

# Colombia - Encuesta Pulso de la Migración - EPM - VII- 2024

**Departamento Administrativo Nacional de Estadística**

report\_generated\_on: October 10, 2024

visit\_data\_catalog\_at: <https://microdatos.dane.gov.co/index.php>

# Información general

## Identificación

NÚMERO DE ID  
DANE-DIMPE-EPM-VII-2024

## Versión

DESCRIPCIÓN DE VERSIÓN  
Ronda V-2023

## Información general

### RESUMEN

El Departamento Administrativo Nacional de Estadística (DANE) en desarrollo de su objetivo misional de producir estadísticas oficiales, que cumplan con los estándares internacionales y que sirvan de base para la toma de decisiones, se propone elaborar la operación estadística Pulso de la Migración. Esta encuesta está dirigida a personas migrantes desde Venezuela y busca producir información relacionada sobre el origen y razones de la migración, acceso a empleo, acceso a servicios de salud, integración socioeconómica, cambios en el flujo de ingresos y remesas, así como percepciones que tienen los migrantes sobre su experiencia de migración, entre otros temas. En ese sentido, esta operación busca ser una herramienta fundamental para la definición e implementación de políticas públicas.

### TIPO DE DATO

Encuesta por muestreo (ssd)

### UNIDAD DE ANÁLISIS

a) Unidad de observación: personas.

b) Unidad de análisis: En este caso están representadas por los hogares y las personas.

c) Unidad de muestreo: Es un muestreo en fases, la primera fase corresponde a la muestra de GEIH, cuyas unidades muestrales son los segmentos. La segunda fase corresponde a todos los hogares y personas de la muestra, que se identifican de acuerdo con las características mencionadas en la población objetivo (al menos un migrante mayor de 15 años en cada hogar).

## Ámbito

### NOTAS

La Encuesta Pulso de la Migración VII ronda indaga sobre diversas temáticas, presentando las principales variables por ronda, a continuación:

#### Ronda 7

- Variables de identificación como: nacionalidad, documento de identificación, documento de regularización, fecha de expedición del documento de regularización, número de contacto, sexo de nacimiento, edad, hacinamiento, personas con las que vive en Colombia, es representante o no del hogar (persona idónea para responder preguntas del hogar).
- Características del hogar: número de personas colombianas y venezolanas en el hogar, convivencia con niños o adolescentes, acceso a educación, deserción, motivos y experiencias de discriminación escolar, consumo de comidas, entre otras.
- Situación migratoria: ubicación geográfica antes y después de migrar, fecha de ingreso a Colombia, forma de migración, intención de permanencia o continuar recorridos migratorios, razones de migración, contactos antes de migrar, reunificación

familiar.

- Salud: el estado de salud, afiliación al sistema de salud y régimen en Colombia, afiliación al sistema de pensiones, dificultades en el acceso al servicio de salud, entre otras.

- Mercado laboral: actividad que ocupó la mayor parte del tiempo antes y después de migrar, ocupación en Venezuela, dificultades y medios utilizados para encontrar trabajo, tipo de contrato, horas semanales de trabajo, acceso a servicios financieros, envío de remesas a Venezuela.

#### KEYWORDS

Actividad económica, Discriminación o actos de discriminación, Frontera Internacional, Imaginario Social, Migración, Migrante, Ocupación, Pasaporte, Remesas, Migración de retorno, Xenofobia, Salud

## Cobertura

### COBERTURA GEOGRÁFICA

La Encuesta Pulso de la Migración VII cubre geográficamente a todo el territorio nacional excluyendo nuevos departamentos<sup>1</sup> y la Isla de San Andrés. Los resultados se presentan para el total de las 23 ciudades y áreas metropolitanas.

1 Nuevos departamentos: Arauca, Casanare, Putumayo, Amazonas, Guaviare, Guainía, Vaupés y Vichada.

### GEOGRAPHIC UNIT

Para una mejor comprensión del tema, es necesario definir los conceptos de cabeceras, centros poblados y rural disperso.

**Cabecera municipal (CM):** 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 (DANE-Conceptos estandarizados).

**Centro poblado:** es un concepto creado por el DANE para fines estadísticos, útil para la identificación de núcleos de población. Se define como una concentración de mínimo veinte (20) viviendas contiguas, vecinas o adosadas entre sí, ubicada en el área rural de un municipio o de un corregimiento departamental. Dicha concentración presenta características urbanas tales como la delimitación de vías vehiculares y peatonales (DANE-Conceptos estandarizados). Este concepto considera:

- **Caserío:** sitio que presenta un conglomerado de viviendas, ubicado comúnmente 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.
- **Inspección de policía:** es una instancia judicial en un área que puede o no ser amanzanada y que ejerce jurisdicción sobre un determinado territorio municipal, urbano o rural y que depende del departamento o del municipio. Es utilizada en la mayoría de los casos con fines electorales. Su máxima autoridad es un inspector de policía.
- **Corregimiento municipal:** es una división del área rural del municipio, la cual incluye un núcleo de población, considerada en los Planes de Ordenamiento Territorial, POT. El artículo 117 de la ley 136 de 1.994 faculta al concejo municipal para que mediante acuerdos establezca esta división, con el propósito de mejorar la prestación de los servicios y asegurar la participación de la ciudadanía en los asuntos públicos de carácter local (DANE-conceptos estandarizados).

**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 (DANE-conceptos estandarizados).

**Área resto municipal:** Delimitación geográfica definida por el DANE para fines estadísticos, alusiva al área comprendida entre el perímetro censal y el límite municipal definido por ordenanza de la Asamblea Departamental. Se caracteriza por la disposición dispersa de viviendas, de centros poblados y de explotaciones agropecuarias existentes en ella (DANE-conceptos estandarizados).

Para la encuesta se han establecido cuatro rondas, en las cuales hay diferentes niveles de atrición. Dependiendo de los niveles de atrición se define la desagregación geográfica adecuada, por ejemplo, para la primera ronda se hace la siguiente desagregación geográfica:

Para la encuesta se estructura la siguiente desagregación geográfica:

- Total 23 ciudades y áreas metropolitanas

El uso de la desagregación geográfica dependerá del tamaño de muestra y del coeficiente de variación, no se aconseja presentar resultados con tamaños de muestras pequeños o con coeficientes de variación superiores al 15%. En este sentido cuando se desee hacer desagregación por regiones lo más apropiado es unir rondas.

UNIVERSO

El universo de estudio en la séptima ronda de la Encuesta Pulso de la Migración está conformado por la población civil no institucional que hacen parte de los hogares en Colombia, excluyendo la isla de San Andrés y nuevos departamentos<sup>1</sup> para el año 2022, con al menos una persona migrante desde Venezuela mayor de 15 años.

1 Nuevos departamentos: Arauca, Casanare, Putumayo, Amazonas, Guaviare, Guainía, Vaupés y Vichada.

Productores y Patrocinadores

INVESTIGADOR(ES) PRIMARIO(S)

Nombre	Affiliation
Departamento Administrativo Nacional de Estadística	Dirección de Metodología y Producción Estadística - DIMPE

OTROS PRODUCTORES

Nombre	Affiliation	Role
Banco Interamericano de Desarrollo		Financiador

FINANCIAMIENTO

Nombre	Abbreviation	Role
Banco Interamericano de Desarrollo	BID	Financiador

OTROS RECONOCIMIENTOS

Nombre	Affiliation	Role
Universidad del Rosario		

Producción de metadato

METADATO PRODUCIDO POR

Nombre	Abbreviation	Affiliation	Role
Ivan Rolando Castillo Prieto	ircastillo@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización - DIRPEN	Coordinador Regulación Estadística
Cesar Mauricio Lopez Alfonso	cmlopeza@dane.gov.co	Dirección de Metodología y Producción Estadística - DIMPE	Director Técnica DIMPE
Alejandro Ramos Hernández	aramosh@dane.gov.co	Dirección de Metodología y Producción Estadística - DIMPE	Coordinador GIT Capital Social
Karen Liseth Atis Ortega	klatiso@dane.gov.co	Dirección de Metodología y Producción Estadística - DIMPE	Contratista
Juliette Carolina Villamil Monroy	jcwillamil@dane.gov.co	Dirección de recolección y acopio - DRA	Coordinadora GIT Personas e Instituciones
Juan Pablo Saenz Perilla	jpsaenzp@dane.gov.co	Dirección de Metodología y Producción Estadística - DIMPE	Temático DIMPE

Juan Jose Rubio Mesa	jjrubiom@dane.gov.co	Dirección de Metodología y Producción Estadística - DIMPE	Diseños Muestrales
Carlos Augusto Muñoz Alfonso	rhzorroc@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización - DIRPEN	Verificador DIRPEN
Katherine Andrea Rojas Acero	karojasa@dane.gov.co	Dirección de Difusión y Cultura Estadística - DICE	Verificador DICE

FECHA DE PRODUCCIÓN DEL METADATO

2024-10-10

ID DEL DOCUMENTO IDD

COL-DANE-EPM-VII-2024

# Muestreo

## Procedimiento de muestreo

La investigación es una encuesta orientada a hogares y personas, por lo tanto, se establece una estrategia muestral adecuada para acceder a mediciones de estas unidades de análisis con procedimientos que aseguran estimadores 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 opta por una muestra en dos fases. En la primera fase se establece un diseño muestral probabilístico (Encuesta GEIH marco 2018), donde se obtiene información de un grupo de personas.

En la segunda fase y tomando en cuenta la información recolectada en la primera fase, se seleccionan todos los hogares con al menos un migrante mayor de 15 años.

Durante las cuatro primeras rondas de esta operación estadística, se estableció una estrategia de muestreo tipo panel, donde la muestra definida en la primera ronda fue usada en las siguientes tres rondas, metodología mediante la cual, dado el carácter específico de migrantes que tiene la población objeto de estudio, se buscó contar con información periódica de las fuentes para conocer las variaciones en los individuos que participan en la encuesta.

Debido al agotamiento de la muestra fruto de las rondas anteriores y considerando la atrición generada con el diseño tipo panel; a partir de quinta ronda, se utiliza un diseño de tipo corte transversal. Para esta última ronda la muestra fue seleccionada de la GEIH entre los meses de diciembre de 2022 a julio de 2023.

### Cálculo de tamaño de muestra:

El tamaño de muestra se definió con el objetivo de garantizar prevalencias de un 10%, con un margen de error de 1 punto porcentual y 95% de confiabilidad, de acuerdo con lo anterior se optó por hacer seguimiento a una muestra de hogares con al menos un migrante venezolano, lo cual equivale aproximadamente a cuatro mil hogares por ronda.

### Selección de muestra:

La selección de la muestra se hace a partir de todos los hogares e individuos identificados en la GEIH que cumplen con la característica definida en la población objetivo, es decir aquellos hogares con al menos un migrante mayor de 15 años. Para obtener el tamaño de muestra definido, se tomó información de la GEIH entre los meses de diciembre de 2022 a julio de 2023.

### Mantenimiento de muestra y manejo de novedades muestrales:

Como se mencionó al inicio de esta sección, la Encuesta Pulso de la Migración fue una encuesta tipo panel en las cuatro primeras rondas, donde la muestra de la primera ronda fue usada en las siguientes rondas. Sin embargo, entre la realización de una ronda y otra desaparecen unidades muestrales, fenómeno conocido como atrición. Para superar los inconvenientes de atrición, en cada ronda siguiente, se refrescó la muestra adicionando hogares con las mismas estructuras poblacionales de la ronda original.

En la ronda 1 la muestra se seleccionó con los encuestados de la GEIH marco 2005 de enero a junio de 2021. En la ronda 2 con el propósito de mantener el tamaño de muestra (4000 hogares), se refrescó la información de los individuos que se pierden de la muestra de la ronda anterior, con los encuestados de junio, julio y agosto de 2021. En la ronda 3 se refrescó la muestra de la ronda 2 con los encuestados de agosto, septiembre y octubre de 2021 y en la última ronda se refrescó la muestra con los encuestados de noviembre y diciembre de 2021.

La siguiente tabla refleja el porcentaje de atrición respecto a la muestra de la ronda 1 que tuvo cada una de las primeras cuatro rondas

Tabla 1: Porcentaje de atrición respecto a la muestra de la ronda 1

Ronda	Encuestas completas	No. Encuestas de Ronda	1	Porcentaje de atrición
1	6966	6966		
2	7633	5677	18.5%	
3	5927	4602	33.9%	
4	8394	3662	47.4%	

Fuente: DANE.

Con el objetivo de mantener el tamaño muestral, a partir de la quinta ronda de la Encuesta Pulso de la Migración los datos mantendrán una estructura de corte transversal aplicando medidas adecuadas para maximizar la tasa de respuesta y de recontacto con las fuentes seleccionadas. Se espera que se realicen múltiples intentos de contacto, con los individuos antes de considerarlos una observación "perdida". Por lo tanto, los individuos deben ser contactados en diferentes momentos del día y en diferentes días de la semana. Una observación se considerará "perdida" sólo si se ha intentado ponerse en contacto con ese hogar varias veces, durante varios días y en diferentes momentos del día, incluidos los fines de semana.

## Desviaciones del diseño de la muestra

### MANTENIMIENTO DE MUESTRA Y MANEJO DE NOVEDADES MUESTRALES

Con el objetivo de mantener el tamaño muestral, a partir de la quinta ronda de la Encuesta Pulso de la Migración los datos mantendrán una estructura de corte transversal aplicando medidas adecuadas para maximizar la tasa de respuesta y de recontacto con las fuentes seleccionadas.

Se espera que se realicen múltiples intentos de contacto, con los individuos antes de considerarlos una observación "perdida". Por lo tanto, los individuos deben ser contactados en diferentes momentos del día y en diferentes días de la semana.

Una observación se considerará "perdida" sólo si se ha intentado ponerse en contacto con ese hogar varias veces, durante varios días y en diferentes momentos del día, incluidos los fines de semana.

Personas esperadas: 7.747

Personas con encuesta completa: 8.020

Cobertura personas con encuesta completa: 103,5%

Hogares esperados: 4.000

Hogares con encuesta completa: 3.695

Cobertura hogares con encuesta completa: 92,4%

## Tasa de respuesta

Cobertura personas: 103,5%

Cobertura de hogares: 92,4%

## Ponderación

Para ampliar información de esta sección, sugerimos a los usuarios consultar el documento metodológico en la sección 2.2.11 Especificaciones de ponderadores

Para construir los factores de expansión se definen primero las poblaciones a donde se deben expandir los individuos. Se toma como referencia la estimación de la población correspondiente a los meses de diciembre de la Gran Encuesta Integrada de Hogares GEIH 2022 y enero a julio de la GEIH 2023. Los factores de expansión usados en la encuesta son heredados de la (GEIH) y estos son posteriormente ajustados por la pérdida de cobertura que hay en el desarrollo de la

operación. Los ajustes de cobertura se describen a continuación.

#### Ajuste de cobertura

Para cada individuo de la muestra, se toman el mismo factor de expansión asignado desde la GEIH y se procede a hacer un ajuste a nivel de los estratos de la GEIH, género, tipo de mirante y mes de selección, que se genera por la pérdida de cobertura que hay en el desarrollo de la operación,

#### ESTIMADORES

Para la encuesta se presentan los estimadores de totales y de razones con sus respectivos estimadores de varianza. Los estimadores de varianza se hacen a través de la metodología de Bootstrap que consiste en remuestreo de la muestra original por medio de permutaciones, que se utilizan posteriormente para estimar la distribución muestral de un estadístico y evaluar la precisión del mismo (Rao, Wu, and Yue, 1992).

#### ESTIMADOR DE TOTALES

El estimador del total para la variable con la muestra en el pos-estrato tiene la siguiente forma:

Con el sub-índice que representa al individuo.

El estimador del total poblacional de la variable es entonces:

En los casos en los cuales se requiera estimar grupos especiales de la población (dominios de estudio), la estrategia es crear otra variable que marque con 1 los individuos que pertenecen al grupo y cero los que no. Posteriormente, esta variable se multiplica a la variable y creando una nueva variable ( $y_1$ ). En este punto se procede de la misma manera para estimar.

#### Estimador de varianza del estimador del total

El procedimiento para el cálculo del estimador de varianza del estimador total, se hace con el método de Bootstrap (Rao, Wu, and Yue, 1992) a través de la función PROC SURVEYMEANS y la opción VARMETHOD=BOOSTRAP, disponible en el software estadístico SAS. La función permite estimar el total de la población de la variable junto con la varianza del estimador en diseños muestrales estratificados.

#### ESTIMADOR DE RAZONES

El estimador de razón se construye haciendo el cociente de los estimadores de dos totales. Es así que si se quiere calcular la razón de los totales y se tiene la siguiente formula:

#### Estimador de varianza del estimador de razones

Para el cálculo de estimación de la varianza del estimador de razones se usa nuevamente el método de Bootstrap con ayuda de la función PROC SURVEYMEANS y la opción VARMETHOD=BOOSTRAP.

#### Errores muestrales

Son varias las formas de medir la precisión de los indicadores. En la investigación se presentan los coeficientes de variación estimados (errores relativos) y los errores marginales de cada indicador cada vez que se publica. Los estimadores de total y razón tienen 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, la función PROC SURVEYMEANS junto con la opción VARMETHOD=BOOSTRAP permite también obtener la estimación del coeficiente de variación de los estimadores.



# Formularios

## Información general

---

El tipo de formulario utilizado en la Encuesta Pulso de la Migración está conformado por un listado de diferentes preguntas agrupadas en diferentes módulos, varían entre rondas de recolección y publicación de resultados. La simplicidad del cuestionario busca recolectar datos de forma rápida y oportuna, siendo un cuestionario diseñado para entrevistas telefónicas que deben durar de 25 a 30 minutos. A continuación, se presenta la estructura del formulario para la quinta ronda de recolección y los objetivos de cada módulo:

**Módulo de identificación:** contiene preguntas cuyo objetivo es el de captar los datos de identificación, sexo y años de la persona encuestada, nacionalidad, acceso a documentos de regularización, convivencia en Colombia, hacinamiento, persona idónea para responder preguntas del hogar (representante)

**Modulo características del hogar:** Busca obtener información sobre número de personas colombianas y venezolanas en el hogar, convivencia con niños o adolescentes, acceso a educación, deserción y discriminación escolar, entre otras.

**Módulo de situación migratoria:** contiene preguntas con las que se busca obtener información acerca del lugar de residencia previo a la situación migrante, intención de permanencia, intención de seguir recorridos migratorios, contactos antes de migrar y reunificación familiar.

**Modulo salud:** indaga sobre el estado de salud, afiliación al sistema de salud en Colombia, afiliación al sistema de pensiones, dificultades en el acceso al servicio de salud.

**Módulo de mercado laboral:** contiene preguntas con las que se busca obtener información sobre la situación de trabajo y actividades realizadas en el país de origen y su situación actual, tipo de contrato, así como las dificultades de acceso al empleo, acceso a servicios financieros y envío de remesas a Venezuela.

**Modulo percepción y discriminación:** busca conocer el motivo y lugares de discriminación, identidad cultural de los migrantes, percepción de pobreza, comparación con el nivel de vida de los padres, y percepción del nivel de vida de los hijos.

Es relevante señalar que en el diseño del formulario de la Encuesta Pulso de la Migración se emplearon criterios para evaluar la factibilidad como su relación con los objetivos de la encuesta, su pertinencia, la claridad para el encuestado, la disponibilidad de fuentes alternativas de información, entre otros. Así mismo, se efectuaron pruebas en el aplicativo de tal forma que no se afectaran las reglas de validación y se verificó que el fraseo de las preguntas fuera adecuado para la comprensión por parte del encuestado/a, de manera que se garantice que la información recolectada tenga el alcance de lo que se pretende medir.

# Recolección de datos

## Fecha de recolección de datos

Start	End	Cycle
2024-04-08	2024-06-08	Ronda 7

## Time Periods

Start	End	Cycle
2024-04-01		Ronda 7

## Modo de recolección de datos

Entrevista telefónica

## Notas de recolección de datos

La Encuesta Pulso de la Migración se realiza mediante entrevista directa de forma telefónica, capturando los datos a través del aplicativo web MySurvey. Este es una plataforma web que contiene un formulario electrónico con el cuestionario, las tablas de validaciones y el control de flujos, lo que permite obtener indicadores de calidad.

El directorio o listado de hogares a encuestar corresponde a una submuestra de la GEIH, de los meses inmediatamente anteriores al periodo de recolección. La selección de la muestra corresponde a las personas migrantes (que hayan migrado en los últimos 5 años o colombianos retornados) según lo definido en el directorio o listado de hogares establecido por el área de diseños muestrales. Este listado de personas es asignado a cada encuestador dependiendo de la sede en la que ha sido contratado.

## Formularios

El tipo de formulario utilizado en la Encuesta Pulso de la Migración está conformado por un listado de diferentes preguntas agrupadas en diferentes módulos, varían entre rondas de recolección y publicación de resultados. La simplicidad del cuestionario busca recolectar datos de forma rápida y oportuna, siendo un cuestionario diseñado para entrevistas telefónicas que deben durar de 25 a 30 minutos. A continuación, se presenta la estructura del formulario para la quinta ronda de recolección y los objetivos de cada módulo:

Módulo de identificación: contiene preguntas cuyo objetivo es el de captar los datos de identificación, sexo y años de la persona encuestada, nacionalidad, acceso a documentos de regularización, convivencia en Colombia, hacinamiento, persona idónea para responder preguntas del hogar (representante)

Modulo características del hogar: Busca obtener información sobre número de personas colombianas y venezolanas en el hogar, convivencia con niños o adolescentes, acceso a educación, deserción y discriminación escolar, entre otras.

Módulo de situación migratoria: contiene preguntas con las que se busca obtener información acerca del lugar de residencia previo a la situación migrante, intención de permanencia, intención de seguir recorridos migratorios, contactos antes de migrar y reunificación familiar.

Modulo salud: indaga sobre el estado de salud, afiliación al sistema de salud en Colombia, afiliación al sistema de pensiones, dificultades en el acceso al servicio de salud.

Módulo de mercado laboral: contiene preguntas con las que se busca obtener información sobre la situación de trabajo y actividades realizadas en el país de origen y su situación actual, tipo de contrato, así como las dificultades de acceso al empleo, acceso a servicios financieros y envío de remesas a Venezuela.

Modulo percepción y discriminación: busca conocer el motivo y lugares de discriminación, identidad cultural de los migrantes, percepción de pobreza, comparación con el nivel de vida de los padres, y percepción del nivel de vida de los

hijos.

Es relevante señalar que en el diseño del formulario de la Encuesta Pulso de la Migración se emplearon criterios para evaluar la factibilidad como su relación con los objetivos de la encuesta, su pertinencia, la claridad para el encuestado, la disponibilidad de fuentes alternativas de información, entre otros. Así mismo, se efectuaron pruebas en el aplicativo de tal forma que no se afectaran las reglas de validación y se verificó que el fraseo de las preguntas fuera adecuado para la comprensión por parte del encuestado/a, de manera que se garantice que la información recolectada tenga el alcance de lo que se pretende medir.

## Data Collectors

Nombre	Abbreviation	Affiliation
Departamento Administrativo Nacional de Estadística	DANE	Gobierno Nacional

## Supervisión

### PROCESO DE APRENDIZAJE A LOS ASISTENTES DE ENCUESTA DE CADA CIUDAD Y AL PERSONAL SELECCIONADO.

Para la Encuesta de Pulso de la Migración (EPM), el proceso de aprendizaje se lleva a cabo a través de videoconferencia con los asistentes técnicos y el personal operativo seleccionado de cada sede. Este proceso es impartido por cada una de las áreas técnicas del DANE Central con las siguientes temáticas:

- Dirección de Metodología y Producción Estadística - DIMPE: temática y diligenciamiento de la encuesta
- Oficina Asesora de Sistemas: aplicativo y transmisión de la información
- Dirección de Difusión y Cultura Estadística - DICE: relacionamiento y sensibilización
- Dirección de Recolección y Acopio -DRA: aspectos operativos
- Gestión Humana: seguridad y salud en el trabajo y protocolo de bioseguridad.
- Diseños muestrales: muestra.

Dado lo anterior, la asistencia de encuesta, una vez tomada esta fase de aprendizaje, tiene la responsabilidad, de replicar en su ciudad este momento de aprendizaje de manera presencial o virtual al personal de campo preseleccionado y seleccionado las veces que se requiera.

### PROCESO DE SENSIBILIZACIÓN Y ACUERDOS DE INTERCAMBIO

El objetivo del proceso de sensibilización es lograr una respuesta oportuna y de calidad por parte de las fuentes de información. Se busca lograr un ambiente de colaboración, acercamiento y conocimiento de la población objetivo mediante la llamada telefónica a cada uno de los hogares y/o personas a los que les será aplicada la encuesta.

Durante la llamada telefónica se pueden establecer citas y horarios en los cuales el grupo de recolección del DANE puede aplicar la encuesta a la fuente de información, en caso de que la sensibilización no encuentre al miembro del hogar que debe responder la encuesta.

Para la Encuesta de Pulso de la Migración el proceso de sensibilización es llevado a cabo por grupo de recolección al mismo tiempo de toma de información a la fuente. Este proceso puede estar acompañado de piezas de comunicación y sensibilización diseñadas por el DANE para facilitar el contacto y la entrega de información de las fuentes.

### ELABORACIÓN DE MANUALES2.

Los manuales constituyen una herramienta eficaz para el entrenamiento del personal operativo, contribuyendo de manera sustantiva a que la transmisión del conocimiento sea homogénea y las personas contratadas desempeñen adecuadamente las actividades y responsabilidades asignadas. Dentro de los instrumentos diseñados por el área logística y temática para el desarrollo de la encuesta, se encuentran los siguientes:

- Manual de conceptos básicos y recolección: en este documento se detalla cada uno de los componentes del operativo de recolección. Así mismo, se define la forma como debe ser diligenciado el cuestionario en cada uno de sus capítulos/módulos. Para cada capítulo/módulo se especifican los objetivos y la conceptualización de cada pregunta.
- Manual operativo: en este se consignan los diferentes aspectos que se deben tener en cuenta para el desarrollo del operativo de campo en materia de: procesos pre-operativos y operativos, recursos, instrumentos de recolección, entre otros.
- Manual de diligenciamiento de formatos: en este manual se presenta de una manera clara y precisa el diligenciamiento de los formatos operativos tipo registro que deben ser utilizados por el personal de campo.

Estos manuales, así como los formatos utilizados en los diferentes procesos operativos de la encuesta se encuentran disponibles para su consulta en la plataforma ISOLUCION, a la cual se puede acceder a través de la intranet institucional u otro enlace oficial.

Para el control del operativo, se diseñaron una serie de reportes que permitirán llevar a cabo la revisión de la información transmitida. El propósito de estos controles es efectuar un seguimiento al trabajo realizado en las distintas ciudades para tomar oportunamente los correctivos necesarios ante los eventuales inconvenientes que se puedan presentar, así como el de asegurar la cobertura, calidad y oportunidad de la información.

Para contar con información que permita garantizar la cobertura y control de calidad de la encuesta, desde la asistencia técnica se debe enviar al DANE Central, la documentación exigida según cronograma.

Desde la asistencia técnica y la supervisión se debe controlar la cobertura diaria de los grupos de trabajo a fin de detectar cuando la cobertura es baja e, igualmente, aquellos que terminan su trabajo muy rápido, esto con el propósito de revisar su forma de trabajo y las cargas que cada uno maneja y, de esta manera, ajustarlas.

Es necesario prestar especial atención cuando reportan muchos casos de no respuesta (rechazos, entrevistas incompletas, ocupados, ausentes temporales, etc.); posiblemente, se trate de personas con técnicas de entrevista o de presentación poco adecuadas. De persistir esta situación, se debe recurrir oportunamente al asistente técnico para evaluar el tema y tomar acciones correctivas.

Algunos mecanismos de control y seguimiento de la información son los siguientes:

#### · Calidad de la recolección

En la Sede/Subsede: se obtiene esta información de la revisión y acompañamiento en las llamadas que realizan las personas con actividades de recolección. Esta información es el insumo para que la asistencia de encuesta/supervisión diligencie el indicador de calidad de recolección, que, posteriormente, es enviado a DANE Central.

En DANE Central: se consolida el indicador de calidad de la recolección reportado por las ciudades, llevando un registro semanal que permita evaluar cuál es el resultado conjunto de la revisión de formularios y acompañamiento en cuanto a las llamadas realizadas por ciudad. En caso de presentarse algún error o novedad en el formato, se reporta a la ciudad para su corrección y/o justificación.

En la Sede/Subsede: en caso de presentarse algún error, la sede/subsede debe corregir y enviar nuevamente a DANE Central.

#### · Inconsistencias y/o alarmas

En DANE Central: se obtiene de la realización de diversas pruebas al resumen de cobertura reportado por las ciudades, con el objetivo de verificar que los datos sean consistentes. En el caso de presentarse inconsistencias y/o alarmas serán reportadas a la ciudad que corresponda. Para la Encuesta Pulso de la Migración (EPM), se generan alarmas a las sedes /subsedes cuando se presentan las siguientes situaciones en campo:

- Bajo número de personas con encuesta completa (EC).
- Alto número de rechazos por parte de las personas a encuestar.
- Alto número de encuestas no efectivas: Ausente Temporal (AT), Sin tiempo para responder, No lo conocen, Ya no hace parte del hogar y Falleció, No migrante, etc.

En la sede: la ciudad debe velar porque las inconsistencias y/o alarmas presentadas sean corregidas a tiempo y dentro de la etapa de recolección.

#### · Comparativo campo vs sistemas

En el DANE Central: se obtiene de comparar la información reportada en el resumen de cobertura (campo) frente a la información de base de sistemas obtenida del aplicativo web. En este comparativo se obtienen las diferencias que existen a nivel de personas con encuesta completa, las cuales son reportadas a cada ciudad para su verificación y corrección según sea el caso.

En la sede: recibe la información y revisa la diferencia para determinar si se debe a información faltante en el resumen de cobertura o si faltó dar por terminada la encuesta en el aplicativo web y así poder realizar los correctivos que sean necesarios.

#### · Indicador de oportunidad

El DANE Central: busca garantizar que la información enviada por las sedes y subsedes sea oportuna. En este sentido, la puntualidad se mide con respecto al envío del resumen de cobertura e indicador de calidad dentro del plazo de tiempo establecido.

En la Sede/Subsede: debe realizar el envío de la información en los tiempos establecidos por el área de Logística.

#### · Indicador de calidad de la recolección

Este indicador corresponde a la evaluación del desempeño del equipo de trabajo que resulta del acompañamiento a las encuestas telefónicas y el seguimiento y la revisión de los formularios o cuestionarios digitalizados en el aplicativo de captura.

En la Sede/: se obtiene esta información de la revisión y acompañamiento telefónico que la Asistencia técnica realiza. Esta información es el insumo para diligenciar el indicador de calidad de recolección.

DANE central: consolida el indicador de calidad de la recolección reportado por las ciudades, llevando un registro quincenal que permita evaluar cuál es el resultado conjunto de la revisión de formularios y acompañamiento telefónico por ciudad. En caso de presentarse algún error o novedad en el formato, se reporta a la ciudad para su corrección y justificación.

En la Sede/Subsede: en caso de presentarse algún error, la sede debe corregir y enviar nuevamente a DANE central.

#### · Indicador de cobertura

Este indicador reporta el avance en la cobertura de los hogares y personas migrantes con encuesta completa lograda en la sede y a nivel nacional.

En la Sede/Subsede: se consolida el resumen de cobertura con base en el formato de resumen acumulado de personas y hogares con el número de la encuesta.

DANE central: cruza los datos del total personas y hogares con encuesta completa contra el total de personas y hogares encontrados reportados en el resumen de cobertura. En caso de encontrar diferencias, se reporta a la sede mediante un comparativo a la sede que haya presentado la inconsistencia.

En la Sede/Subsede: si se presentan diferencias, la sede deberá verificar y justificar por qué se presentaron y corregir si es el caso dentro de la semana de recolección.

Para el control operativo, el equipo logístico, diseñó un conjunto de formatos que facilitan el seguimiento tanto de la etapa de recuento, como de la fase de sensibilización, supervisión y recolección de la información

# Procesamiento de datos

## Edición de datos

---

### CREACIÓN DE LA BASE DE DATOS (ORACLE)

Es necesario contar con el motor de bases de datos y en un esquema creado para la investigación que permita el cargue y almacenamiento general de los datos. La base de datos para la investigación asignada es una base de datos relacional donde la Encuesta Pulso de la Migración, tiene un único registro por persona.

Se debe asignar el número de la encuesta asociado a la investigación y definir los tipos de datos, textos de las preguntas, listas de valores asociadas, subpreguntas, entre otros. De igual manera, definir los formularios.

Adicionalmente, se establecen los flujos y validaciones de cada una de las preguntas. En la estructura del formulario se debe asociar a cada pregunta el código equivalente con el código de pregunta en Oracle. Así mismo, como parte del diseño se crea un ETL (Extract, Transform and Load) en Pentaho Data Integration y se carga la información a la tabla de la base de datos.

### CONSOLIDACIÓN DE ARCHIVOS DE DATOS

#### CARGUE DE INFORMACIÓN

El proceso de carga de información a la base de datos se realiza seleccionando los archivos planos generados con la herramienta propia de la plataforma MySurvey y con un ETL (Extract, Transform and Load) en Pentaho Data Integration creado, se cargan los datos.

Al iniciar el proceso de carga se verifica la identificación única para cada encuesta, garantizando que no exista duplicidad en la información.

#### DICCIONARIO DE DATOS

En el diccionario de datos se describen, los metadatos de todas las variables incluidas en las tablas generadas con la información de la operación estadística, entre ellos el nombre de la variable, su código, tipo, extensión, el texto de la pregunta respectiva y de las alternativas de respuesta.

El diccionario se construye de acuerdo con la codificación de las preguntas con las que cuenta el formulario de la encuesta. En la medida en que el formulario se revisa y ajusta periódicamente, esta se actualiza.

#### REVISIÓN Y VALIDACIÓN

Documenta el procedimiento establecido para revisar y validar el archivo de datos que se conforma a partir de la recolección o el acopio, teniendo en cuenta la posibilidad de volver a consultar al encuestado idóneo o la fuente de datos.

De manera adicional, en conjunto con los equipos de logística y sistemas, se realiza una revisión semanal de la información recolectada para detectar inconsistencias susceptibles de ser verificadas y corregidas en campo con el fin de garantizar la calidad de la información recolectada.

### DISEÑO DE INSTRUMENTOS DE EDICIÓN ( VALIDACIÓN Y CONSISTENCIA) E IMPUTACIÓN DE DATOS

#### CONSISTENCIA DE INFORMACIÓN

A través del desarrollo de sentencias Oracle, se identifican posibles inconsistencias por duplicidad a nivel de personas. Se generan reportes que permiten realizar constantemente el monitoreo y control a la información que se ha recolectado.

Posteriormente, en el esquema diseñado en ORACLE se crean procedimientos para validar la información de la base de

datos, flujos, variables sin información, rangos y estructura, con el fin de obtener una base de datos consistente.

Es así que cada operación queda almacenada en su propio esquema de base de datos. Esto permite que los datos se mantengan aislados de la información de otras operaciones 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.

#### DISEÑO PARA LA GENERACIÓN DE CUADROS DE RESULTADOS

El equipo temático de la encuesta recibe el archivo en formato DTA por parte del equipo de sistemas con las especificaciones para la generación de los cuadros de salida que se van a publicar. Posteriormente se reciben los factores de expansión por parte del equipo de muestras, con los cuales el equipo de temática genera los cuadros de salida definitivos, que permiten la visualización de los resultados de la encuesta. Estos cuadros se generan en formato XLS y son aquellos que se encuentran en el Anexo publicado.

Revisados los cuadros generados por parte del equipo temático, se realizan ajustes, en caso de ser necesario. Los cuadros de salida se generan para los dominios total nacional. Para el análisis y difusión de los resultados de la EPM se han definido los cuadros de salida desagregados por sexo, edad y tamaño del hogar.

# Estimación de datos

## Estimación de error de la muestra

El proceso de análisis de resultados está comprendido por:

### MÉTODOS DE ANÁLISIS DE RESULTADOS

Los métodos, las técnicas y los procedimientos que permiten verificar la coherencia y calidad de la Encuesta Pulso de la Migración:

#### · Análisis de consistencia

Pueden existir en la operación errores que no dependen de la selección de la muestra, como los que se pueden originar en el diseño del formulario, directamente de la fuente de información por mala interpretación de las instrucciones de diligenciamiento, del recolector por interpretación equivocada del instructivo de diligenciamiento o por deficiencias conceptuales, durante la captura, durante el manejo de los archivos, en la generación de resultados o en el manejo de las novedades por mala asignación.

Como se mencionó anteriormente, el supervisor de campo es el encargado de la revisión y análisis de consistencia de la información recolectada en el aplicativo web. Finalizado el periodo de recolección es el equipo logístico el que consolida y depura las bases de posibles inconsistencias. Estas bases se ponen a disposición del equipo donde se genera nuevamente un listado de posibles inconsistencias para someterse a revisión.

#### · Análisis de contexto

Se hace un análisis de contexto con los resultados de las rondas anteriores. Con el análisis de contexto se busca determinar qué tan acorde están los resultados con respecto a mediciones similares o paralela a los temas tratados por la EPM. Lo anterior, permite verificar la consistencia de la información en relación con los grandes grupos de actividades, pero sin que ello pueda entenderse como una comparación entre encuestas, pues las metodologías difieren en cada caso.

#### · Análisis de comparabilidad

Las comparaciones se realizan con los resultados mensuales de la EPM de acuerdo con las prevalencias de las diferentes preguntas que recoge la encuesta.

#### · Análisis estadístico

Se realizan los siguientes análisis:

- El análisis descriptivo, que ayuda a observar el comportamiento de la muestra en estudio, a través de tablas, gráficos.
- Se analiza en la muestra la estructura de los indicadores a partir de la distribución de frecuencias, y se detectan posibles inconsistencias.
- 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.
- Se verifica que, al momento de realizar desagregación de resultados, los coeficientes de variación no sean tan altos que impidan la inferencia estadística.

#### · Análisis univariado

Consiste en el análisis de cada una de las variables estudiadas por separado, es decir, el análisis está basado en una sola variable. Las técnicas más frecuentes de análisis univariado son la distribución de frecuencias, el análisis de las medidas de tendencia central y de las medidas de dispersión de la variable. La distribución de frecuencias de la variable requiere la observación de cómo están distribuidas las categorías de la variable, pudiendo presentarse como valores absolutos o en términos relativos, estas suelen presentarse acompañadas de gráficos estadísticos (como histogramas, tortas, etc.) para facilitar su comprensión y análisis.

#### · Análisis bivariado

No aplica.

#### · Análisis multivariado

No aplica.

### 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, solo con las encuestas completas, se genera un archivo que remite al área de sistemas. Dicha área, luego de realizar ese proceso, reenvía el archivo al área temática para su verificación y aprobación y posterior publicación en la página web de DANE.

De acuerdo con los protocolos de anonimización, no se incluyen variables relacionadas con la identificación de las viviendas como son su dirección, barrio, número telefónico, etc., y datos de identificación de las personas. Debido a que la EPM es una submuestra de la GEIH, se permite que las variables de directorio, secuencia y orden sean usadas como llaves de identificación y así usar los datos de la GEIH del mismo mes del año anterior. Con este proceso se busca garantizar la reserva estadística y a su vez facilitar el uso de la información.

### VERIFICACIÓN DE LA ANONIMIZACIÓN DE MICRODATOS

El equipo temático revisará que las variables, que por confidencialidad deben ser anonimizadas en la base de datos, no se encuentren publicadas en la página web de DANE.

### COMITÉS DE EXPERTOS

Dado que la Encuesta Pulso de la Migración es una estadística experimental, no existe un comité experto oficial, sin embargo, se tienen mesas de trabajo con Banco Mundial, USAID, el Banco Interamericano de Desarrollo para definir las temáticas, cuestionarios, la muestra, la metodología de recolección y se discuten los resultados de manera que se respalda la oficialidad de la encuesta.