

Colombia - Gran Encuesta Integrada de Hogares - GEIH - Factores de Expansión Departamental 2007 - 2025

Dirección de Metodología y Producción Estadística - DIMPE

report_generated_on: May 5, 2026

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

Información general

Identificación

NÚMERO DE ID

DANE-DIMPE-GEIH-FACTORES-DE-EXPANSION-DEPARTAMENTAL-2021-2025

Información general

RESUMEN

El Departamento Administrativo Nacional de Estadística - DANE ha desarrollado encuestas de hogares desde finales de la década del 60 cuando adelantó encuestas de fuerza de trabajo e ingresos y gastos de manera transversal.

La recolección de la Gran Encuesta Integrada de Hogares empezó el 7 de agosto de 2006 en su módulo central de mercado laboral e ingresos y, a partir del 11 de septiembre, con su módulo de gastos de los hogares. A partir del 10 de julio de 2006 se amplió la cobertura de la ECH a 11 ciudades adicionales a las trece principales ciudades y áreas metropolitanas, al resto de cabeceras y al resto rural; cobertura que en la actualidad mantiene la GEIH.

En este documento podrá encontrar la evolución histórica de la medición del mercado laboral en Colombia y las principales características técnicas de la Gran Encuesta Integrada de Hogares.

La explicación del marco conceptual y diseño temático de la encuesta respaldada por recomendaciones de organismos internacionales que regulan y estandarizan los conceptos con el fin de asegurar la comparabilidad entre los países en materia de mercado laboral.

La descripción del diseño estadístico permite tener claridad sobre el marco muestral y los procedimientos de definición de la muestra y el alcance de la misma en términos de estimación y análisis de resultados. Así mismo el documento expone cada proceso de la producción estadística desde las actividades preparatorias al operativo de campo, pasando por la recolección, la transmisión y procesamiento de datos y los métodos de control de calidad, para finalizar con los procesos de análisis y difusión de los datos.

A este documento lo acompañan como anexos el formulario de la encuesta, el listado de cuadros disponibles al público por medio del banco de datos y el listado de clasificaciones internacionales adoptadas por la Gran Encuesta Integrada de Hogares.

TIPO DE DATO

Encuesta por muestreo (ssd)

UNIDAD DE ANÁLISIS

UNIDAD DE MUESTREO

La unidad de muestreo es la Medida de Tamaño, MT, o segmento. La medida de tamaño o segmento está conformada por un promedio de diez viviendas las cuales se indagan todas, se incluyen todos los hogares de cada vivienda y cada persona de cada hogar.

UNIDAD DE OBSERVACIÓN

La unidad de observación son aquellos conjuntos de elementos que componen el universo con una característica específica. En este caso la unidad básica de observación es el hogar.

Ámbito

NOTAS

A continuación se relacionan las variables más importantes de la operación estadística:

- Vivienda: Tipo de vivienda y características físicas (material de paredes y pisos).

- Datos del hogar: Conexión a servicios públicos, privados - comunales, valor pagado por consumo y calidad de los mismos, conexión y uso de servicio sanitario, obtención de agua para consumo, lugar y energía para preparar alimentos, eliminación de basuras, tipo de tenencia de la vivienda, tenencia de bienes en el hogar.
- Registro de personas: Identificación del residente habitual.
- Características generales: Sexo, edad, parentesco, estado civil.
- Seguridad social en salud: Cobertura del Sistema General de Seguridad Social en Salud -SGSSS- por regímenes, persona que paga afiliación y cobertura.
- Educación: alfabetismo, asistencia escolar, máximo nivel educativo alcanzado y último año aprobado o que esté cursando y títulos o diplomas obtenidos.
- Fuerza de trabajo: PEA (ocupados y desocupados)
- Ocupados:
 - Empleo principal (rama de actividad, ocupación, tipo de contrato, acceso a prestaciones, tiempo trabajado y posición ocupacional).
 - Asalariados (medios de búsqueda, remuneración mensual, horas extras, pagos en especie, subsidios, primas y bonificaciones).
 - Independientes (formas de trabajo, registro mercantil, contabilidad, ganancia u honorarios netos).
 - Asalariados e independientes (duración del empleo, horas normales y efectivas trabajadas, honorarios, tamaño de la empresa, sitio de trabajo, Afiliación a pensiones, caja de compensación familiar y ARP, duración entre empleo anterior y actual).
 - Empleo secundario (horas trabajadas, posición ocupacional, remuneración mensual, tamaño de la empresa y sitio de trabajo)
 - Empleo con insuficiencias de horas y situaciones de empleo inadecuado (por competencias e ingresos).
 - Calidad del empleo
- Desocupados: Duración de búsqueda de trabajo, historial laboral, ingresos y seguridad social.
- Inactivos: trayectoria laboral, ingresos y seguridad social.
- Otras actividades. Actividades no remuneradas y ayudas
- Otros Ingresos. Otros ingresos diferentes a los ingresos laborales
- Fecundidad. Cantidades de hijos nacidos vivos y la influencia de la paternidad o maternidad en el mercado laboral
- Migración. Dinámicas migratorias a nivel nacional de la población objetivo
- Trabajo Infantil. Actividades y razones del trabajo infantil
- Formación para el trabajo. Dinámicas de formación para el trabajo, sus características y participación en procesos de certificación de competencias a nivel nacional de la población objetivo

TOPICS

Topic	Vocabulary	URI
Ingresos, propiedad y ahorro/inversión	CESSDA	http://www.nesstar.org/rdf/common
Empleo	CESSDA	http://www.nesstar.org/rdf/common
Desempleo	CESSDA	http://www.nesstar.org/rdf/common

Situación económica e indicadores	CESSDA	http://www.nesstar.org/rdf/common
Economía rural	CESSDA	http://www.nesstar.org/rdf/common
Formación continua [6.4]	CESSDA	http://www.nesstar.org/rdf/common
Vivienda [10.1]	CESSDA	http://www.nesstar.org/rdf/common
Fertilidad [14.2]	CESSDA	http://www.nesstar.org/rdf/common
Disponibilidad y utilización de servicios sociales específicos [15.3]	CESSDA	http://www.nesstar.org/rdf/common

KEYWORDS

Asalariados, Desocupados (D):, Desempleo abierto, Desempleo oculto, Empleo asalariado:, Empleo total, Empleos independientes, Horas-hombre trabajadas:, Horas extras, Horas ordinarias, Ocupados (OC):, Ocupación, Población económicamente activa (PEA):, Población económicamente inactiva (PEI):, Población en edad de trabajar (PET), Población objetivo:, Población total (PT), Porcentaje de la población en edad de trabajar (PET), Prestaciones sociales:, Subempleo objetivo:, Subempleo por insuficiencia de horas, Subempleo por situaciones de empleo inadecuado:, Subempleo subjetivo, Subempleo y condiciones de empleo inadecuado:, Tasa de subempleo (TS):, Tasa bruta de participación (TBP), Tasa de desempleo (TD), Tasa de ocupación (TO):, Tasa global de participación (TGP), Trabajador familiar:, Trabajador por cuenta propia

Cobertura**COBERTURA GEOGRÁFICA**

Tiene cobertura nacional que permite obtener resultados para cabeceras y resto, ciudades y áreas metropolitanas, grandes regiones y total por departamento.

GEOGRAPHIC UNIT

Tiene cobertura nacional que permite obtener resultados por zona urbana y rural, cinco grandes regiones (Región Atlántica, Región Oriental, Región Central, Región Pacífica y Región Bogotá) y total por departamentos: 23 departamentos, las 13 grandes ciudades con sus áreas metropolitanas y 11 ciudades intermedias, que son en su orden:

13 GRANDES CIUDADES CON SUS ÁREAS METROPOLITANAS :

Bogotá

Medellín - Valle de Aburrá

Cali - Yumbo

Barranquilla - Soledad

Bucaramanga - Floridablanca - Girón - Piedecuesta

Manizales - Villamaría

Pasto

Pereira - Dosquebradas - La Virginia

Ibagué

Cúcuta - Villa del Rosario - Los Patios - El Zulia

Villavicencio

Montería

Cartagena

11 CIUDADES INTERMEDIAS :

Tunja

Florencia

Popayán

Valledupar

Quibdó

Neiva

Riohacha

Santa Marta

Armenia

Sincelejo

San Andrés.

UNIVERSO

El universo para la Gran Encuesta Integrada de Hogares está conformado por la población civil no institucional, residente en todo el territorio nacional.

Productores y Patrocinadores

INVESTIGADOR(ES) PRIMARIO(S)

Nombre	Affiliation
Dirección de Metodología y Producción Estadística - DIMPE	

OTROS PRODUCTORES

Nombre	Affiliation	Role
Dirección de Metodología y Producción Estadística - Gran Encuesta Integrada de Hogares	Departamento Administrativo Nacional de Estadística - DANE	Equipo Técnico

FINANCIAMIENTO

Nombre	Abbreviation	Role
Departamento Administrativo Nacional de Estadística	DANE	Ejecutor

Producción de metadato

METADATO PRODUCIDO POR

Nombre	Abbreviation	Affiliation	Role
Ivan Rolando Castillo Prieto	ircastillop@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización- DIRPEN	Coordinador Regulación
Diego Alejandro Martínez Jimenez	damartinezj@dane.gov.co	Dirección de Metodología y Producción Estadística - DIMPE	Documentador GEIH
Carlos Augusto Muñoz Alfonso	camunoza@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización- DIRPEN	Actualizador DIRPEN

Katherine Andrea Rojas Acero	karojasa@dane.gov.co	Dirección de Difusión y Cultura Estadística - DICE	Verificador DICE
------------------------------------	----------------------	--	------------------

FECHA DE PRODUCCIÓN DEL METADATO

2025-05-08

ID DEL DOCUMENTO IDD

COL-DANE-GEIH-FACTORES-DE-EXPANSION-DEPARTAMENTAL-2021-2025

Muestreo

Procedimiento de muestreo

MARCO MUESTRAL

Está constituido por el inventario cartográfico y el listado de viviendas obtenidas de la Muestra Maestra de Hogares, con actualizaciones permanentes y nuevos recuentos de edificaciones y viviendas a través de la misma.

TIPO DE MUESTREO

Teniendo en cuenta los objetivos y las características del marco, se optó por una muestra probabilística, estratificada, de conglomerados, multietápica, los cuales se presentan a continuación:

- PROBABILÍSTICA

Cada unidad de la población objetivo tiene una probabilidad de selección conocida y superior a cero. Este tipo de muestra permite establecer anticipadamente la precisión deseada en los resultados principales, y calcular la precisión observada en todos los resultados obtenidos.

- ESTRATIFICADA

Este método asegura una mejor precisión de la muestra, al disminuir la varianza de las estimaciones, los criterios de estratificación son:

- El primer estrato corresponde a las 24 capitales y áreas metropolitanas con dominios de estudio independientes. Cada ciudad o área metropolitana es autorrepresentada. Para la estratificación y selección de la muestra, las áreas geográficas se organizaron, según los principios cartográficos establecidos, en sectores, secciones y manzanas con la información del número de viviendas a nivel de manzana y el estrato socioeconómico.

- El segundo estrato corresponde al resto urbano y al resto de cabeceras, centros poblados y rural disperso. Los municipios se estratificaron de acuerdo con los siguientes criterios:

- Geográficos, a nivel de regiones constituidas por varios departamentos.

- Socioeconómicos, a nivel municipal, con los siguientes indicadores:

- Nivel de urbanización, en términos de la cantidad de población de las cabeceras municipales.

- Estructura urbano-rural de la población municipal (% de población en cabecera). Proporción de la población con necesidades básicas insatisfechas (NBI), la cual, a su vez, se clasificó en 4 intervalos: A [0-29,8%]; B [29,8%-42,7%]; C [42,7%-57,3%]; y D [57,3%-100%].

- Tamaño poblacional del estrato.

Cada municipio con 7.000 o más habitantes en su población total, se constituyó en Unidad Primaria de Muestreo, UPM. Los de menor población se agruparon con uno vecino para constituirse en UPM.

Las UPM cuyo tamaño era similar al tamaño promedio de los estratos, se constituyeron en estratos de "certeza"; para efectos probabilísticos tuvieron probabilidad 1. Las demás UPM se agruparon en estratos de no certeza, aplicando las variables de estratificación en el orden de prioridad especificado antes.

DE CONGLOMERADOS

Corresponde a la unidad final de muestreo, que es la medida de tamaño o segmento; es el área que contiene un promedio

de diez viviendas, en la cual se investigan todas las viviendas, todos los hogares y todas las personas.

La concentración de la muestra en conglomerados presenta grandes beneficios en el nivel operativo; lo cual reduce, en forma considerable, los tiempos y los costos, y facilita la supervisión.

MULTIETÁPICA

Unidades Primarias de Muestreo (UPM). Hay UPM de inclusión forzosa y de inclusión probabilística:

- Son de inclusión forzosa las áreas metropolitanas y el resto de capitales departamentales que son autorrepresentadas, pues su tamaño muestral permite estimaciones trimestrales separadas para cada una de ellas, y aquellos municipios de tamaño intermedio cuya población es similar o mayor que el tamaño promedio de los estratos de su departamento.

- Son de inclusión probabilística las UPM seleccionadas con PPT de los estratos conformados por dos o más UPM (estratos de no certeza). La técnica de selección controlada utilizada permitió mediante un procedimiento probabilístico incrementar las probabilidades de selección de las combinaciones preferidas de UPM que podían ser seleccionadas en un departamento, adicionalmente a lo obtenible con un muestreo estratificado corriente.

En el caso de la Muestra Maestra la selección controlada proporcionó un mejor balance interregional en las características de las UPM seleccionado. El resultado final de este ejercicio fue lograr la máxima reducción posible en el error estándar de estimación de los resultados.

Unidades Secundarias de Muestreo (USM):

En las cabeceras municipales y centros poblados, las USM son grupos de manzanas contiguas del mismo municipio, que contienen un mínimo de 12 medidas de tamaño (MT) con un promedio de 10 viviendas, y en la zona rural dispersa, la USM es una sección o sector censal. El tamaño de la USM permite un control adecuado de la distribución y rotación de la muestra con cada encuesta. Para la selección de las USM se aplica el siguiente procedimiento (PPT):

- Conformación, delimitación y listado de las USM por estrato socioeconómico (solo cabeceras municipales) de cada municipio seleccionado, siguiendo el orden de la nomenclatura de la cartografía censal: sector, sección y manzana.
- Cálculo del número de medidas de tamaño (MT), dividiendo por 10 (el número de viviendas) y aproximando a un valor entero.
- Cálculo de intervalo de muestreo (IM), dividiendo el total de MT existentes en una zona por el número de USM requeridas en la muestra de la zona.
- Selección de un número de arranque entre 1 y el valor del intervalo, el cual indica la primera USM de la muestra de la zona; y selección del resto de USM de la zona, sumando sucesivamente el IM al número de arranque.

Unidades Terciarias de Muestreo (UTM). El tercer paso es la selección dentro de cada USM de una o más UTM, dependiendo de las requeridas, del tamaño y cantidad de USM preseleccionadas en cada municipio.

Cada UTM es un segmento con un promedio de 10 viviendas contiguas (5 a 14 viviendas), es decir, una medida de tamaño (MT). En un paso intermedio, se seleccionan manzanas en las cabeceras municipales, centros poblados y segmentos de sección rural claramente delimitados, con límites naturales (quebradas, ríos, etc.) o sociales (carreteras, caminos, etc.). Los segmentos de los sectores rurales se conforman previamente con base en fotografías aéreas y en lo posible, se procura que tengan un promedio de 10 viviendas; sin embargo, por insuficiente detalle cartográfico, los segmentos pueden tener dos o más MT.

La selección se hace con probabilidad proporcional al número de MT (PPT), así una manzana o segmento rural puede contener una o más MT, y las que tienen menos de 5 viviendas se juntan con una manzana vecina del mismo estrato socioeconómico.

Las áreas urbanas no urbanizadas se anexan al área o manzana vecina muy cercana con una o más viviendas con el objeto de que tengan probabilidad de selección y para que puedan captar cualquier proceso futuro de urbanización. Si solo se requiere una MT y la manzana o porción solo tiene una MT, la muestra queda constituida por todas las viviendas y hogares existentes. Si la manzana o segmento rural tienen dos o más MT se procede al recorrido y delimitación de segmentos de una MT, en el terreno. Finalmente, se hace la asignación aleatoria del orden en que entrarán los segmentos al proceso de

rotación.

Unidades Cuartas de Muestreo (UCM). Cada UCM es un segmento de 10 viviendas contiguas (5 a 14 viviendas), es decir una medida de tamaño MT y su selección es aleatoria.

Finalmente se hace una asignación aleatoria del orden en que entrarán los segmentos al proceso de rotación. En este proceso de selección por etapas, solo se requiere una actualización cartográfica detallada de las áreas seleccionadas.

Para el estrato de resto de cabeceras, centros poblados y rural disperso se seleccionó:

- En una primera etapa, la UPM, utilizando la técnica de selección controlada dentro de cada estrato.
- Para la segunda etapa se seleccionó en la zona urbana la manzana, y en la al resto de cabeceras, centros poblados y rural disperso la sección, o sea la USM.
- En la tercera etapa se seleccionó el segmento o UTM.

En este proceso de selección por etapas sólo se requiere una actualización cartográfica detallada de las áreas seleccionadas.

DEFINICIÓN TAMAÑO DE LA MUESTRA

Inicialmente el tamaño de la muestra mensual correspondía aproximadamente a 23 000 hogares. En el año 2000 con la implementación de la ECH la muestra maestra fue ampliada de 165 a más de 240; con 30 000 hogares en 13 áreas y 7 500 en el resto de cabeceras, centros poblados y rural disperso. Durante 2004 se amplió la muestra maestra, con un total de 44 400 hogares, con 30 000 hogares en 13 áreas y 14 400 en el resto de cabeceras, centros poblados y rural disperso. Para el año 2006 con la implementación de la GEIH, se amplió la muestra a 11 ciudades más; con 17 600 hogares adicionales, para un total de 62 000 hogares. Con el nuevo marco generado por el censo 2005, la nueva muestra implementada desde 2009 está conformada por 437 municipios y anualmente se visitan aproximadamente 248.028 hogares, concentrados en 22.548 segmentos. La muestra mensual es de 20.669 hogares, 18.790 viviendas y 1.879 segmentos.

Los tamaños de muestra se calculan con una precisión deseada de la variable tasa de desempleo no superior a un error estándar relativo del 5% y una tasa de desempleo del 10%. Los cálculos se realizan con las fórmulas correspondientes al tipo de diseño muestral. Se ajustan con base en el efecto de los conglomerados en el diseño (deff), que es una relación, para cada dominio, entre la varianza real de este diseño de conglomerados y la que se obtendría con un diseño aleatorio simple de elementos.

Tasa de respuesta

COMPONENTES DEL FACTOR DE EXPANSIÓN

- Factor básico de expansión (F). Aplicado a los datos muestrales, da a cada elemento de la muestra el peso o representación que le corresponde en el universo investigado. En consecuencia, mediante su aplicación, se estiman numéricamente, en forma aproximada, las características de la población objetivo.
- Peso de submuestreo (Ph). Está dado por segmento y es teóricamente igual a 1 para todos los segmentos en razón de que representan una medida de tamaño. El desarrollo dinámico del marco de muestreo en algunos casos puede ser mayor o menor que 1, de acuerdo con la densidad de viviendas del segmento en el momento de hacer la encuesta, y por tal motivo modifica el factor básico de expansión en el segmento.
- Ajuste de cobertura por no respuesta (Rh). Cuando las tasas de no respuesta varían en los subgrupos de la población de diferentes características, el ajuste normal es asignar a los hogares y a las personas no encuestadas el promedio de las características de los encuestados en el mismo segmento. Esto se logra corrigiendo el factor básico de expansión por un nuevo factor resultante de la razón entre el número de hogares seleccionados en un segmento y el número de hogares encuestados en el mismo segmento.
- El factor final de expansión (Wh). Es el producto de los anteriores y se aplica a cada segmento de la ciudad o área investigada.

AJUSTE POR LOS VALORES POBLACIONALES

Las estimaciones del universo de estudio pueden ser mejoradas, si se equiparan a nivel de las desagregaciones geográficas los totales de la población obtenidos de la muestra expandida, con los totales del censo de población proyectados a la fecha de la encuesta. Dicho mejoramiento se basa en la premisa de que la estructura por desagregaciones geográficas, proyectada a partir de un censo de población reciente, es más exacta que la estimada a partir de la muestra.

Se aplican los estimadores de regresión utilizando la información auxiliar correspondiente a las Proyecciones de Población del CENSO 2005, en este caso, se ha usado la información auxiliar por grupos de edad y sexo. Este proceso se realiza a través de los Métodos de Calibración que son procedimientos que utilizan información auxiliar relacionada con las variables de estudio, con el fin de mejorar la precisión y reducir los sesgos en las estimaciones.

ESTIMADOR DEL TOTAL

Para aplicar el método de calibración se establecieron los siguientes ocho grupos:

SEXO:

- Hombres
- Mujeres

Y para cada uno los siguientes grupos de EDAD:

- De 0 a menores de 12 años (para el caso urbano, en lo rural 10 años).
- 12 a menores de 25 años
- 25 a menores de 55 años
- 55 años y más

Con la información del total de personas en las proyecciones de población, en cada grupo de calibración se construye el vector de totales.

Donde cada variable x da la cantidad de personas en cada grupo de la vivienda. (Se hace a nivel de vivienda ya que es la unidad de análisis más agregada, de esta manera a la vivienda le corresponde un único factor de expansión que es el mismo de los hogares y personas que viven en ella).

El cálculo del ajuste por calibración se realizó utilizando la macro Clan 97 v3.1 software creado por Statistics Sweden que corresponde a un conjunto de rutinas en el programa SAS para la estimación puntual y de errores en encuestas por muestreo.

ESTIMADOR DE RAZÓN

Las tasas, razones, proporciones y promedios, generadas a partir de este diseño muestral son de la forma de una razón, en la cual el numerador y el denominador son variables aleatorias.

ESTIMADOR DE LA VARIANZA PARA UNA RAZÓN

Para estimar la varianza del estimador de una razón se debe transformar la variable en una función lineal, para esto se utiliza el método de Linealización de Taylor.

Ponderación

Se diseñó una muestra para obtener una precisión esperada medida en términos del error de muestreo menor o igual a 5% con un nivel de confiabilidad del 95%, para los principales indicadores de fuerza laboral nacionales.

Formularios

Información general

DISEÑO DE INSTRUMENTOS

El instrumento desarrollado para la recolección de información de la GEIH, cuenta con una serie de preguntas e incorpora los siguientes capítulos y módulos

LETRA: A

NOMBRE DEL CAPÍTULO: Identificación

NÚMERO DE PREGUNTAS: 19

LETRA: B

NOMBRE DEL CAPÍTULO: Vivienda

NÚMERO DE PREGUNTAS: 5

LETRA: C

NOMBRE DEL CAPÍTULO: Datos del hogar

NÚMERO DE PREGUNTAS: 15

LETRA: D

NOMBRE DEL CAPÍTULO: Registro de personas

NÚMERO DE PREGUNTAS: 4

LETRA: E

NOMBRE DEL CAPÍTULO: Características generales

NÚMERO DE PREGUNTAS: 8

LETRA: F

NOMBRE DEL CAPÍTULO: Seguridad social en salud

NÚMERO DE PREGUNTAS: 7

LETRA: G

NOMBRE DEL CAPÍTULO: Educación

NÚMERO DE PREGUNTAS: 5

LETRA: H

NOMBRE DEL CAPÍTULO: Fuerza de trabajo

NÚMERO DE PREGUNTAS: 14

LETRA: I

NOMBRE DEL CAPÍTULO: Ocupados

NÚMERO DE PREGUNTAS: 65

LETRA: J

NOMBRE DEL CAPÍTULO: Desocupados

NÚMERO DE PREGUNTAS: 14

LETRA: K

NOMBRE DEL CAPÍTULO: Inactivos

NÚMERO DE PREGUNTAS: 10

LETRA: L

NOMBRE DEL CAPÍTULO: Otras actividades

NÚMERO DE PREGUNTAS: 1

LETRA: M

NOMBRE DEL CAPÍTULO: Ingresos no laborales

NÚMERO DE PREGUNTAS: 4

LETRA: N

NOMBRE DEL CAPÍTULO: Módulo de Micronegocios

NÚMERO DE PREGUNTAS: 6

LETRA: O

NOMBRE DEL CAPÍTULO: Módulo Formación para el trabajo *

NÚMERO DE PREGUNTAS: 16

LETRA: P

NOMBRE DEL CAPÍTULO: Módulo de Fecundidad

NÚMERO DE PREGUNTAS: 11

LETRA: Q

NOMBRE DEL CAPÍTULO: Módulo de Migración

NÚMERO DE PREGUNTAS: 7

LETRA: R

NOMBRE DEL CAPÍTULO: Módulo de Trabajo Infantil **

NÚMERO DE PREGUNTAS: 23

*A partir de 2013 aplica en el segundo trimestre de cada año.

** A partir de 2012 aplica en el cuarto trimestre de cada año.

El equipo de Temática Social elabora un Manual de diligenciamiento y conceptos básicos, que acompaña al formulario, el cual es guía para los recolectores en campo.

Recolección de datos

Fecha de recolección de datos

Start	End	Cycle
2021	2021	N/A
2022	2022	Semanal, Mensual, Trimestral y Anual
2023	2023	Semanal, Mensual, Trimestral y Anual
2024	2024	Semanal, Mensual, Trimestral y Anual
2025	2025	Semanal, Mensual, Trimestral y Anual

Time Periods

Start	End	Cycle
2025		Semanal, Mensual, Trimestral y Anual

Modo de recolección de datos

Entrevista personal asistida con DMC (dispositivo móvil de captura)

Notas de recolección de datos

SENSIBILIZACIÓN

El objetivo del Sistema de Sensibilización, es lograr una respuesta oportuna y de calidad por parte de las fuentes de información para la investigación estadística al DANE. Esta busca lograr un ambiente de colaboración, acercamiento y conocimiento de la población objetivo mediante la visita a cada uno de los hogares a los que les será aplicada la encuesta, por parte de los sensibilizadores, con el fin de informarles que el hogar ha sido seleccionado para ser encuestado. En esta primera visita, se pueden establecer citas y horarios en los cuales los encuestadores del DANE pueden aplicar la encuesta.

CAPACITACIÓN

Cada Dirección Territorial y subsele capacita al personal convocado en temas como: cartografía, recuento, recolección, conceptos básicos y manejo del Dispositivo Móvil de Captura (DMC), realiza los exámenes escritos y de campo que sean necesarios y selecciona el personal. Para cada uno de los cursos de capacitación deben aplicarse evaluaciones de: asistencia al curso, prueba de conocimientos y manejo del flujo del cuestionario.

SELECCIÓN DE PERSONAL

En la página web del DANE Central se publican los términos de referencia. El DANE Central envía a las Direcciones Territoriales y subsele el nombre de las personas que se inscribieron; luego las Direcciones Territoriales y subsele reciben las hojas de vida de estas personas y verifican si cumplen el perfil requerido. Como criterios para la selección del personal se evalúa el grado de aprehensión, manejo de los conceptos y de los flujos del cuestionario. Finalmente las personas que cumplen el perfil requerido son llamadas para capacitación mediante publicación en la página web.

Se hacen evaluaciones sobre todos los conceptos temáticos y del trabajo de campo, y de igual forma se hacen simulaciones de encuestas para medir la destreza en el trabajo de campo que hace parte de la evaluación definitiva.

Por último se realiza una prueba de campo, la cual es una evaluación que consiste en que cada uno de los que están participando en el curso, van a terreno y realizan una o dos encuestas, las cuales se califican y hacen parte de la evaluación. Esta prueba de campo tiene las mismas características de la recolección, es decir, participan recolectores, supervisores y sensibilizadores.

ESQUEMA OPERATIVO, MÉTODO Y PROCESAMIENTO PARA LA RECOLECCIÓN

Para manejar eficientemente el trabajo de campo, se conforma un equipo operativo de carácter temporal que funciona con base en cuatro niveles jerárquicos:

Asistente técnico Gran Encuesta Integrada de Hogares. A nivel regional es el responsable de la organización operativa de la encuesta; de él dependen los grupos de trabajo adscritos a su Dirección Territorial, con quienes debe mantener comunicación permanente.

Coordinador de campo Gran Encuesta Integrada de Hogares. En las Direcciones Territoriales del DANE, es el responsable de las actividades preliminares de la encuesta y la supervisión a los grupos de trabajo en campo, y en las subsedes tiene a su cargo el manejo de la encuesta en todos sus aspectos administrativos y operativos.

Apoyo Logístico Gran Encuesta Integrada de Hogares. El Apoyo Logístico colabora con las actividades de recuento y logísticas, en la elaboración de los informes operativos. Es el encargado de la revisión, preparación y distribución de los requerimientos para la ejecución de los procesos de recolección de la información y, de la realización de por lo menos una re-encuesta semanal a un hogar perteneciente a cada grupo de trabajo, con fin de confirmar si la información recolectada es consistente.

Apoyo informático. En las Direcciones Territoriales del DANE, es el encargado de transmitir la información y realizar el cruce de segmentos faltantes.

Supervisores. Coordinan y controlan el trabajo en los segmentos en la ciudad y en los municipios que se les asignen, y reportan su trabajo al coordinador de campo de la encuesta.

Encuestador. Conocido también como recolector, es la persona encargada de obtener la información requerida, digitándola en el DMC conforme a las normas y los conceptos establecidos. La labor del encuestador es exigente y requiere un conocimiento preciso de los conceptos y las normas que rigen la encuesta, además de un compromiso orientado a lograr la cobertura y los niveles de calidad propuestos.

Para la ejecución del trabajo de campo, se utiliza un fólder por segmento (carga de trabajo) denominado "carpeta de muestra", en cuya carátula se encuentra la identificación del área, que se va a trabajar el cual contiene:

- Un plano (cuando así se requiera) del municipio o del segmento en que se encuentra ubicado para orientación del grupo de trabajo.
- El formato de recuento GEIH1.
- Listado de edificaciones y viviendas GEIH2.

La encuesta utiliza un informante directo para las personas de 18 años y más, y para aquellas de 10 a 17 años que trabajen o estén buscando trabajo. Para los demás se acepta informante idóneo (persona del hogar mayor de 18 años, que a falta del informante directo pueda responder correctamente las preguntas). No se acepta información de empleados del servicio doméstico, pensionistas, vecinos o menores, excepto cuando el menor de edad es el jefe del hogar o cónyuge.

El sistema utilizado para la recolección es el de barrido y rutas, para el resto de cabeceras, centros poblados y rural disperso.

Formularios

DISEÑO DE INSTRUMENTOS

El instrumento desarrollado para la recolección de información de la GEIH, cuenta con una serie de preguntas e incorpora los siguientes capítulos y módulos

LETRA: A
NOMBRE DEL CAPÍTULO: Identificación
NÚMERO DE PREGUNTAS: 19

LETRA: B
NOMBRE DEL CAPÍTULO: Vivienda
NÚMERO DE PREGUNTAS: 5

LETRA: C

NOMBRE DEL CAPÍTULO: Datos del hogar

NÚMERO DE PREGUNTAS: 15

LETRA: D

NOMBRE DEL CAPÍTULO: Registro de personas

NÚMERO DE PREGUNTAS: 4

LETRA: E

NOMBRE DEL CAPÍTULO: Características generales

NÚMERO DE PREGUNTAS: 8

LETRA: F

NOMBRE DEL CAPÍTULO: Seguridad social en salud

NÚMERO DE PREGUNTAS: 7

LETRA: G

NOMBRE DEL CAPÍTULO: Educación

NÚMERO DE PREGUNTAS: 5

LETRA: H

NOMBRE DEL CAPÍTULO: Fuerza de trabajo

NÚMERO DE PREGUNTAS: 14

LETRA: I

NOMBRE DEL CAPÍTULO: Ocupados

NÚMERO DE PREGUNTAS: 65

LETRA: J

NOMBRE DEL CAPÍTULO: Desocupados

NÚMERO DE PREGUNTAS: 14

LETRA: K

NOMBRE DEL CAPÍTULO: Inactivos

NÚMERO DE PREGUNTAS: 10

LETRA: L

NOMBRE DEL CAPÍTULO: Otras actividades

NÚMERO DE PREGUNTAS: 1

LETRA: M

NOMBRE DEL CAPÍTULO: Ingresos no laborales

NÚMERO DE PREGUNTAS: 4

LETRA: N

NOMBRE DEL CAPÍTULO: Módulo de Micronegocios

NÚMERO DE PREGUNTAS: 6

LETRA: O

NOMBRE DEL CAPÍTULO: Módulo Formación para el trabajo *

NÚMERO DE PREGUNTAS: 16

LETRA: P

NOMBRE DEL CAPÍTULO: Módulo de Fecundidad

NÚMERO DE PREGUNTAS: 11

LETRA: Q

NOMBRE DEL CAPÍTULO: Módulo de Migración

NÚMERO DE PREGUNTAS: 7

LETRA: R

NOMBRE DEL CAPÍTULO: Módulo de Trabajo Infantil **

NÚMERO DE PREGUNTAS: 23

*A partir de 2013 aplica en el segundo trimestre de cada año.

** A partir de 2012 aplica en el cuarto trimestre de cada año.

El equipo de Temática Social elabora un Manual de diligenciamiento y conceptos básicos, que acompaña al formulario, el cual es guía para los recolectores en campo.

Data Collectors

Nombre	Abbreviation	Affiliation
Departamento Administrativo Nacional de Estadística	DANE	Gobierno Nacional

Supervisión

INSTRUMENTOS DE CONTROL PARA SUPERVISIÓN

La recolección de la información está bajo la dirección del supervisor, quien fundamentalmente controla el trabajo de campo que realizan los encuestadores, y cuyos resultados dependen en gran parte de la iniciativa para la solución de problemas y dificultades, del liderazgo y la capacidad organizativa del trabajo.

Las actividades de control que debe llevar a cabo un supervisor son:

- Control de desempeño de las actividades diarias del encuestador.
- Control de calidad, con el objetivo de verificar la calidad y consistencia de la información recolectada mediante la verificación de conocimientos de conceptos del encuestador o por medio de reentrevistas.
- Control de visitas dirigido a revisar que el encuestador haya concurrido directa y personalmente a los hogares de las viviendas que debe entrevistar, y el cumplimiento de las citas concertadas para lograr la encuesta completa.
- Control de cobertura, cuyo propósito es controlar que los hogares de las viviendas encuestadas correspondan exactamente a las viviendas seleccionadas en la muestra; y en caso de rechazo por parte del hogar, tratar de persuadir al informante de la importancia de la investigación, recordando su confidencialidad para que suministre los datos.
- Como parte de las funciones del supervisor también se encuentra la evaluación de la entrevista, para verificar el buen manejo de los DMC por parte del encuestador y seguimiento a las instrucciones de recolección dadas para recoger la información de mercado laboral.

Procesamiento de datos

Edición de datos

CONSOLIDACIÓN DE ARCHIVOS

Los datos provenientes del área geográfica en donde se aplica la investigación se consolidan diariamente, y pueden ser monitoreados por el área temática y sistemas

Para enviar la información al DANE Central, se utiliza la herramienta Swin que genera un archivo comprimido asignándole un nombre, que indique la investigación, ciudad y fecha a que corresponde la información contenida; luego se copia al servidor destinado para el proceso de transmisión, utilizando el protocolo FTP en la carpeta asignada para ello y confirmar inmediatamente que la información está en dicho servidor, se envía mediante correo electrónico.

De igual forma, en el DANE Central se realizan las copias de respaldo necesarias y se aplican los planes de contingencia que garanticen el normal funcionamiento del operativo.

NORMAS DE VALIDACIÓN Y CONSISTENCIA

Validación de rangos de acuerdo con la estructura de las preguntas. Cuando la pregunta tiene predefinidos las opciones de respuesta, es necesario controlar las opciones a mostrar dadas las restricciones que existen sobre unicidad o valor máximo posible de ocurrencia.

Validación de los universos. A este proceso lo definen tres aspectos: el primero, cuando la pregunta define un flujo o salto dependiendo de la opción; el segundo cuando los datos de la vivienda se toman una sola vez, así exista más de un hogar; el tercero, cuando se define en el universo según la edad de la persona residente en el hogar.

Para algunas variables relacionadas con la aplicación de conceptos básicos de la Gran Encuesta Integrada de Hogares que son determinantes en la calidad de la estructura de la información, el DMC permite que el entrevistador confirme la respuesta que da el entrevistado. Estas variables son: Número de hogares en la vivienda; nombres y apellidos; sexo; edad; parentesco; y todas las variables relacionadas con valores de los diferentes formularios.

En el diseño del programa se garantiza que en la misma pantalla se muestren el capítulo que está relacionada con las variables como las diferentes preguntas que por su alta correlación debe controlarse su consistencia.

Los anteriores aspectos hacen parte del documento "Especificaciones de normas técnicas de validación y consistencia" que tienen de forma detallada cada una de las preguntas de los formularios.

VERIFICACIÓN DE LA CONSISTENCIA INTERNA DE LOS DATOS Y AJUSTES

El proceso para controlar e identificar los datos faltantes, inválidos o inconsistentes se lleva a cabo mediante el análisis de la variable "Incompleta" que toma valor 0 cuando la encuesta se encuentra completa y se puede utilizar directamente de la base de datos, y toma valor 1 cuando la encuesta debe ser revisada por datos faltantes inválidos o inconsistentes.

IMPUTACIÓN Y/O AJUSTES DE COBERTURA

El procesamiento es centralizado por DANE Central, e inicia desde el momento en que las Direcciones Territoriales envían la información; los datos son transmitidos por FTP (File Transfer Protocol) y posteriormente cargados en la base de datos Oracle.

En esta operación estadística no se realiza el proceso de imputación, cuando los miembros de los hogares se niegan a responder la encuesta, esta no respuesta es ajustada con los factores de expansión, lo que se explica en el numeral 3.6.1 Componentes del factor de expansión

Estimación de datos

Estimación de error de la muestra

ANÁLISIS ESTADÍSTICO

Los principales indicadores que se utilizan en el análisis estadístico de los resultados son:

Cobertura de viviendas: se obtiene de cruzar la información de viviendas iniciales seleccionadas, contra viviendas finales encontradas.

Cobertura de hogares: se obtiene de cruzar la información del total de encuestas completas contra el total de hogares encontrados.

Calidad de la recolección: se obtiene de identificar el número de errores cometidos durante la recolección.

Errores e inconsistencias: se obtiene de la realización de diversas pruebas que ayudan a constatar que los cálculos estén bien hechos y que los datos sean coherentes.

Alarmas: se obtienen al hacer una búsqueda de datos que se salgan de los estándares, como: alto número de viviendas vacantes, alto número de rechazos, bajo promedio de personas, entre otros.

Informe de cobertura campo vs Sistemas: se obtiene de cruzar la información que las sedes y subsedes envían al DANE Central en los resúmenes de cobertura sobre viviendas, hogares y personas encontradas, contra la información que es transmitida al área de sistemas.

Indicador de puntualidad: busca garantizar que la información enviada por las sedes y subsedes sea oportuna.

Tasa de respuesta: se obtiene al medir en número de encuestas efectivas respecto a las esperadas. El objetivo de este indicador es identificar el grado de eficacia en el que se encuentra la investigación con respecto al proceso operativo de recolección.

Para garantizar la calidad de la información es necesario tener en cuenta las siguientes normas:

- El personal de campo y oficina, a saber: recolectores y supervisores, ha de ser evaluado periódicamente.
- Por lo menos dos veces a la semana, el líder de la GEIH en cada ciudad debe hacer una reunión con el personal de campo para solucionar casos especiales, analizar el desarrollo del proyecto en cuanto a cobertura, no respuesta, cartografía, manejo de la muestra, etc.
- Los recolectores deben ser rotados de supervisor y de zonas de la ciudad.
- Mensualmente, el líder de la GEIH de cada ciudad debe enviar al DANE Central un informe de la justificación de la variación de las tasas de mercado laboral según formato establecido.
- Las cargas de trabajo deben ser evacuadas en su totalidad en la semana respectiva, y por ningún motivo dejar trabajo de recolección para la semana siguiente.
- El personal de campo debe tomar obligatoriamente un día de descanso a la semana, según convenga, para el desarrollo óptimo del operativo.
- En cada ciudad se debe conformar un "Comité Técnico de la Gran Encuesta Integrada de Hogares". Este comité tendrá como funciones principales hacer seguimiento, evaluar los aspectos metodológicos y operativos de la encuesta y producir las recomendaciones que considere pertinentes, las cuales deben ser enviadas al DANE Central junto con el informe mensual.

ANÁLISIS DE CONTEXTO

Este procedimiento está orientado al análisis tanto de la consistencia interna de los datos como del comportamiento del mercado laboral y su relación con otras variables macroeconómicas que sirvan de base para la formulación, el seguimiento y la evaluación de políticas.

Para realizar el análisis de contexto se llevan a cabo diferentes tipos de análisis como son:

- El análisis descriptivo ayuda a observar el comportamiento de la muestra en estudio, a través de tablas, gráficos, intervalos de confianza y estadísticas de tendencia y dispersión.
- Se analiza en la muestra la estructura de los indicadores por dominios de estudio a partir de la distribución de frecuencias, y se detectan posibles inconsistencias y valores atípicos. Se verifica la cobertura a nivel de segmento, viviendas, hogares y persona, y se observa la distribución de la pérdida de muestra para realizar los respectivos ajustes de la no respuesta.
- El análisis inferencial tiene como objetivo realizar la estimación de la muestra a la población objetivo. En este proceso se aplican los factores de expansión y se revisa la inferencia a la población objetivo establecida para los correspondientes dominios de estudio.
- 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.
- El secretario técnico de la GEIH, el coordinador de Temática Social, el director de Metodología y Producción Estadística, el subdirector y el director del DANE revisan la consistencia de la información sobre los principales indicadores de mercado laboral IML y posteriormente se elaboran y organizan otros productos como son el boletín de prensa y la presentación de resultados. En este proceso se realiza análisis comparativos y se analiza la consistencia interna de la información frente a otras variables económicas.
- Temática Social analiza el dato por divulgar frente al mismo periodo del año inmediatamente anterior y respecto a otros periodos, y los de las demás variables según conceptos de fuerza de trabajo.

COEFICIENTE DE VARIACIÓN

Para medir la magnitud de la variabilidad de la distribución muestral del estimador, denominado error muestral, se usan el error estándar y el coeficiente de variación.

El coeficiente de variación se define como la relación porcentual del error estándar o raíz cuadrada de la varianza del estimador y el estimador, multiplicado por 100.

El valor de este coeficiente, expresado en porcentaje, permite evaluar la calidad de un procedimiento de estimación.