, usos mineros o extractivos. El número de unidades residenciales por área es

menor a las zonas urbanas. Departamento Administrativo Nacional de Estadística - DANE. (2018). Censo Nacional de Población y Vivienda CNPV, Manual de Conceptos 2018.

Resto: Corresponde al área geográfica cubierta por el centro poblado y rural disperso. (Departamento Administrativo Nacional de Estadística).

GEOGRAPHIC UNIT DESAGREGACIÓN GEOGRÁFICA

La ENTIC Hogares tendrá una desagregación geográfica nacional por cabecera y centros poblados - rural disperso y Total por departamentos.

- Total Nacional: Total nacional y desagregado por cabecera y centro poblado - rural disperso.

- Departamental: Total por departamento

UNIVERSO UNIVERSO DE ESTUDIO

El universo para la Encuesta Tecnologías de la Información y Comunicación en Hogares (ENTIC) Hogares está conformado por la población civil no institucional residente en todos los municipios del territorio nacional, excluyendo la parte rural de San Andrés

Productores y Patrocinadores

INVESTIGADOR(ES) PRIMARIO(S)

Nombre	Affiliation
Edgar Arlex Garzón Alarcon	Dirección de Metodología y Producción Estadística - DIMPE
Edgar Eduardo Guayazán Sierra	Dirección de Metodología y Producción Estadística - DIMPE

OTROS PRODUCTORES

Nombre	Affiliation	Role
Dirección de Metodología y Producción Estadística-GIT Temática de industria	Gobierno Nacional	Equipo Técnico

FINANCIAMIENTO

Nombre	Abbreviation	Role
Convenio Interadministrativo 827-2024 suscrito entre el Fondo Único de Tecnologías de la Información y las Comunicaciones, el Departamento Administrativo Nacional de Estadística DANE - FONDANE	Fondo Único de TIC - DANE - FONDANE	Ejecutor

FECHA DE PRODUCCIÓN DEL METADATO

2026-04-01

ID DEL DOCUMENTO IDD

COL-DANE-ENTIC-H-2024

Muestreo

Procedimiento de muestreo

MARCO ESTADÍSTICO MUESTRAL

La muestra de la ENTIC hogares de 2024 se selecciona en dos etapas de muestreo. Cada etapa tiene un marco muestral asociado, la primera se usa un marco de municipios con una estructura de estratificación ajustada a las necesidades de las investigaciones sociales y la segunda corresponde con medidas de tamaño o segmentos estratificada por dominio geográfico y tiene su origen en el marco geoestadístico nacional, producto del Censo Nacional de Población y Vivienda (CNPV) del 2018.

Marco Muestral de municipios

El marco de municipios es un marco que enlista cada municipio del país (vigencia censo 2018), de modo que cuenta con información auxiliar de las poblaciones de cabecera y centro poblado - rural disperso, indicaciones del índice de pobreza multidimensional, porcentaje de centro poblado y rural disperso. De igual forma, incluye los estratos de muestreo de la primera etapa.

Marco Muestral de Medidas de Tamaño o Segmentos

Es un marco construido a partir de la información estadística de densidad de viviendas en las manzanas y secciones rurales del marco geoestadístico nacional, sirve para identificar y ubicar geográficamente las USM, por lo tanto, enlista cada conglomerado de 10 viviendas del país, asociándole la manzana o sección rural a la que pertenece.

Para construir el marco se toma el número de viviendas que hay en cada manzana o sección rural del marco geoestadístico nacional y luego se divide en tamaños de aproximadamente 10 viviendas. Como resultado se obtiene el número de medidas de tamaño teóricas asociadas a cada manzana y sección rural del país. Con esta información se crea un marco nuevo de muestreo donde las filas son los conglomerados; por ejemplo, si una manzana tiene 30 viviendas entonces al dividir en 10, significa que se tienen 3 medidas de tamaño y que, en consecuencia, en el marco de conglomerados, van a aparecer 3 filas para representar los 3 conglomerados.

Con este marco se selecciona directamente la muestra de segmentos en la segunda etapa de esta encuesta, luego de la selección se identifican las manzanas o secciones asociadas a la muestra y se procede a crear un marco de viviendas segmentado. Para el caso de cabeceras y centro poblado se desarrolla un operativo de campo dirigido a las manzanas asociadas, a estas medidas de tamaño, donde se procede a enlistar todas las viviendas contenidas en las manzanas. Este marco se segmenta en medidas de tamaño de 10 viviendas. En lo que respecta a la zona rural dispersa, cada sección se segmenta creando polígonos que, según información de foto satelital contengan un aproximado de 10 viviendas. En ambos casos la segmentación se enumera y posteriormente se encuesta la medida de tamaño elegida en la primera parte.

DISEÑO MUESTRAL

La ENTIC-Hogares es una encuesta dirigida a hogares y personas, por lo tanto, se establece una estrategia muestral adecuada para acceder a mediciones de dichas unidades de análisis con procedimientos que aseguran estimadores aproximadamente insesgados de los parámetros bajo una precisión y confianza preestablecida de antemano. A continuación, se describirá en detalle el plan muestral.

Tipo de muestreo:

Teniendo en cuenta los objetivos y las características de la encuesta se utiliza una muestra probabilística, multietápica, estratificada y de conglomerados.

a) Probabilística: Se satisfacen los siguientes criterios, a saber:

1. Se conoce de ante mano todas las posibles muestras.
2. Cada posible muestra tiene su respectiva probabilidad de selección.
3. El procedimiento da a cada elemento del universo de estudio una probabilidad de inclusión conocida y mayor que cero.
4. Se tiene un mecanismo aleatorio dentro del cual a cada posible muestra tiene una probabilidad no nula de selección.

b) Multietápica: Para lograr la muestra de viviendas, hogares y personas se seleccionaron secuencialmente las unidades de muestreo (UPM y USM) en dos etapas

- En la primera se seleccionan municipios con un muestreo sin reemplazamiento proporcional al tamaño.
- En la segunda se seleccionan conglomerados de 10 viviendas con un muestreo sistemático con un ordenamiento geográfico previo.

c) Estratificada: Para la ENTIC Hogares se utiliza el sistema de estratificación de municipios creado desde 2020 para UPM y otra basada en un criterio geográfico para USM.

Respecto a la estratificación de UPM, en la primera etapa de la muestra se generan grupos de municipios dentro de cada departamento. En el primer grupo de cada departamento se establece el estrato de inclusión forzosa que contiene un único municipio que corresponde a la capital de departamento.

El segundo grupo lo componen los municipios considerados como áreas metropolitanas de las capitales de departamento. Las otras agrupaciones de municipios se hicieron a partir de un Análisis de Componentes Principales (ACP) y un posterior agrupamiento tomando en cuenta la proyección de los municipios sobre los nuevos ejes factoriales y la población. Para lograrlo, se tomaron los resultados del Censo 2018, el índice de pobreza multidimensional, sus 15 indicadores y el porcentaje de urbanización. Con el propósito de buscar un diseño muestral auto ponderado, los grupos consolidados tienen aproximadamente la misma población, aunque esta característica no se puede asegurar; por tal motivo, van a existir unos estratos más grandes que otros y algunos de inclusión forzosa de municipios con una gran población, es decir, que no corresponden al tamaño promedio de los estratos.

En la segunda etapa la estratificación se hace a nivel de cabecera, centro poblado y rural disperso. Por último, se hace una estratificación implícita de los segmentos al ordenar el marco según sector, sección y manzana y luego seleccionar a través de un muestreo sistemático la muestra de USM. Este procedimiento garantiza que la muestra quede uniformemente distribuida sobre las áreas geográficas.

d) De conglomerados: Los conglomerados están definidos por la unidad secundaria de muestreo (USM) y corresponden a 10 viviendas (contiguas) en promedio donde se encuestan todas las viviendas, los hogares y las personas que lo conforman.

Cálculo del tamaño de muestra

Para establecer un tamaño de muestra, en primera instancia, se debe identificar un indicador relevante en la investigación al cual se le asigne un tamaño de muestra suficiente para que el estimador de este indicador cumpla con ciertas características de precisión y confianza.

En los casos donde el diseño muestral es complejo, además de lo anterior, se debe tomar como referencia un muestreo aleatorio simple de elementos (MAS) estableciendo un tamaño de muestra primario, el cual deberá ser ajustado por un efecto del diseño muestral (Deff) y otros parámetros como la tasa de no respuesta esperada y las subdivisiones de la población donde, en esta última, se procura garantizar suficiente muestra para su respectiva desagregación. Se invita al lector a revisar el desarrollo teórico del tema en (Cochran, 1977) o en las metodologías publicadas por Naciones Unidas sobre encuestas de hogares (Naciones Unidas, 2008).

El tamaño de la muestra de la ENTIC hogares 2024 corresponde a la encuesta de 5.271 segmentos en 421 municipios y en los cuales se proyecta alcanzar 53.300 hogares. Una vez realizado el operativo de recolección en campo, se obtuvo información de 53.570 hogares con encuestas completas.

Desviaciones del diseño de la muestra

SELECCIÓN DE MUESTRA

En la primera etapa, la UPM o municipio se selecciona dentro de cada estrato de muestreo con un algoritmo de selección proporcional al tamaño (PPT) con covariable de tamaño dada por el censo total de personas para 2018.

Para la segunda, se ordena el marco por sector, sección y manzana luego, se hace una segmentación del marco muestral, de tal manera que la USM tenga aproximadamente 10 viviendas en promedio. Luego dentro de cada estrato de la etapa se hace una selección sistemática de segmentos con único arranque. Para asegurar un tamaño de muestra fijo se usa el método de "Fractional Interval Method".

Todos los procedimientos se hacen con PROC SURVEYSELECT del software SAS (SAS Institute Inc, 2019). Por lo cual, las especificaciones de algoritmos usados se pueden revisar en la página web del programa.

Mantenimiento de muestra y manejo de novedades muestrales:

En el desarrollo normal de la operación estadística, se identifican diversos aspectos que pueden causar sesgos tales como la desactualización de marco, la imposibilidad de acceder a la muestra seleccionada con el procedimiento probabilístico predefinido debido a inconvenientes en campo y al agotamiento de la fuente.

La desactualización de marco se genera con el tiempo debido a que en las áreas geográficas la población puede variar. Para controlar la desactualización del marco se procede a tomar en cuenta estos cambios con un factor de corrección. En la medida que las manzanas varíen en su número de viviendas las medidas de tamaño también lo harán en proporción ajustándolas en su factor de expansión con un factor de corrección que toma en cuenta el cambio del número de viviendas de la manzana respecto a lo medido en el censo. La fórmula se define como el número de segmentos actuales sobre el número de segmentos que aparecían en el marco al momento de seleccionar la medida de tamaño. Posteriormente, se realiza un proceso en el cual, con la información del recuento se va actualizando el marco de medidas de tamaño en lo que respecta al número de segmentos y luego, se ajusta la muestra según el resultado de la actualización.

Con respecto a la de pérdida de muestra, lo que se solicita es hacer revisitas en la medida de lo posible. En los casos donde se presentan novedades como pérdida de segmentos por problemas de orden público o por rechazo de segmentos, se tiene protocolos que se siguen para evitar la pérdida; sin embargo, en los casos donde es inevitable la pérdida de segmentos se opta por reemplazarlos con segmentos lo más parecidos posible al segmento original (estos casos son escasos).

Tasa de respuesta

AJUSTES DE COBERTURA (O AJUSTE DE COBERTURA POR NO RESPUESTA)

El ajuste de cobertura se realiza en los casos donde se pierden viviendas, hogares y personas, el procedimiento describe la construcción de un factor de cobertura (FCOB). Este factor restituye a nivel segmento las pérdidas ocasionadas por el operativo de campo. Es importante resaltar que, para evitar sesgos de selección, cuando se pierde de forma parcial o total la información de una persona dentro del hogar, el hogar se da como perdido y se toma en cuenta en el factor de cobertura.

El resultado final de los indicadores por cobertura en campo, se obtuvieron de la siguiente manera:

- A nivel de viviendas: Total de viviendas finales / Total viviendas finales esperadas.
- A nivel de hogares: Total de hogares con encuesta completa/ Total de hogares con encuesta completa esperados.
- A nivel de persona: Total de personas con encuesta completa/ Total de personas con encuesta completa esperadas.

Para este cálculo, la cantidad de viviendas, hogares y personas esperadas corresponden a la información derivada del proceso de sensibilización.

Dominio	Cobertura viviendas	Cobertura hogares con Encuesta Completa	Cobertura personas Encuesta Completa
TOTAL NACIONAL	99,75%	98,79%	99,54%

Ponderación

ESPECIFICACIONES DE PONDERADORES

El software utilizado para el tratamiento de los ponderadores es un desarrollo propio del DANE en el programa SAS (Statistical Analysis System). Los resultados muestrales se desarrollan conforme a un ajuste por variable exógena, que es un estimador independiente de población, y permite mejorar las estimaciones referentes a las desagregaciones geográficas; se igualan los totales estimados con base en la muestra, con los totales del censo de población proyectados a la fecha de la encuesta, en este caso el CNPV de 2018. Lo anterior se apoya en la premisa de que la estructura por desagregaciones geográfica, proyectada a partir de un censo de población reciente, es más exacta que la estimada a partir de la muestra. A continuación, se muestra las etapas en las que se construyen los ponderadores:

1. Factores de corrección de cobertura: Se calcula la cobertura de la muestra donde se identifican los segmentos y hogares perdidos en campo. Luego con la información de los conteos se construyen los factores de corrección en la base.
2. Factores básicos: Se adiciona a la base los factores de expansión del diseño muestral. Se multiplican con los factores de corrección de la etapa anterior y como resultado se tiene un factor básico.
3. Factores techo: Se toma en cuenta la población de cada estrato de muestreo para calibrar el factor anterior a las

poblaciones censales del año 2018 de los estratos de muestreo preestablecidos en el diseño.

4. Factores de expansión calibrados: Se usan las proyecciones poblacionales publicadas por el DANE, para ajustar según el modelo de calibración expuesto anteriormente el factor de techo.

El cálculo del ajuste por calibración se realiza utilizando la macro Clan 97 v3.1. (creada por Statistics Sweden), la cual corresponde a un conjunto de rutinas en el programa SAS para la estimación puntual y de errores en encuestas por muestreo.

Finalmente, a nivel del software STATA, se utilizan los objetos de base de datos, que se construyeron en el ítem anterior, y se realiza la programación para el cálculo de variables auxiliares (porcentajes, máximos y mínimos) y la generación del formato de salida. Una vez los cuadros se generan pasan al equipo de muestras para calcular los coeficientes de variación.

Formularios

Información general

DISEÑO DEL CUESTIONARIO

En coherencia con el objetivo de obtener la información necesaria para la actualización de los indicadores de tenencia y uso de las TIC a nivel de viviendas, hogares y personas, cuyos resultados sirvan como insumo para la definición de políticas que permitan diseñar y ejecutar planes sociales.

El formulario diseñado para la ENTIC Hogares incluye cinco (5) módulos permanentes de la operación estadística.

- Módulo I: Localización y vivienda.

En este módulo se hace una identificación de la ubicación de las viviendas y los hogares, así como un control de calidad sobre la recolección de las encuestas y preguntas a nivel de vivienda.

-Módulo II: Tenencia de Bienes y Servicios TIC en el Hogar.

En este módulo se capta información relacionada con la tenencia Bienes y Servicios TIC en el hogar (Tenencia telefónica, computadores, televisores y dispositivos tecnológicos; servicios de conexión a internet, televisión y radio, máxima velocidad de internet).

-Módulo III: Características y composición del hogar (para todas las personas del hogar)

En este módulo se busca identificar las personas que conforman el hogar y establecer su parentesco con el jefe(a) del hogar. Además, se realiza una caracterización demográfica de todos los integrantes en aspectos como sexo, edad, estado civil, autorreconocimiento étnico, lugar de nacimiento y aspectos de migración.

-Módulo IV: Uso de TIC (para todas las personas de 5 años y más)

El objetivo del módulo es medir el uso de las Tecnologías de la Información y las Comunicaciones (TIC) por parte de las personas, con énfasis en el uso del computador (de escritorio, portátil y tableta), la internet y la telefonía móvil celular. Entre otros aspectos, se indaga por lugares de acceso, frecuencia de uso, actividades realizadas a través de Internet, utilización de herramientas de Inteligencia Artificial (IA), tenencia y uso de la telefonía móvil celular y uso de la señal de radio.

-Módulo V: Seguridad y uso responsable del Internet (para todas las personas de 5 años y más)

En este módulo se busca identificar los mecanismos que usan las personas durante el uso del Internet para proteger de posibles problemas como captura de virus, pérdidas económicas y divulgación no consentida de videos o imágenes, entre otras. Así mismo, se indaga por los incidentes de seguridad digitales ocurridos y las acciones que toman las personas después de haber sido víctima de uno de estos.

Así mismo, con el fin de garantizar la coherencia de la información recolectada y la secuencia lógica de las preguntas, el cuestionario fue diseñado en conjunto con las especificaciones de validación donde se documentan las condiciones para mantener la consistencia entre las variables y por lo tanto, el correcto flujo de la información, junto con la caracterización de estas.

Recolección de datos

Fecha de recolección de datos

Start	End	Cycle
2024-07-19	2024-12-24	N/A

Time Periods

Start	End	Cycle
-------	-----	-------

Notas de recolección de datos

DISEÑO DE LA RECOLECCIÓN/ ACOPIO

Con la finalidad de garantizar un óptimo desarrollo del operativo de campo se desarrolla el diseño de la recolección en ocho partes: 1) Método y estrategias de recolección de datos; 2) Estructura organizacional del operativo; 3) Esquema de entrenamiento; 4) Conformación del equipo; 5) Proceso de sensibilización y acuerdos de intercambio; 6) Elaboración de manuales; 7) Diseño de las estrategias de comunicación y 8) Diseño de estrategias de seguimiento y control.

Actitud de las fuentes

Las sedes manifestaron que en un porcentaje considerable las fuentes se encontraron dispuestas a brindar la información requerida y fueron receptivas al acercamiento de los equipos de recolección dando respuesta a la encuesta. No obstante, la disposición de las personas tiende a ser menor en los segmentos ubicados en las capitales y áreas metropolitanas que en áreas rurales.

En cuanto a la actitud según el estrato trabajado; mientras que en los estratos más bajos las personas mostraron una mejor recepción hacia el suministro de información, en los altos fueron más reacios a responder las preguntas, aludiendo en especial temas relacionados con la disponibilidad de tiempo. De otra parte, la percepción de inseguridad a nivel nacional produjo desconfianza de los hogares en diferentes zonas y estratos de la muestra asignada.

Formularios

DISEÑO DEL CUESTIONARIO

En coherencia con el objetivo de obtener la información necesaria para la actualización de los indicadores de tenencia y uso de las TIC a nivel de viviendas, hogares y personas, cuyos resultados sirvan como insumo para la definición de políticas que permitan diseñar y ejecutar planes sociales.

El formulario diseñado para la ENTIC Hogares incluye cinco (5) módulos permanentes de la operación estadística.

- Módulo I: Localización y vivienda.

En este módulo se hace una identificación de la ubicación de las viviendas y los hogares, así como un control de calidad sobre la recolección de las encuestas y preguntas a nivel de vivienda.

-Módulo II: Tenencia de Bienes y Servicios TIC en el Hogar.

En este módulo se capta información relacionada con la tenencia Bienes y Servicios TIC en el hogar (Tenencia telefónica, computadores, televisores y dispositivos tecnológicos; servicios de conexión a internet, televisión y radio, máxima velocidad de internet).

-Módulo III: Características y composición del hogar (para todas las personas del hogar)

En este módulo se busca identificar las personas que conforman el hogar y establecer su parentesco con el jefe(a) del hogar. Además, se realiza una caracterización demográfica de todos los integrantes en aspectos como sexo, edad, estado civil, autorreconocimiento étnico, lugar de nacimiento y aspectos de migración.

-Módulo IV: Uso de TIC (para todas las personas de 5 años y más)

El objetivo del módulo es medir el uso de las Tecnologías de la Información y las Comunicaciones (TIC) por parte de las

personas, con énfasis en el uso del computador (de escritorio, portátil y tableta), la internet y la telefonía móvil celular. Entre otros aspectos, se indaga por lugares de acceso, frecuencia de uso, actividades realizadas a través de Internet, utilización de herramientas de Inteligencia Artificial (IA), tenencia y uso de la telefonía móvil celular y uso de la señal de radio.

-Módulo V: Seguridad y uso responsable del Internet (para todas las personas de 5 años y más)

En este módulo se busca identificar los mecanismos que usan las personas durante el uso del Internet para proteger de posibles problemas como captura de virus, pérdidas económicas y divulgación no consentida de videos o imágenes, entre otras. Así mismo, se indaga por los incidentes de seguridad digitales ocurridos y las acciones que toman las personas después de haber sido víctima de uno de estos.

Así mismo, con el fin de garantizar la coherencia de la información recolectada y la secuencia lógica de las preguntas, el cuestionario fue diseñado en conjunto con las especificaciones de validación donde se documentan las condiciones para mantener la consistencia entre las variables y por lo tanto, el correcto flujo de la información, junto con la caracterización de estas.

Data Collectors

Nombre	Abbreviation	Affiliation
Departamento Administrativo Nacional de Estadística	DANE	Gobierno Nacional

Supervisión

DISEÑO DE LA ESTRATEGIA DE SEGUIMIENTO Y CONTROL

Para el control de calidad de la información se utilizan normas de validación y consistencia y controles operativos de calidad tales como el resumen de cobertura, las alarmas e inconsistencias, el monitoreo, los indicadores de cobertura de vivienda, hogares y personas y el Indicador de Calidad de la Recolección (ICR). Estos serán explicados a continuación:

- Especificaciones de validación y consistencia En primera instancia, el grupo temático define, con base en las características de la investigación, una serie de validaciones y consistencias entre preguntas, filtros y flujos del formulario. Éstas son implementadas en el aplicativo dispuesto en el DMC por la Oficina de Sistemas, junto con validaciones adicionales definidas por el grupo de la DRA para el seguimiento y control de los resultados operativos.

Este proceso se efectúa mediante pruebas directas en las que participan los equipos que hacen parte de la encuesta (DRA, Sistemas, Diseños muestrales y Temática). Se destaca que esta etapa se realiza antes de iniciar el operativo en campo, dado que el propósito es hallar y corregir posibles errores o dificultades que contenga el formulario de la encuesta.

- Instrumentos de control para supervisión Para el proceso de control, el grupo de la Dirección de Recolección y Acopio de la encuesta diseñó un conjunto de formatos que facilitan el monitoreo y supervisión tanto del proceso de recuento, como de sensibilización y recolección de la información. Esto permite comparar los datos obtenidos en campo con el reporte generado en la base de datos de encuestas transmitidas. Entre ellos se enmarcan:

- Resumen acumulado de viviendas, hogares y personas por segmento.

- Resumen de cobertura. • Formato diario de supervisión de campo.

- Informe técnico de coordinación de campo. Las sedes son responsables de verificar y asegurar la calidad de la información recolectada y transmitida al DANE Central.

El seguimiento a la calidad inicia con el supervisor de campo, quien hace la respectiva verificación confirmando que todas las encuestas estén totalmente diligenciadas y con la calidad establecida, de modo que esta es una actividad que hace al momento de cerrar la encuesta.

Ahora bien, desde DANE Central se hará verificación de la información durante el operativo para que las sedes tomen correctivos oportunamente a cada una de las inconsistencias. Semanalmente se generan reportes de cobertura, los cuales son enviados a los asistentes de encuesta para que verifiquen el estado de cobertura del operativo, además de observar los avances en la cobertura estimada.

Resumen de cobertura El objetivo de este formato es llevar un control de manera resumida del número de viviendas, hogares y personas encontradas en los segmentos asignados y trabajados durante la recolección para detectar posibles inconsistencias y alarmas en la información recolectada. Este formato es utilizado en el operativo urbano, en las cabeceras municipales y en el área rural disperso. El principal usuario del formato es la persona encargada de la coordinación de campo, quien consolida la información de la supervisión a cargo para remitir semanalmente estos resúmenes a la Dirección de Recolección y Acopio en el DANE central vía FTP. De este reporte se obtienen los indicadores de cobertura operativa y resultados que permiten medir el proceso de seguimiento.

Detección de alarmas y corrección de inconsistencias Del resumen operativo de cobertura se obtienen las inconsistencias detectadas en la revisión semanal. Al realizar la revisión, se reportan las cifras que son motivo de justificación por parte de las sedes tales como la baja cobertura, segmentos con menos de 10 viviendas encontradas, alto número de vacantes, entre otros. Las inconsistencias deben ser corregidas por la ciudad correspondiente, así mismo las alarmas deben ser justificadas en el menor tiempo posible.

Identificación de errores temáticos y solicitud de revisión Además de las inconsistencias revisadas durante el operativo de campo por el equipo de la DRA, en el proceso de carga de las bases de datos el equipo temático identifica posibles errores de diligenciamiento, enviando reportes a las territoriales para verificar si son errores efectivos o si existe una justificación para las alertas generadas.

Monitoreo Periódicamente se realizan monitoreos a los grupos de campo para observar el desempeño de éstos y detectar falencias que pueden afectar la calidad de la información recolectada. En este orden, se observan aspectos como la presentación personal; la lectura adecuada de las preguntas; la no-inducción de respuestas; los cambios de capítulo; la realización de buenos sondeos; el registro adecuado de los miembros del hogar y el manejo de conceptos. De igual forma, se verifica la correcta distribución de materiales de trabajo, el diligenciamiento de formatos, el manejo del transporte, la revisión de las encuestas por parte del personal de supervisión, la ubicación en el segmento y la capacidad para crear un ambiente propicio con la fuente que permita desarrollarla. Todos los aspectos para mejorar son socializados con los responsables de cada proceso y se hace seguimiento a la puesta en marcha de las soluciones y correctivos que sean necesarios.

Informe de cobertura campo vs sistemas Semanalmente, luego de consolidar la totalidad de la información de cobertura de campo de cada una de las sedes, se efectúa un cruce de variables entre la base de sistemas y el resumen operativo de cobertura, comparando la cantidad de viviendas, hogares y personas a nivel de segmento. La información consolidada y comparada se envía a las sedes para su respectiva revisión y corrección. De esta forma, las sedes pueden detectar información con errores de digitación en los resúmenes de cobertura y la posible falta de información o duplicidad de viviendas en la base de sistemas. A partir de esto, las ciudades deben hacer los ajustes respectivos al resumen de cobertura y transmitirlo nuevamente vía FTP. Si la información de campo es correcta, deben verificar las inconsistencias en la base de sistemas y enviar nuevamente los archivos respectivos comunicando la novedad al ingeniero encargado en el DANE Central.

Indicadores de cobertura en viviendas, hogares y personas Semanalmente se realiza seguimiento a las tasas de cobertura operativa, con el fin de detectar oportunamente comportamientos atípicos que permitan presumir debilidades en el desempeño de los grupos de trabajo en campo, y así efectuar indagaciones más profundas para tomar los correctivos respectivos.

- Cobertura de viviendas: se obtiene de cruzar la información de las viviendas finales encontradas en el operativo de campo contra el total de viviendas encontradas en sensibilización * 100. Los datos son obtenidos de lo reportado en el resumen de cobertura y de lo encontrado en sensibilización.
- Cobertura de hogares: se obtiene de cruzar la información de hogares con encuesta completa sobre cantidad de hogares obtenidos en sensibilización * 100. Los datos son obtenidos de lo reportado en el resumen de cobertura y de lo encontrado en sensibilización.
- Cobertura de personas: se obtiene de cruzar la información del total de personas con encuesta completa contra la cantidad de personas obtenidas en sensibilización * 100. Los datos son obtenidos de lo reportado en el resumen de cobertura y de lo encontrado en sensibilización.

Indicador de Calidad de la Recolección (ICR) El objetivo de este indicador es determinar el grado de calidad y confiabilidad de la información enviada por la persona encargada de la coordinación de campo de cada sede. Este indicador se mide a través del número de inconsistencias detectadas en el acompañamiento al personal encuestador en campo. Esto permite hacer seguimiento continuo a las diferentes sedes que presentan bajos indicadores de calidad de la información, con el fin de determinar las posibles fallas y sugerir recomendaciones o medidas de mejoramiento. Se calcula un indicador de calidad para cada persona del grupo de recolección, el cual es obtenido con el promedio aritmético de los puntajes alcanzados para los formularios que fueron sometidos al proceso de supervisión y control de calidad, mediante un formato dispuesto

especialmente para este fin. Dentro del cálculo del tipo de error cometido se crea una variable de penalización sobre el resultado del indicador para hacer distinción de dichas faltas, las cuales se penalizan de acuerdo con el tipo de inconsistencia.

Procesamiento de datos

Edición de datos

DISEÑO DE PROCESAMIENTO

A continuación, se presenta el diseño de las herramientas tecnológicas de software y hardware, que se implementan para el procesamiento de los datos (software y hardware), los programas requeridos para la grabación, la consolidación y el almacenamiento.

Consolidación de archivos de datos

Recibidos los archivos, la información es descargada del buzón FTP y almacenada en una estructura jerárquica que permite realizar la descrición de los datos. Una vez se seleccionan los archivos que se deben cargar en la base de datos, se realiza el proceso de carga utilizando herramientas de ETL (extracción, transformación y carga).

Cargue de información

Al iniciar el proceso de carga se verifica la identificación única para cada encuesta, garantizando que no exista duplicidad en la información y se asigne un directorio único a cada encuesta.

Una vez cargados los datos, se genera un log que indica el estado o posibles fallas presentadas durante el proceso y se procede a trasponer toda la información. Este proceso permite poblar las tablas creadas en la base de datos para cada capítulo de la investigación. Así, cada tabla se relaciona con cada nivel manejado en la encuesta (vivienda - hogar - persona).

Consistencia de información

A través del desarrollo de sentencias DML (Data Manipulation Language, por sus siglas en inglés), se identifican posibles inconsistencias por duplicidad a nivel de personas y problemas de selección de segmentos a trabajar. Así mismo, se generan reportes de cobertura y frecuencias que permiten realizar constantemente el monitoreo y control a la información que se ha recolectado y cargado en las bases de datos. Después, en el esquema, definido en el motor de base de datos, se crean procedimientos para validar la información de la base de datos, flujos, variables sin información, rangos y estructura, para obtener una base de datos consistente. Así, cada investigación se almacena en su esquema de base de datos. Esto permite que los datos se mantengan aislados de la información de otras investigaciones y los permisos de acceso a los usuarios se realicen mediante los mecanismos de roles y privilegios propios del sistema manejador de base de datos.

Codificación

En la estructura del formulario se debe asociar a cada pregunta el código equivalente con el de pregunta en las tablas asignadas para cada capítulo en el esquema de datos, incluyendo el número de formulario. Adicionalmente, como parte del diseño se crea un ETL (Extracción Transformación y Carga, por sus siglas en inglés) y se carga la muestra a una tabla de la base de datos, para poder realizar el cargue de la información correspondiente.

Diccionario de datos

En el diccionario de datos se describen por temáticas las características de todas las variables que son (o serán) utilizadas en la operación estadística en concordancia con las tablas de datos generadas. Así mismo, este apartado ayuda a definir las especificaciones de estimación (para operaciones estadísticas por muestreo). Todo esto, se realiza con base el formulario de la encuesta realizada, que será actualizado cada año.

El diccionario de datos de la ENTIC hogares, facilita el cumplimiento de las siguientes tareas: -Configuración de la base de datos: permite visualizar la estructura de la base de datos por capítulos, variables temáticas y/o por unidades de observación para todas las preguntas aplicadas con el instrumento de recolección en el operativo de campo.

También, permite realizar consultas rápidas sobre la caracterización de variables, presentando la codificación de las variables, tipo, extensión y posibles respuestas.

- Reglas de edición (validación y consistencia) e imputación: permite realizar una revisión de las variables para su dominio

geográfico en términos de viviendas, hogares y personas. Así mismo, es utilizado por el equipo de temática para realizar las respectivas pruebas de inconsistencias y ser informadas oportunamente para resolverse con el equipo de campo.

Revisión y validación

Las sedes realizan los procesos de revisión y verificación y luego se transmite la base de datos al DANE Central, en esta etapa, los analistas de la operación estadística realizan la revisión de la consistencia de la información. En el evento en que se detecten inconsistencias, las mismas se resuelven en la fase de devolución, en la cual, la sede debe realizar las respectivas observaciones. Estos procesos de aclaración o corrección de la información se realizan por teléfono, correo electrónico o visitas a la fuente cuando la situación lo amerite.

Una vez la base de datos es aprobada por el equipo de la Dirección de Recolección y Acopio (DRA), la misma es remitida al equipo temático de la operación estadística, el cual realiza una revisión de calidad a través de un análisis de la información a nivel de agregados totales por departamento, cabecera municipal y centros poblados y rural disperso y agregados por sexo. Para ello, se recurre a calcular variaciones, contribuciones y participaciones; sin embargo, si tras el análisis se encuentran inconsistencias o aclaraciones sobre la información, se retorna al grupo de analistas del DANE Central, quienes deben aclarar las respectivas correcciones de la información, consultando directamente con las fuentes de información. Posterior al desarrollo de dicho proceso, la base de datos se entrega depurada al equipo de temática que luego del análisis y aprobación de las aclaraciones o ajustes realizados a la base, procede a la elaboración de los productos para publicación.

Posteriormente, se realiza un comité interno en el cual un grupo de funcionarios de las diferentes áreas internas del DANE (Dirección de Recolección y Acopio, la Dirección de Regulación, Planeación, Estandarización y Normalización, la Dirección de Metodología y Producción Estadística y la Dirección de Síntesis y Cuentas Nacionales) realizan el análisis de la información tanto a nivel de la información agregada como a nivel del contexto durante el período de referencia. De igual forma, el equipo de DIRPEN realiza una validación de las métricas de calidad para la iteración de la ENTIC Hogares que se piensa publicar.

Diseño de instrumentos de edición (validación y consistencia) e imputación de datos

En el procesamiento de datos se utiliza el programa informático SAS (Statistical Analysis System) diseñado para el análisis de datos y soluciones estadísticas. Para esto se lleva a cabo la verificación interna de los datos y ajustes; ajustes de cobertura y generación de cuadros de salida y el proceso de anonimización.

-Verificación de la consistencia interna de los datos y ajustes Se realiza la revisión de totales y subtotales para cada variable dependiendo del dominio geográfico en términos de viviendas, hogares y personas. Así mismo, se generan las frecuencias de las principales variables y se realizan análisis de las respectivas distribuciones de los valores extremos y de los totales entre otros. Finalmente, como resultado de este proceso se realiza un reporte de inconsistencias que es enviado al equipo de temática, el cual solicita los respectivos ajustes y corrección de éstas cuando son necesarias.

- Imputación y ajustes de cobertura Una vez cumplidas las etapas de consistencia, depuración, validación de la información y revisión de frecuencias y cumplida la totalidad de los procesos estadísticos que garantizan la calidad y cobertura de la investigación, se realiza la conformación de la base solamente con las encuestas completas.

El componente temático de la encuesta hace entrega de los documentos con las especificaciones para la generación de cuadros de salida. Éstos son desarrollados y posteriormente se ejecutan los programas que permiten generarlos sobre la base de datos de encuestas completas y se efectúan los análisis de resultados, los cuales son generados en HTML o XLS.

Revisados los cuadros generados por parte del equipo temático, se realizan ajustes a los programas en caso de ser necesario, posteriormente se reciben los factores de expansión por parte del equipo de muestras y se generan los cuadros de salida definitivos, que permiten la visualización de los resultados de la encuesta.

Otros procesamientos

Diseño para la generación de cuadros de resultados

El diseño de los cuadros de resultados describe la totalidad de los procesos desarrollados para el llenado de cuadro de salidas a partir de programas o aplicativos informáticos, con base en el diseño de los cuadros y las especificaciones de procesamiento (a nivel de celda), de cada una de las variables o de los cruces de variables, desde la base de datos depurada de la operación estadística.

El proceso se desarrolla de la siguiente forma:

- El equipo temático de la encuesta envía las especificaciones para la generación de los cuadros de salida al equipo de sistemas. ? El grupo de sistemas programa estos resultados para la generación de cuadros en formato HTML o XLS. ? Revisados los cuadros generados por parte del equipo temático, se realizan ajustes a los programas, en caso de ser necesario. ? Se reciben los factores de expansión por parte del equipo de muestras y se generan los cuadros de salida definitivos, que permiten la visualización de los resultados de la encuesta.

- Los cuadros de salida se generan a nivel del total nacional, cabecera, centros poblados y rural disperso, y por departamentos.

Estimación de datos

Estimación de error de la muestra

DISEÑO DEL ANÁLISIS

Para el análisis y discusión de los resultados se tiene en cuenta el análisis estadístico y el análisis de contexto.

Métodos de análisis de resultados

- Análisis de consistencia: Este análisis incluye i) la comparación de resultados entre preguntas complementarias o relacionadas aplicadas en el mismo año, ii) el análisis del comportamiento de los principales indicadores básicos provenientes del módulo TIC de la Encuesta de Calidad de Vida ECV periodo 2023, para los dominios comunes (total nacional, cabecera, centros poblados y rural disperso) iii) y el contraste de los resultados por dominios (por ejemplo, cabecera frente a centros poblados y rural disperso, y departamentos entre sí). Lo anterior permite identificar tendencias básicas, presencia de valores válidos, información atípica y consistencia entre variables. Estos análisis utilizan técnicas estadísticas de chequeo que permiten tener una visión más amplia y robusta del comportamiento del fenómeno estudiado.

- Análisis de contexto: El análisis de contexto consiste en la comparación de los resultados obtenidos en la ENTIC Hogares con los arrojados por otras investigaciones adelantadas el mismo año o un año cercano. La principal operación estadística que se usa en esa comparación es la Encuesta de Calidad de Vida (ECV) del DANE, la cual cuenta con preguntas comparables a las incluidas en la ENTIC Hogares y que, por su aplicación continua, facilitaría contar con un período de referencia similar.

- Análisis de comparabilidad: Con el propósito de contrastar la coherencia y la consistencia de los resultados y realimentar los análisis realizados, se define el uso de estándares estadísticos y se tienen en cuenta procedimientos de comparabilidad con otras operaciones estadísticas similares que se produzcan en el SEN, en otros países, o de organismos estadísticos internacionales (por ejemplo, de la OCDE y la UIT).

- Análisis estadístico: El análisis estadístico de los resultados de la ENTIC Hogares se efectúa en dos etapas: Al inicio se realiza un análisis descriptivo a nivel de muestra. En ese análisis se examina la estructura de los indicadores por dominios de estudio a partir de la distribución de frecuencias y se detectan posibles inconsistencias y valores atípicos. Posteriormente, se verifica la cobertura en el nivel de segmento, viviendas, hogares y personas, y se observa la distribución de la pérdida de muestra para realizar los respectivos ajustes de la no respuesta. Después se realiza el análisis inferencial para estimar los parámetros de la población objetivo a partir de la muestra. En este proceso se aplican los factores de expansión y se revisa la inferencia a la respectiva población objetivo, la cual es establecida para los correspondientes dominios de estudio. Además, se verifica que los ajustes del factor de expansión no generen sesgos en las estimaciones y que sus errores muestrales sean aceptables de acuerdo con los parámetros establecidos por dominios de estudio.

- Análisis univariado En la ENTIC Hogares, se emplea principalmente distribuciones de frecuencias para las preguntas con varias opciones de respuesta, logrando presentarse como valores absolutos o en términos relativos, éstas suelen presentarse acompañadas de gráficos estadísticos (como histogramas, tortas, entre otras) para facilitar su comprensión y análisis.

- Análisis bivariado: Dentro del análisis estadístico de los datos, se tendrá en cuenta los límites (tanto inferior como superior) por medio de los cuales la estimación de los datos e indicadores puede oscilar, teniendo en cuenta, los coeficientes de variación de los datos calculados.

- Análisis multivariado Teniendo en cuenta que la producción de información estadística en la ENTIC Hogares es en su mayoría descriptiva, no se empleará este tipo de análisis.

Anonimización de microdatos

El equipo de temática determina las variables que por confidencialidad deben ser anonimizadas en la base de datos que se va a publicar y a partir de la base de datos final, es decir, únicamente con las encuestas completas, se construye un archivo que es remitido a la Oficina de sistemas. Quien luego de realizar ese proceso, reenvía el archivo al área temática para su verificación, aprobación y posterior publicación en la página web del DANE.

La ENTIC Hogares no incluye en la base de datos algunas variables correspondientes a la identificación de las viviendas como son la dirección, barrio, número telefónico y datos de identificación de las personas tales como nombre y apellido, fecha de nacimiento, nombre de la empresa donde trabaja, entre otros.

En las bases compartidas por la Sala de Procesamiento Especializado Externo (SPEE), el protocolo de anonimización consiste en eliminar las variables de identificación de las fuentes (tales como documento de identidad, barrio, nombre, dirección, teléfono, entre otras) y la información es publicada tal y como se recolectó, teniendo en cuenta que los usuarios firman acuerdos de confidencialidad en el manejo adecuado de la información estadística. Es fundamental destacar que la información disponible en la SPEE cumple con los protocolos de anonimización dispuestos por el DANE para garantizar la confidencialidad y uso de los datos sin violar la reserva estadística, de conformidad con el Código de Buenas Prácticas para las Estadísticas Oficiales. Además, cuenta con un equipo de profesionales encargados de verificar que los resultados generados por los investigadores preserven en todo momento la reserva estadística.

Verificación de la anonimización de microdatos

Para las bases que se comparten en la Sala de Procesamiento Especializado Externo (SPEE), el procedimiento consiste en verificar la eliminación de las variables de identificación de las fuentes (tales como documento de identidad, barrio, nombre, dirección, teléfono, entre otras).

Adicionalmente, se incluye la documentación pertinente para que el usuario pueda consultar la metodología y diccionario de datos.

Para eliminar el riesgo de identificación de las fuentes, en este caso los hogares y personas que suministran los datos para la generación de la información estadística, se debe llevar a cabo un proceso de anonimización que permita proteger la privacidad de las fuentes y preservar el aprovechamiento de los datos a disposición de los usuarios. Por lo anterior, este procesamiento se basa en el documento Guía para la anonimización de bases de datos en el Sistema Estadístico Nacional publicado por el DANE.

Comités de expertos

Se realizarán reuniones a través de una mesa de trabajo de la ENTIC Hogares, conformada por personas expertas externas e internas del DANE, con el fin de realizar seguimiento a los principales indicadores sociales comparando con fuentes o estadísticas producidas por otras entidades. Esto permitirá hacer contraste entre las diferentes metodologías, sus coberturas y los resultados que cada una arroja. Este es un proceso pertinente puesto que así se garantiza la calidad de la información y la participación de los usuarios en el proceso de análisis y socialización de resultados.

Del mismo modo, se cuenta con una mesa de trabajo permanente con el MINTIC, como entidad líder del sector y secretaria técnica de la Mesa Técnica de Estadísticas TIC, en donde se encuentran los principales actores que participan en el seguimiento y desarrollo de la política de tecnología en el país.

Otras formas de estimar datos

Errores muestrales

Son diversas las formas en las cuales se puede medir la precisión de los indicadores. En la encuesta se presentan los coeficientes de variación estimados (errores relativos) y los errores marginales de cada indicador. Los estimadores de total y razón se asumen con una distribución normal y sus medias y varianzas son las presentadas en la anterior sección.

El coeficiente de variación se define como la relación porcentual del error estándar del estimador y el valor esperado del estimador.

El valor de este coeficiente, expresado en porcentaje, permite evaluar la calidad de un procedimiento de estimación. Se consideran como excelentes las mediciones menores de 5%, buenas las que están entre 5% y 10% y aceptables las que están entre 10% y 15%.

El error marginal mide la amplitud media del intervalo de confianza. Con este valor se puede definir el límite inferior y superior del intervalo.