

**DIRECCIÓN DE REGULACIÓN, PLANEACIÓN,
ESTANDARIZACIÓN Y NORMALIZACIÓN
-DIRPEN-**

COORDINACIÓN DE REGULACIÓN ESTADÍSTICA

**GUÍA PARA LA ELABORACIÓN DEL DOCUMENTO
METODOLÓGICO DE OPERACIONES ESTADÍSTICAS**

Mayo - 2020

**DEPARTAMENTO ADMINISTRATIVO
NACIONAL DE ESTADÍSTICA
(DANE)**

JUAN DANIEL OVIEDO ARANGO

Director

RICARDO VALENCIA RAMÍREZ

Subdirector

MARÍA FERNANDA DE LA OSSA ARCHILA

Secretaría General

DIRECTORES TÉCNICOS

ANTONIO JOSÉ AVENDAÑO AROSEMENA

Dirección de Metodología y Producción Estadística

LILIANA ACEVEDO ARENAS

Dirección de Censos y Demografía

JULIETH ALEJANDRA SOLANO VILLA

Dirección de Regulación, Planeación,
Estandarización y Normalización

JOVANA ELIZABETH PALACIOS MATALLANA

Dirección de Síntesis y Cuentas Nacionales

SANDRA LILIANA MORENO MAYORGA

Dirección de Geoestadística

MAURICIO ORTIZ GONZÁLEZ

Dirección de Difusión, Mercadeo y Cultura
Estadística

© DANE, 2020

Prohibida la reproducción total o parcial sin
permiso o autorización del Departamento
Administrativo Nacional de Estadística,
Colombia.

LUISA FERNANDA SUÁREZ LEÓN

Coordinadora de Regulación

LUIS ALEJANDRO MONTENEGRO RAMIREZ

CLAUDIA MILENA SANCHEZ CRUZ

Coordinación de Regulación

SONIA MARCELA NARANJO MORALES

Corrección de estilo

CONTENIDO

Introducción	6
I. Estructura del documento metodológico	8
INTRODUCCIÓN	10
1. ANTECEDENTES	10
2. DISEÑO DE LA OPERACIÓN ESTADÍSTICA	10
2.1. DISEÑO TEMÁTICO	11
2.1.1. <i>Necesidades de información</i>	11
2.1.2. <i>Formulación de objetivos</i>	11
2.1.3. <i>Alcance</i>	13
2.1.4. <i>Marco de referencia</i>	13
2.1.5. <i>Definición de variables y construcción de indicadores estadísticos</i>	14
2.1.6. <i>Plan de resultados</i>	15
2.1.7. <i>Estándares estadísticos utilizados</i>	16
2.1.8. <i>Diseño del cuestionario</i>	17
2.1.9. <i>Normas, especificaciones o reglas de edición e imputación de datos</i>	18
2.2. DISEÑO ESTADÍSTICO	18
2.2.1. <i>Universo de estudio</i>	19
2.2.2. <i>Población objetivo</i>	19
2.2.3. <i>Cobertura geográfica</i>	19
2.2.4. <i>Desagregación geográfica</i>	20
2.2.5. <i>Desagregación temática</i>	20
2.2.6. <i>Fuentes de datos</i>	20
2.2.7. <i>Unidades estadísticas</i>	20
2.2.8. <i>Período de referencia</i>	21
2.2.9. <i>Período de recolección/acopio</i>	21

2.2.10.	<i>Marco estadístico</i>	22
2.2.11.	<i>Diseño muestral</i>	22
2.2.12.	<i>Ajustes de cobertura (o ajuste de cobertura por no respuesta)</i>	24
2.2.13.	<i>Especificaciones de ponderadores</i>	24
2.3.	DISEÑO DE LA RECOLECCIÓN/ ACOPIO	25
2.3.1.	<i>Métodos y estrategias de recolección o acopio de datos</i>	25
2.3.2.	<i>Estructura organizacional del operativo y conformación del equipo</i>	26
2.3.3.	<i>Esquema de entrenamiento del personal</i>	27
2.3.4.	<i>Convocatoria y selección del personal</i>	27
2.3.5.	<i>Proceso de sensibilización y acuerdos de intercambio</i>	28
2.3.6.	<i>Elaboración de manuales</i>	28
2.3.7.	<i>Diseño de la estrategia de comunicación y plan de contingencias</i>	28
2.3.8.	<i>Diseño de la estrategia de seguimiento y control</i>	29
2.3.9.	<i>Diseño de sistemas de captura de datos</i>	29
2.3.10.	<i>Transmisión de datos</i>	30
2.4.	DISEÑO DEL PROCESAMIENTO	30
2.4.1.	<i>Consolidación de archivos de datos</i>	31
2.4.2.	<i>Codificación</i>	31
2.4.3.	<i>Diccionario de datos</i>	31
2.4.4.	<i>Revisión y validación</i>	32
2.4.5.	<i>Diseño de instrumentos de edición (validación y consistencia) e imputación de datos.</i>	32
2.4.6.	<i>Diseño para la generación de cuadros de resultados</i>	32
2.5.	DISEÑO DEL ANÁLISIS	33
2.5.1.	<i>Métodos de análisis de resultados</i>	33
2.5.2.	<i>Anonimización de microdatos</i>	35
2.5.3.	<i>Verificación de la anonimización de microdatos</i>	35
2.5.4.	<i>Comités de expertos</i>	36
2.6.	DISEÑO DE LA DIFUSIÓN Y COMUNICACIÓN	36
2.6.1.	<i>Diseño de los sistemas de salida</i>	36
2.6.2.	<i>Diseño de productos de comunicación y difusión</i>	36
2.6.3.	<i>Entrega de productos</i>	37
2.6.4.	<i>Estrategia de servicio</i>	37
2.7.	DISEÑO DE LA EVALUACIÓN DE LAS FASES DEL PROCESO	37
2.8.	DISEÑO DE LOS SISTEMAS DE PRODUCCION Y FLUJOS DE TRABAJO	38
3.	DOCUMENTACIÓN RELACIONADA	38

GLOSARIO	38
BIBLIOGRAFÍA	38
ANEXOS	39
Bibliografía	40

Introducción

El DANE, en su compromiso por fortalecer la calidad de la información estadística, y como coordinador del Sistema Estadístico Nacional (SEN) y de acuerdo con lo indicado por el artículo 155 de la Ley 1955 del 2019 establece las condiciones y las características que deben cumplir las estadísticas oficiales en Colombia y brinda lineamientos, normas y estándares que favorezcan la producción estadística a nivel nacional. En este escenario, se elabora la Norma Técnica de Calidad del Proceso Estadístico (NTCPE 1000) y los Lineamientos para el Proceso Estadístico para el SEN, en los que se exige la documentación de cada una de las fases del proceso en el desarrollo de una operación estadística.

En consecuencia, se brindan directrices que permiten realizar la documentación metodológica de forma estandarizada, recopilando los elementos requeridos a lo largo de las fases del proceso estadístico. El documento se constituye en el metadato principal, en dos sentidos: precisa los criterios técnicos-estadísticos para el desarrollo de la operación estadística, y proporciona la información necesaria para el uso y la interpretación correcta de la información estadística.

Dada su importancia, el documento comienza a construirse en la fase 2 y puede concluir al final de la fase 3, luego de llevar a cabo las pruebas y dependiendo de los hallazgos de estas será actualizado para disponer de una versión definitiva. Por lo anterior, el DANE pone a disposición una versión unificada y actualizada de la guía donde se presenta el paso a paso para la elaboración de la documentación metodológica para operaciones estadísticas que usan diversas fuentes de datos.

El objetivo de esta guía es suministrar las pautas generales relativas a la estructura, la presentación y el contenido del documento metodológico, empleando criterios estandarizados¹ y armonizados²; de la misma manera que tiene en cuenta referentes internacionales.

Este documento se estructura en tres partes: en la primera parte se presentan las generalidades de las operaciones estadísticas; la segunda hace una descripción de cada uno de los elementos que hacen parte de la documentación metodológica, y en la tercera se relaciona la bibliografía de referencia útil sobre los

¹ Entiéndase como la homogenización de la información en aspectos tales como conceptos, clasificaciones y métodos utilizados en la producción de los datos con el fin de comparar sus resultados.

https://www.dane.gov.co/files/planificacion/fortalecimiento/cuadernillo/Estandarizacion_documentos_metodologicos.pdf

² Proceso que busca un acuerdo común acerca del manejo de los datos, de tal forma que la información producida sea de fácil interpretación, además de poderse comparar con series de datos producidos por pares internacionales.

tópicos tratados. Cabe resaltar que el documento metodológico es un estándar que permite la conservación de la memoria de las operaciones estadísticas.

I. Estructura del documento metodológico

En la fase de diseño se llevan a cabo los diseños de la operación estadística que describen los métodos, las técnicas, los procedimientos y las estrategias que se aplicarán para reunir, procesar los datos y analizar e interpretar los resultados. Consecuentemente, se requiere que los diseños sean documentados para asegurar la memoria de la operación y estandarizar el desarrollo de esta.

Para permitir una visión clara de la forma en que se lleva a cabo la planeación y el desarrollo de la operación estadística dentro del documento metodológico se describen aspectos como: las necesidades de información, los objetivos generales y específicos; la contextualización temática; los métodos y los componentes estadísticos que harán parte de la operación, entre otros. De igual manera comprende los diseños: temático; estadístico; recolección/acopio, procesamiento, análisis, difusión y comunicación, evaluación, y el diseño de los sistemas de producción y de los flujos de trabajo.

La estructura del documento metodológico de las operaciones estadísticas, que se adelanten en el SEN, está organizada teniendo en cuenta los subprocesos del proceso estadístico, desde la perspectiva del diseño (fase 2), con el fin de brindar todos los elementos técnicos y metodológicos que enmarcan el desarrollo de la operación en las demás fases.

Gráfica 1. Estructura del documento metodológico

INTRODUCCION

1. ANTECEDENTES

2. DISEÑO DE LA OPERACIÓN ESTADÍSTICA

2.1. DISEÑO TEMÁTICO

- 2.1.1. Necesidades de información
- 2.1.2. Formulación de objetivos
- 2.1.3 Alcance
- 2.1.4 Marco de referencia
- 2.1.5 Definición de variables y construcción de indicadores estadísticos
- 2.1.6 Plan de resultados
- 2.1.7 Estándares estadísticos utilizados
- 2.1.8 Diseño del cuestionario
- 2.1.9 Normas, especificaciones o reglas de edición e imputación de datos

2.2. DISEÑO ESTADÍSTICO

- 2.2.1 Universo de estudio
- 2.2.2 Población objetivo
- 2.2.3 Cobertura geográfica
- 2.2.4 Desagregación geográfica
- 2.2.5 Desagregación temática
- 2.2.6 Fuentes de datos
- 2.2.7 Unidades estadísticas
- 2.2.8 Período de referencia
- 2.2.9. Período de recolección/acopio
- 2.2.10 Marco estadístico (censal o muestral)
- 2.2.11 Diseño muestral
- 2.2.12 Ajustes de cobertura (o ajuste de cobertura por no respuesta)
- 2.2.13 Especificaciones de ponderadores

2.3. DISEÑO DE LA RECOLECCION/ACOPIO

- 2.3.1 Métodos y estrategias de recolección o acopio de datos
- 2.3.2 Estructura organizacional del operativo y conformación del equipo
- 2.3.3 Esquema de entrenamiento de personal
- 2.3.4 Convocatoria y selección de personal
- 2.3.5 Proceso de sensibilización y acuerdos de intercambio
- 2.3.6 Elaboración de manuales
- 2.3.7 Diseño de las estrategias de comunicación y plan de contingencias
- 2.3.8 Diseño de la estrategia de seguimiento y control
- 2.3.9 Diseño de sistemas de captura
- 2.3.10 Transmisión de datos

2.4 DISEÑO DEL PROCESAMIENTO

- 2.4.1 Consolidación de archivos de datos
- 2.4.2 Codificación
- 2.4.3 Diccionario de datos
- 2.4.4 Revisión y validación
- 2.4.5 Diseño de instrumentos de edición (validación y consistencia) e imputación de datos
- 2.4.6 Diseño para la generación de cuadros de resultados

2.5 DISEÑO DEL ANÁLISIS

- 2.5.1 Métodos de análisis
- 2.5.2 Anonimización de microdatos
- 2.5.3 Verificación de la anonimización de microdatos
- 2.5.4 Comités de expertos

2.6. DISEÑO DE LA DIFUSIÓN Y COMUNICACION

- 2.6.1 Diseño de sistemas de salida
- 2.6.2 Diseño de productos de comunicación y difusión
- 2.6.3 Entrega de productos
- 2.6.4 Estrategia de servicio

2.7. DISEÑO DE LA EVALUACIÓN DE LAS FASES DEL PROCESO

2.8. DISEÑO DE LOS SISTEMAS DE PRODUCCION Y FLUJOS DE TRABAJO

3. DOCUMENTACIÓN RELACIONADA

GLOSARIO

BIBLIOGRAFÍA

ANEXO

A continuación, se presenta una descripción sobre el contenido de cada uno de los ítems a relacionar, con el fin de apoyar a los usuarios en la documentación metodológica de las operaciones estadísticas:

INTRODUCCIÓN

Es un resumen claro y sencillo que describe el contexto temático para todos los usuarios de la operación estadística. Se recomienda abordar los siguientes elementos:

- La descripción y la ubicación del contexto temático de la operación estadística e incorporar los principales antecedentes teóricos y prácticos.
- La justificación y la importancia de la operación estadística donde se identifica tanto la fuente de datos como las razones por las cuales se realiza. Junto a esto se describe su importancia para el país.
- El propósito general de la operación estadística que se genera a partir del objetivo general y un resumen de los objetivos específicos que se consideren más relevantes.
- La presentación del documento metodológico se debe realizar de forma organizada y describir la estructura del documento; se aconseja realizar una descripción mínima de todos los capítulos o las secciones que componen la metodología.

1. ANTECEDENTES

Presenta el inicio y la evolución histórica de la operación estadística, así como sus modificaciones más relevantes. Esta información permite contextualizar a los usuarios sobre los estudios y el trabajo que se ha adelantado y evidencia la experiencia acumulada en la realización de la operación estadística e identifica las características de los estudios que abordan temáticas similares. Además, brinda la posibilidad de construir una visión comparativa acerca de los cambios metodológicos que pueden incidir en la obtención y el análisis de los resultados.

2. DISEÑO DE LA OPERACIÓN ESTADÍSTICA

El diseño determina los métodos y los procedimientos con los que se reunirán, procesarán y analizarán los datos para cumplir los objetivos y satisfacer las necesidades de información.

Los diseños se validan y a su vez se pueden ajustar de acuerdo con los resultados de las pruebas que se desarrollan en la fase 3 (construcción), de la misma manera los diseños también se actualizan con el rediseño de la operación estadística a partir de nuevas necesidades de información, el cambio de las existentes y los resultados de la evaluación final que se realiza en la fase 8 (evaluación).

2.1. DISEÑO TEMÁTICO

Documenta las necesidades de información que se han consultado y analizado y que han permitido formular los objetivos de la operación estadística e incluye: el alcance, la contextualización y la delimitación del fenómeno u objeto de estudio desde el punto de vista teórico, conceptual, normativo, soportado con referentes nacionales e internacionales. Igualmente, incluye variables, indicadores estadísticos y el plan de resultados.

Asimismo, relaciona los estándares estadísticos que requiere la operación estadística para lograr la integración de los datos, la comparabilidad y la interoperabilidad de los sistemas de información.

2.1.1. Necesidades de información

Esta fase del diseño documenta las necesidades de información identificadas y confirmadas en interacción con los usuarios, para ser satisfechas con la operación estadística. Esta descripción contiene los requerimientos sectoriales de información, de organismos internacionales y el aporte de la operación estadística para el diseño de políticas públicas, cuando sea del caso. A su vez, es necesario describir los criterios utilizados para la priorización de necesidades y presentar los resultados de la aplicación de dichos criterios, al igual que los mecanismos.

2.1.2. Formulación de objetivos

Los objetivos de la operación estadística delimitan de manera clara y concreta las necesidades de información, que son acordadas con los usuarios de la información y están sujetas a las restricciones presupuestales. Los objetivos no requieren ni preámbulos ni explicaciones posteriores, lo mejor es expresarlos de manera concreta y que no tengan opción de distorsión ni de interpretaciones incorrectas.

Este apartado relaciona los objetivos definidos para la operación estadística. Estos se dividen en objetivo general y objetivos específicos.

- a) Objetivo general:** documenta el logro o el propósito por alcanzar de la operación estadística. También puede considerarse como la representación global del resultado al que se quiere llegar y que debe estar acorde con la pregunta central de investigación planteada, en términos del tipo de análisis, objeto, la población de estudio, delimitación geográfica y temporal.

Para las operaciones estadísticas a partir de registros administrativos el objetivo debe corresponder a la finalidad de la información estadística que se producirá y no a la del registro administrativo del que proviene la información.

- b) Objetivos específicos:** documentan los propósitos que se pretende lograr, exponen de manera clara y precisa los alcances y resultados que se desean obtener en las diferentes etapas o fases del proceso estadístico, y dar una respuesta o solución al objetivo general. Los objetivos específicos deben ser precisos, medibles, evaluables, viables y seguir un diseño metodológico.

Los objetivos deben cumplir con las siguientes características:

- Ser claros y estar muy bien definidos de manera que no haya duda en aquellos que son responsables de participar en su logro.
- Flexibles para ser modificados cuando las circunstancias de la operación estadística lo requieran. Tener productos (intermedios y finales) asociados.
- Coherentes y tener en cuenta que deben responder a las necesidades identificadas y priorizadas.

La formulación de un objetivo contiene los siguientes elementos asociados a las respectivas preguntas que se pretenden responder:

Objetivo = verbo + unidades de análisis o casos + propósito + alcance geográfico + espacio temporal.

Que corresponde a:

Preguntas = ¿Qué? + ¿Sobre qué? + ¿Para qué? + ¿Dónde? + ¿Cuándo?

Con el fin de ilustrar un objetivo se presenta el siguiente ejemplo: “Proporcionar información básica sobre el tamaño y la estructura de la fuerza de trabajo (empleo, desempleo e inactividad) de la población del país, así como de las características sociodemográficas de la población colombiana”.

Respuesta a *qué*: proporcionar información básica; *sobre qué*: fuerza de trabajo; *para qué*: determinar el tamaño y la estructura de la fuerza de trabajo; *dónde*: Colombia; *cuándo*: continuamente (esta es una operación estadística de carácter continuo).

2.1.3. Alcance

Describe la delimitación temática, geográfica y temporal de la operación estadística. Con el alcance se precisan los aspectos que se investigarán y sobre los que se presentarán los resultados; adicionalmente, relaciona aquellos temas que no serán tomados en cuenta justificando su exclusión.

2.1.4. Marco de referencia

Su construcción se realiza a partir de la revisión de literatura existente relacionada con el objetivo de la operación estadística, así como aspectos que en su conjunto contextualizan el fenómeno de estudio de la operación. Está compuesto por: el marco teórico, el marco conceptual, el marco legal o normativo y los referentes internacionales y nacionales empleados.

a) Marco teórico: corresponde a la literatura que sustenta la operación estadística, desde la óptica temática y el resumen de la revisión realizada de los aportes que fundamentan el análisis de teorías, planteamientos y enfoques. Del mismo modo el marco teórico permite comprender el problema, delimitar el problema de estudio, conocer la teoría temática de estudios y a buscar la generación de nuevos enfoques en la forma de abordar el problema a resolver.

b) Marco conceptual: corresponde a la descripción de los principales conceptos o términos técnicos que contextualizan la operación estadística.

Es necesario precisar los conceptos más importantes empleados en la operación estadística y sus interrelaciones; estos deben ser expresados claramente sin extenderse en detalles. Deben estar escritos de manera clara y concisa para que el equipo de trabajo, los entrevistados y los usuarios de la

información, se familiaricen con los mismos. Esto facilita el desarrollo de la operación, el procesamiento, el análisis y la adecuada lectura de los resultados.

Es importante que los conceptos que se utilicen en la operación estadística estén estandarizados por el DANE para garantizar la comparabilidad, la integración y la interoperabilidad de la información estadística.

c) Marco legal o normativo: describe la normatividad (acuerdos, constitución, sentencias de las cortes, leyes, decretos, resoluciones u otros actos administrativos) en la que se circunscribe la operación estadística. La recopilación de las normas legales se debe presentar en orden cronológico incluyendo su título, su objetivo y el organismo que la expide.

d) Referentes internacionales: describen los principales trabajos y las recomendaciones realizadas por organizaciones internacionales y que están relacionados con la temática de la operación estadística. Tales organizaciones pueden ser: la Organización de Naciones Unidas, los institutos nacionales de estadística, la Organización Internacional de Migraciones (OIM), la Organización Mundial de la Salud (OMS), la Organización Internacional del Trabajo (OIT), el Fondo Monetario Internacional (FMI), entre otras.

e) Referentes nacionales: en este punto se describen los principales trabajos realizados por entidades u organizaciones nacionales que han adelantado proyectos similares y que brindan aportes temáticos, metodológicos y de aseguramiento de la calidad.

2.1.5. Definición de variables y construcción de indicadores estadísticos

Enuncie y describa las principales variables³ a medir con la operación estadística y a emplear en el diseño de los indicadores estadísticos y de la operación, teniendo en cuenta que las categorías definidas correspondan con las utilizadas en operaciones similares para permitir la integración de datos y la comparabilidad de las estadísticas.

³ Desde el momento en que se definen las variables es necesario establecer y documentar aquellas, en las cuales los microdatos deben ser anonimizados para garantizar la confidencialidad y reserva de las unidades de información; con el propósito de dar cumplimiento a la Ley 1266 de 2008, que dicta las disposiciones generales del habeas data y regula el manejo de la información contenida en bases de datos personales, la Ley 1581 de 2012, que brinda las disposiciones generales para la protección de datos personales, la Ley 1712 de 2014, que se constituye en la ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y la Ley 79 de 1993 (art. 5) sobre reserva estadística; así mismo, con el Código Nacional de Buenas Prácticas estadísticas (principio 11).

La documentación describe la metodología definida para obtener los indicadores que se calculan o estiman para la operación estadística, presentar sus objetivos, fórmulas y la correspondiente forma de interpretación. Es fundamental que se elaboren y documenten con precisión y se identifiquen las variables necesarias para su cálculo porque inciden además en el diseño de los cuadros de resultados. Los indicadores están estrechamente relacionados con los objetivos, son expresiones cualitativas o cuantitativas que permiten describir características, comportamientos o fenómenos de la realidad, a través de una variable o la relación entre las variables de la operación estadística estos pueden ser: índices, razones, proporciones, promedios, porcentajes, totales o indicadores especiales⁴.

2.1.6. Plan de resultados

Describe de manera general el plan de resultados propuesto para la operación estadística, los resultados esperados, el tipo de reportes o de informes, cuadros de salida, gráficos, mapas temáticos, sistemas geoespaciales, atlas estadísticos y otros instrumentos que se van a generar de acuerdo con los objetivos de la operación estadística, los cuales deben describir los requerimientos de los usuarios de la información.

- **Cuadros de salida**

Los cuadros de salida o de resultados son un conjunto organizado de datos que están diseñados con el propósito de materializar los requerimientos de los usuarios. Estos presentan los resultados de la operación estadística, con sus respectivos niveles de desagregación temática y geográfica, junto con las anotaciones, los encabezados, los pies de página, etc., que facilitan el análisis y la interpretación de dichos resultados. Es importante que el título indique su contenido de forma clara y concisa y presente: el lugar al que se refiere la información, el fenómeno que se presenta, las variables que se subordinan y la fecha de referencia.

Dentro del contenido metodológico se pueden relacionar los títulos de los cuadros de salida para conocer los cruces de variables que se efectuarán y los cuadros se incluyen en un anexo o en un documento adjunto, dependiendo del volumen de estos.

⁴ También denominados *indicadores compuestos* o *sintéticos* corresponden a la agregación de indicadores parciales de un determinado componente en un único indicador.

2.1.7. Estándares estadísticos utilizados

Este apartado relaciona los estándares estadísticos que se usan en las diferentes fases de la operación estadística y que permiten la integración, la comparabilidad y la interoperabilidad de la información estadística.

Los estándares pueden ser:

- *Clasificaciones estadísticas*⁵.
- *Conceptos estandarizados*⁶.
- *Nomenclaturas*⁷.
- *Variables*⁸.
- *Unidades estadísticas*⁹.
- *Data Documentation Initiative DDI y Dublin Core*¹⁰.
- *Statistical Data and Metadata Exchange (SDMX)*¹¹.

⁵ Son aquellas que agrupan fenómenos (económicos, culturales o sociales) de manera sistemática con criterios preestablecidos y en función del uso de la clasificación. Está conformada por: un conjunto de categorías que pueden ser asignadas a una variable específica registrada en una encuesta estadística o en un registro administrativo y se utilizan en la producción y la presentación de estadísticas.

⁶ Son definiciones unificadas basadas en estándares estadísticos internacionales fundamentales para la armonización estadística, con el fin de proporcionar los elementos necesarios para la construcción y fortalecimiento de los marcos conceptuales. Estos conceptos deben estar armonizados en estándares estadísticos internacionales fundamentales para la armonización estadística, con el fin de proporcionar los elementos necesarios para la construcción y fortalecimiento de los marcos conceptuales.

⁷ Son conjuntos de códigos numéricos o alfanuméricos que sirven para identificar diferentes elementos. Ejemplo: países, unidades territoriales, operaciones estadísticas, entidades estatales.

⁸ Es una característica de una unidad observada, que puede asumir un conjunto de valores los cuales pueden ser una medida numérica o una categoría de una clasificación asignada. Nota: para el caso del directorio estadístico representa una Característica que al ser medida en diferentes empresas es susceptible de adoptar diferentes valores.

⁹ Corresponde a las unidades de observación, las unidades de análisis y, cuando aplique, las unidades de muestreo.

¹⁰ Son estándares que, acompañados de un software especializado de uso libre, el Nesstar Publisher, permiten realizar procesos de documentación de operaciones estadísticas donde los resultados son visibles en un portal centralizado llamado Archivo Nacional de datos (ANDA). En dicho software se configuran las plantillas que permiten realizar el proceso de transmisión, difusión de metadatos y microdatos y enlaza y preserva todo el material relacionado de la operación estadística que el responsable de la operación estadística considere de gran importancia (documento metodológico, ficha metodológica, bases de datos, guías, manuales etc.).

¹¹ es un estándar impulsado por Eurostat, FMI, OCDE, Naciones Unidas, Banco Mundial, Banco Central Europeo y el BIS, que fortalece los procesos de integración a través de proyectos de interoperabilidad que permitan a los sistemas de información de las entidades modernizar sus intercambios de datos estadísticos y los metadatos necesarios para entender como están estructurados, y a su vez presentar los datos en series de tiempo.

Todos los estándares que utilicen las operaciones estadísticas deben corresponder a versiones actualizadas y adaptadas para Colombia, teniendo en cuenta los referentes internacionales para garantizar la comparabilidad y la integración de la información estadística.

2.1.8. Diseño del cuestionario

Describa la estructura del cuestionario definido para la operación estadística (aplica para aquellas operaciones estadísticas que obtienen los datos mediante este instrumento).

El contenido del cuestionario se define de acuerdo con los objetivos de la operación estadística, traduciendo estos objetivos y necesidades de información en indicadores, cuadros de salida y variables para llegar finalmente a las preguntas. La cantidad de preguntas debe reducirse a las estrictamente necesarias para cumplir con el plan de resultados y para obtener la información estadística que se presentará en los cuadros de salida.

El diseño del cuestionario debe basarse en los indicadores que expresan los resultados y en la forma como van a ser recolectados los datos. El cuestionario se estructura generalmente en bloques temáticos o capítulos conformados por grupos de preguntas que tienen una secuencia lógica.

Es necesario conocer y documentar los objetivos de cada uno de los módulos o los capítulos que lo componen, así como las preguntas, lo cual se convierte en un recurso fundamental al momento de evaluar la modificación o la exclusión de una pregunta del cuestionario o de evaluar el instrumento de recolección o de su aplicación en campo, incluso para el análisis de los resultados. El formato del cuestionario debe hacer parte de los anexos.

Para el caso de operaciones estadísticas continuas es fundamental definir los criterios necesarios para incluir o eliminar preguntas del formulario o cuestionario, ya que se puede afectar el diseño y la calidad de la información que se recolecte, al aumentar la carga al encuestado.

Otro aspecto esencial es la definición de pruebas al cuestionario. Las pruebas buscan que el cuestionario sea coherente con el fenómeno de estudio y con la secuencia lógica de las preguntas y los flujos de información. Dada su importancia, se requiere la documentación del plan de pruebas que incluye: tipos de prueba, objetivos, desarrollo metodológico, operativo y logístico, instrumentos, tiempos, recursos y

resultados esperados. Posterior a la ejecución de dichas pruebas se deben elaborar informes de resultados con los correspondientes aportes al diseño.

2.1.9. Normas, especificaciones o reglas de edición e imputación de datos

Es fundamental documentar las reglas que se aplicarán durante la recolección y el procesamiento al conjunto de datos, con el fin de garantizar su completitud y consistencia. Generalmente las reglas se refieren a: los valores que en forma individual pueden asumir las variables (proceso de validación); las relaciones que existen entre unas variables y otras (proceso de consistencia); la estructura global que se espera de un conjunto de datos, y la forma como se asigna el valor de un dato faltante o inconsistente (proceso de imputación).

Las reglas deben ser exhaustivas para garantizar los criterios mencionados anteriormente, puesto que cualquier omisión en alguna situación implicaría inconsistencias en la base de datos final e induciría a errores en los resultados y en las decisiones o las acciones.

En el documento metodológico se describen, de forma general, las reglas o las especificaciones establecidas para la edición y la imputación de los datos, y en un documento específico se registra con todo detalle el conjunto de reglas o especificaciones. De igual manera se deben relacionar dentro del documento metodológico los manuales de validación y consistencia y los manuales de imputación (cuando apliquen).

2.2. DISEÑO ESTADÍSTICO

Consiste en la definición, la construcción y/o la formulación del marco estadístico, el universo, la población, las unidades estadísticas, los periodos y el diseño muestral, este último para el caso de operaciones estadísticas por muestreo.

Para el caso de operaciones estadísticas por censo se diseña el marco censal y para las operaciones por muestreo se define el diseño muestral donde se determina el tamaño, la selección de la muestra y establece la metodología estadística de estimación y dependiendo del tipo de muestreo se calculan los factores de expansión proporcionando las fórmulas para el cálculo de varianzas y los criterios de calidad de los resultados.

Para el caso de operaciones estadísticas que tienen registros, que trabajan a partir de ellos o que integran datos provenientes de otras operaciones estadísticas u otras fuentes de datos no estadísticos; se deben caracterizar las fuentes que contienen las unidades de observación y variables de estudio y define los criterios metodológicos necesarios para fortalecer las fuentes y realizar el aprovechamiento estadístico.

Los elementos básicos que cada operación estadística debe contener en su diseño estadístico son:

2.2.1. Universo de estudio

Describe el total de unidades o elementos a los que se refiere la operación estadística o constituyen la colectividad de interés y que satisfacen una definición común. Refleja los objetivos del estudio, por cuanto los describe en términos de contenido, unidades, espacio y tiempo y debe justificar la exclusión o la inclusión de grupos en el universo.

2.2.2. Población objetivo

Es el conjunto de unidades o elementos definidos en un espacio y en un tiempo sobre los que se desea obtener los datos y generar estadísticas.

Puede ser igual al universo o un subconjunto de éste. En la documentación se explican las dificultades y razones por las cuales se limita la población, razones que pueden ser de diversa índole: presupuestales, operativas, políticas, económicas o sociales, o no son objeto de estudio.

Es necesario especificar las unidades y sus características, así como delimitarlas geográfica y temporalmente.

2.2.3. Cobertura geográfica

Describe la extensión territorial sobre la que se ejecuta una operación estadística; puede ser nacional, regional, departamental (DANE, 2019).

2.2.4. Desagregación geográfica

Presenta el nivel de detalle de una división territorial con el que se agrega la información para su difusión, de acuerdo con la cobertura geográfica (total, nacional, departamental, entre otros). Esto se determina en la planeación de la operación estadística (INEGI, 2010b).

2.2.5. Desagregación temática

Hace referencia al nivel de detalle del tema de estudio con que se va a generar y difundir la información estadística, igualmente se establece en la planeación de la operación estadística.

2.2.6. Fuentes de datos

Describe las fuentes mediante las cuales se obtiene la información estadística las cuales pueden ser primarias o secundarias.

La *fente primaria* es aquella información que se obtiene directamente de las unidades que suministran los datos y corresponden a operaciones estadísticas por censo y muestreo. En la documentación para este criterio se debe proporcionar el número de fuentes.

La *fente secundaria* es aquella que ofrece información sobre el tema que se va a investigar, pero que no son la fuente original de los hechos o las situaciones, es decir, que no se obtiene directamente de las unidades sino a partir de otras operaciones estadísticas o de registros administrativos. Cuando este sea el caso debe registrarse la cantidad de fuentes y relacionarlas.

2.2.7. Unidades estadísticas

Se define el conjunto de unidades empleadas para el diseño, la medición y la presentación de resultados de las operaciones estadísticas. Las unidades estadísticas son:

- a) **Unidad de observación:** elemento o conjunto de elementos sobre los que se hace medición de las diferentes variables en una operación estadística, que se obtiene la información sobre sus características dependiendo de la operación estadística. Estas pueden ser: personas, hogares,

viviendas, cultivos, empresas establecimientos, parcelas o predios agrícolas, áreas administrativas, entre otras; pueden existir varias unidades de observación en la misma operación estadística.

b) Unidad de análisis: elemento de estudio en una medición y sobre la que se presentan los resultados o las conclusiones de la operación estadística y puede la operación estadística puede tener varias unidades de análisis. Por ejemplo, en la operación estadística “Estadísticas Vitales” la unidad de análisis son los nacidos vivos y los fallecidos, mientras que para las encuestas agropecuarias pueden ser el predio, la finca y la unidad productora.

Para operaciones estadísticas por muestreo se debe tener en cuenta:

c) Unidad de muestreo: describe el elemento o el conjunto de elementos definidos o establecidos a ser seleccionados en la muestra. Una operación estadística puede tener varias unidades de muestreo y es necesario describir sus características.

En un diseño en varias etapas se debe definir una unidad muestral para cada una de ellas; las unidades de muestreo en cada etapa se submuestra de la unidad (más grande) elegida en la etapa anterior. Adicionalmente, se definen específicamente todas las unidades de observación y se diseñan las unidades de análisis correspondientes a los diferentes niveles de desagregación con los cuales se presentarán los resultados.

2.2.8. Período de referencia

Documenta el intervalo de tiempo al que hace referencia el conjunto de información de la operación estadística (año, semestre, trimestre mes, semana o día). Igualmente, precisa el intervalo de tiempo al que van referidas las variables, dado que una operación estadística puede tener diferentes periodos de referencia por la variedad de temas o la frecuencia de ocurrencia de los hechos objeto de estudio.

2.2.9. Período de recolección/acopio

Describe el intervalo de tiempo que se ha determinado para recolectar o acopiar los datos, según el tipo de operación estadística. Para el caso de operaciones por censo o muestreo, el tiempo que dure el operativo de campo y para el caso de operaciones a partir de registros administrativos o derivadas, el intervalo establecido para el acopio de las bases de datos.

2.2.10. Marco estadístico (censo o muestral, aplica para operaciones estadísticas por censo o por muestreo, en los demás casos omite este apartado)

Este se define como el instrumento que permite identificar y ubicar los elementos de la población objetivo. Se denomina marco muestral o marco censo según el tipo de operación estadística ya sea por muestreo o por censo respectivamente. Su diseño debe contener las variables definidas que permiten caracterizar, identificar y ubicar las unidades de observación, así como la metodología de construcción y actualización del marco.

Es necesario documentar las especificaciones, los criterios y los procedimientos a que hace referencia la información contenida e indicar el medio en que se encuentra el marco (listado en papel, cartografía, medios magnéticos, entre otros). Adicionalmente, el marco generado a partir de registros administrativos es útil para controlar la cobertura de las fuentes que suministran los datos necesarios en la generación de los resultados y analizar el contexto de las cifras con la información auxiliar que contenga.

2.2.11. Diseño muestral

Consiste en la definición del diseño, la selección de la muestra y la estimación de resultados, es necesario especificar las etapas, las fases y los algoritmos empleados en los procesos de selección. Para definirlo es de utilidad contar con elementos conceptuales básicos como los son universo, población, marco muestral, muestra y unidades estadísticas.

La metodología del plan muestral y su documentación deben incluir detalladamente toda la formulación matemática para su aplicación rigurosa durante el procesamiento y la fase de análisis; es necesario incluir en la documentación los ajustes realizados a partir de los hallazgos de las pruebas piloto desarrolladas.

Tipo de muestreo: describe el método de selección de la muestra. Justifica el diseño teniendo en cuenta las características de las variables, el marco muestral, los supuestos acerca de la distribución poblacional, el tipo de estudio, la población, etc. Es decir, define el diseño muestral empleado para lograr los objetivos de la operación estadística.

Detalla las ventajas de ese diseño en particular, argumentando, en los casos en que aplique, las etapas, las fases, los criterios de conformación, fijación y tamaño de los estratos o de los conglomerados y, los

mecanismos de selección de la muestra (tales como: Algoritmo Coordinado Negativo, Algoritmo Fan-Muller & Rezucha, entre otros).

Cálculo del tamaño de la muestra: documenta la prelación de los parámetros que se van a estudiar, parámetro base para la estimación, la metodología para el cálculo de los tamaños de muestra del estudio por etapas, estratos, niveles o áreas, según corresponda, determinando aspectos como: los elementos que intervienen; los costos de la operación estadística; los mecanismos de afijación¹², precisión y confiabilidad; efectos de diseño, así como la metodología para el mantenimiento de la muestra, cuando se requiera, de acuerdo con el tipo de medición. Describe los procedimientos utilizados y la frecuencia con que se aplican. Igualmente, la documentación metodológica debe presentar los algoritmos elaborados para la selección de la muestra.

Selección de la muestra: documenta la aplicación de la estrategia de selección de las unidades de muestreo de acuerdo con el método establecido garantizando que la selección sea aleatoria y que la probabilidad de selección de las unidades sea la que establece el método elegido. Todo ello a partir del marco muestral (actualizado) que permite acceder a la población objetivo y una vez determinado el diseño muestral.

En esta parte se aplican las fórmulas a utilizar en el cálculo del tamaño de muestra de la operación estadística, bien sea por etapas, estratos, niveles o áreas; determinando los elementos que intervienen y siguiendo el diseño muestral establecido. Es necesario anexar a la metodología del diseño los algoritmos desarrollados.

¹² Afijación: estrategia de distribución / reparto de la muestra en los diferentes estratos. Esta puede ser:

Afijación simple: consiste en el reparto a partes iguales de la muestra entre los estratos conocidos.

Afijación proporcional: consiste en el reparto proporcional de la muestra entre los distintos estratos, en base al número de efectivos de cada uno de los mismos (así se mantiene constante el coeficiente de elevación).

Afijación óptima: se trata del reparto de la muestra entre los estratos atendiendo al tamaño y a la varianza de cada uno de los mismos, manteniendo una dispersión constante para toda la muestra.

¹² Los Métodos de Calibración (Lundström, S. y Särndal, C.E. (1999 y 2005) 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. Entre algunos métodos de calibración tenemos: Calibración por Regresión, método Clanman, entre otros.

Las consecuencias por las debilidades que pueda presentar un marco muestral y las imprecisiones en la estructura del universo de población general estimada a partir de la muestra expandida se pueden corregir al máximo con base en un factor de ajuste estructural y de totales, en algunos casos se utiliza la información auxiliar por grupos de sexo, edad, estrato, nivel académico, etc... En los métodos de calibración se utiliza información auxiliar completa que tenga alta correlación con las variables principales de la operación estadística, y de esta manera poder producir estimaciones más eficientes. <file:///C:/Users/Alexander/Downloads/922-925-1-PB.pdf>

Ponderadores: describe los criterios aplicados para la generación de los ponderadores y/o factores de expansión y para determinar el tratamiento informático necesario. De igual forma se requiere la presentación de los tipos de ajustes que se llevan a cabo y su tratamiento informático.

Metodología de estimación: este procedimiento tiene las siguientes características: a) describe el procedimiento estadístico establecido para estimar los valores de cada uno de los parámetros a considerar en la población objetivo, b) presenta las fórmulas de los estimadores que se utilizan para obtener los parámetros, c) define y justifica la metodología elegida, la fórmula y sus componentes, d) indica cómo se realiza el cálculo de los factores de expansión y los elementos que allí intervienen, y e) describe los métodos de ajuste al factor de expansión, bien sea por no-respuesta total o ajustes por variables exógenas de alta calidad que captan la dinámica de la población objetivo.

Cálculo de precisión de los resultados: describe la metodología utilizada para estimar los errores de muestreo y su presentación para determinar el nivel de confianza. Reseña el diseño del método de estimación de la varianza. Da a conocer las fórmulas para calcular el error estándar y/o el coeficiente de variación de todos los estimadores.

Rotación: para el caso de las operaciones estadísticas por muestreo que son continuas es fundamental documentar los aspectos relacionados con la rotación de la muestra, que corresponden a las técnicas o los mecanismos diseñados para mantener la muestra actualizada.

2.2.12. Ajustes de cobertura (o ajuste de cobertura por no respuesta)

Describe los procedimientos para los ajustes de cobertura de la operación estadística. El ajuste de cobertura se da en los casos de pérdida de muestra (operaciones estadísticas por muestreo) para evitar sesgos en las estimaciones y en los censos, se ajusta por la no respuesta en los elementos (unidades) del universo que no se midieron por alguna razón.

2.2.13. Especificaciones de ponderadores

Luego de establecer las formulaciones matemáticas para el cálculo de los factores de expansión y las varianzas (en las operaciones estadísticas por muestreo), se elaboran y documentan las respectivas especificaciones para el desarrollo de los aplicativos de software necesarios para realizar las estimaciones o ponderaciones.

2.3. DISEÑO DE LA RECOLECCIÓN/ ACOPIO¹³

En esta sección se describe la manera como se obtienen los datos necesarios en una operación estadística, aquí se establecen aspectos de conformación de los equipos de trabajo, la cantidad de personal, los perfiles y los roles que se emplean para la recolección o acopio y el método. Adicionalmente se determinan los medios necesarios para la captura o el acopio y los mecanismos para la transmisión de los datos; también se establecen los mecanismos de seguimiento y control y se asegura la conformación de la base de datos originales de la operación estadística.

Para las operaciones estadísticas que utilizan fuentes alternativas relacionadas con instrumentos tecnológicos de medición (por ejemplo: sensores remotos, pluviómetros, hidrómetros, sicrómetros, anemómetros, etc.) la documentación describe:

- El uso de instrumentos tecnológicos de medición.
- Los procedimientos para la calibración de los instrumentos y para las pruebas con el fin de verificar su funcionamiento.
- El mecanismo para la consecución de imágenes satelitales y la identificación de los polígonos de interés (cuando aplique).
- El procedimiento de medición correspondiente de las variables establecidas metodológicamente y para la generación del archivo de datos originales.

2.3.1. Métodos y estrategias de recolección o acopio de datos

Se documenta el método (o métodos) a utilizar para la recolección o el acopio de los datos de las unidades de observación y se define la estrategia para el operativo de campo, si aplica. Por ejemplo: entrevista telefónica, entrevista directa, autodiligenciamiento por correo, autodiligenciamiento con cuestionario electrónico, autodiligenciamiento asistido con dispositivo móvil de captura (DMC), autodiligenciamiento con cuestionario impreso, entrevista por convocatoria, entre otros, teniendo en cuenta que en el desarrollo de una misma operación se pueden utilizar diferentes técnicas dependiendo de las necesidades y los alcances definidos.

¹³ La recolección refiere a la obtención de los datos directamente en campo, mientras que el acopio corresponde a la recepción de un archivo o base de datos.

Adicionalmente, se pueden prever y documentar otras estrategias como establecer puntos fijos, donde, en casos particulares, se pueden recopilar datos de personas que por motivos de vacaciones no se encuentran en su lugar de residencia; esta misma estrategia se puede emplear en la entrada de condominios combinada con la visita a las viviendas. En otros casos, en lugares lejanos con baja densidad de viviendas y personas se puede realizar convocatoria para efectuar las entrevistas en un lugar determinado, de fácil concurrencia.

Para operaciones que usan fuentes de datos secundarias se describe el procedimiento y la plataforma tecnológica para el acopio de los datos (estructurados¹⁴ o no estructurados¹⁵), especificando las variables requeridas, las técnicas para verificar la calidad de los datos, el protocolo para la transmisión y recepción; y la conformación de la base de datos originales de la operación estadística.

2.3.2. Estructura organizacional del operativo y conformación del equipo

Aquí se documentan: las estructuras establecidas geográfica y operativamente para el desarrollo del trabajo cuando se lleva a cabo en campo; la logística para la distribución de materiales y el flujo de las comunicaciones, teniendo en cuenta el tipo de fuente de datos; el método de recolección; el tipo de entrevista, y la cobertura geográfica definida, en el caso de operaciones estadísticas por censo o muestreo.

Documente los mecanismos o estrategias establecidas para el alistamiento y organización de los materiales y equipos, envío y transporte de estos elementos y su disposición en los puntos de uso; al igual que la logística para el retorno de los elementos devolutivos. Igualmente, es necesario documentar el esquema de custodia de los equipos que se usen en campo que se usen para la recolección, como es el caso de los Dispositivos Móviles de Captura (DMC).

Los esquemas geográficos son importantes a la hora de: la organización y la distribución de material (cartografía o mapas, fotografía, cuestionarios, formatos, manuales, etc.); el transporte, la selección y el

¹⁴ Son datos que cumplen con un modelo que implica un formato, longitud y tamaño predefinidos y que se ajustan a tablas de registros o filas con campos de significados fijos y relaciones o enlaces entre las tablas. Fuente: basado en la definición de Enterprise Big data framework. <https://www.bigdataframework.org/data-types-structured-vs-unstructured-data/> (recuperado en 24/04/2020). Es el caso, por ejemplo, de las operaciones estadísticas que acopian archivos de datos de operaciones tradicionales (censos y por muestreo) o de registros administrativos.

¹⁵ Son los datos que no tienen un modelo de datos predefinido, su estructura y modelo son flexibles y no se ajustan a un modelo relacional. Fuente: basado en la definición de Enterprise Big data framework. <https://www.bigdataframework.org/data-types-structured-vs-unstructured-data/> (recuperado en 24/04/2020).

entrenamiento de personal; la planeación de las entrevistas; la organización de la supervisión; el monitoreo, y el control.

Se documentan los roles definidos para llevar a cabo la recolección o acopio, se precisan las cantidades de personal, los perfiles acordes para cada rol y las responsabilidades de cada uno de ellos, siempre teniendo en cuenta la disponibilidad presupuestal, los objetivos de la operación estadística, la estructura geográfica y operativa, el método de recolección, el tipo de entrevista, el rendimiento promedio para el diligenciamiento, entre otros factores.

2.3.3. Esquema de entrenamiento del personal

Describe el modelo, los métodos, las modalidades, las estrategias y los materiales pedagógicos para realizar el entrenamiento del personal que recolecta o acopia los datos. Se requiere precisar, entre otros aspectos, la estructura, los procesos o los subprocesos que la componen, los medios y las ayudas audiovisuales, los perfiles, el número de los instructores o los formadores, la logística y demás requerimientos que deben cumplirse para llevar a cabo el proceso de entrenamiento.

Es importante describir otros aspectos como la infraestructura local en los sitios donde se va a desarrollar el entrenamiento, la disponibilidad de recursos tecnológicos, las estrategias ante la ausencia de los perfiles requeridos para cada uno de los roles, si se prevé el pago de alquiler de equipos y espacios, entre otros recursos. De igual manera, documenta si la logística para el entrenamiento la realizará el personal de la entidad productora o se contratará a través de un operador logístico.

Igualmente, se documentan los mecanismos de evaluación y verificación de asimilación de conceptos y contenidos, como talleres, simulaciones y pruebas en campo.

2.3.4. Convocatoria y selección del personal

Documenta los procedimientos generales para la convocatoria, selección y contratación del personal, teniendo en cuenta los perfiles requeridos para los diferentes roles, incluyendo los instructores cuando sea necesario. Puede considerarse, por ejemplo, que luego de la convocatoria del personal se realice una primera selección con unos criterios básicos, se lleve a cabo el entrenamiento (virtual o presencial), se evalúen los conocimientos y destrezas adquiridas, y al personal con los mejores resultados se le contrate para desarrollar las actividades determinadas.

A su vez, documenta los procedimientos que deben ser acordes con la legislación vigente y la política de contratación de la entidad productora de información estadística y el esquema de recolección establecido.

2.3.5. Proceso de sensibilización y acuerdos de intercambio

Documenta el conjunto de acciones encaminadas a generar conciencia en la población objetivo, acerca de la importancia de la operación estadística. Describe los mecanismos, los textos, los medios y los instrumentos que serán utilizados en la sensibilización¹⁶, al igual que los actores que serán multiplicadores de los contenidos a divulgar e instancias objeto de la actividad, para optimizar la comunicación con la unidad de información o de reporte. Incluye un plan de difusión de los objetivos, los métodos y las características de la recolección, así como la importancia y los beneficios de los datos que requiere la operación estadística.

Es importante tener presente, para el caso de operaciones estadísticas que utilizan fuentes de datos secundarias, la inclusión de acuerdos de intercambio con los responsables de estas fuentes, para asegurar la entrega de los archivos en las condiciones y los plazos requeridos.

2.3.6. Elaboración de manuales

Esta sección relaciona los manuales, las guías o los instructivos elaborados para cada rol o instrumento, con el fin de dar claridad frente a los conceptos y los procesos establecidos; presenta las directrices definidas para el desarrollo de la operación estadística y los procedimientos estandarizados. Es importante incorporar en este apartado un vínculo para que el usuario pueda acceder a ellos.

2.3.7. Diseño de la estrategia de comunicación y plan de contingencias

Documente el flujo de comunicaciones teniendo en cuenta la jerarquía organizacional establecida y los actores involucrados, este último cuando se trate de fuentes secundarias, los medios y los recursos disponibles, los diferentes niveles de complejidad en las interacciones, señalando los roles o actores que intervienen, las fuentes a consultar en el surgimiento de un evento y la solución que debe efectuarse.

¹⁶ Es un proceso de comunicación, activo y creativo, que promueve la transformación, el cambio de actitud y el comportamiento en la sociedad. Pretende lograr una toma de conciencia respecto a una determinada temática (DANE. Material de sensibilización para censistas y supervisores. 3er. CNA. 2014).

Todos los miembros de la estructura funcional deben conocerlo y acatarlo porque solo a través de este mecanismo se solucionarán efectivamente los eventos que se presenten.

Asimismo, se documenta un plan de contingencias en el que se proporcionan las orientaciones para el manejo y la solución de situaciones imprevistas que permitan mantener bajo control el proceso y cumplir las metas establecidas. Las estrategias planteadas deben comunicarse siguiendo el flujo de comunicaciones. Igualmente, se plantean acciones para monitorear los resultados de las estrategias en los casos contingentes.

2.3.8. Diseño de la estrategia de seguimiento y control

Documente la estrategia y los mecanismos que permitan asegurar el cabal cumplimiento de las metas de las actividades de recolección o acopio, de acuerdo con la programación establecida. Para ello es fundamental definir el seguimiento y el control a cada una de las actividades determinadas y del desempeño de cada uno de los roles.

En caso de operativos de campo los instrumentos y las herramientas corresponden a:

- Recorrido de los recolectores que permite identificar la cobertura.
- Avances diarios de cobertura.
- Criterios de observación de entrevistas.
- Reentrevistas para corroborar los datos obtenidos en algunas unidades de observación.
- Rendimientos diarios por rol.
- Visualización de indicadores operativos.
- Tablero de control, entre otros.

Para el caso de las operaciones en que se realiza el acopio de los datos se requiere establecer algún mecanismo para el seguimiento de las actividades previstas en el diseño.

2.3.9. Diseño de sistemas de captura de datos

En desarrollo de las operaciones se incorpora el uso de tecnologías en los procesos para hacerlas más oportunas y de mayor eficiencia en el manejo de datos. Cualquiera que sea el instrumento desarrollado, debe tener estas características generales: fácil manejo; que presente ayudas o aclaraciones de ser

necesario; que permita la transmisión ágil y fácil de los datos; y que garantice la seguridad e integridad de la información. Asimismo, al conformar bases de datos es necesario asegurarse que se dispone de la tecnología para operar grandes volúmenes de datos.

Esta sección documenta los sistemas de captura que se establezcan y desarrollen, teniendo en cuenta el método de recolección o acopio definido, los desarrollos tecnológicos a emplear, los tiempos y recursos financieros disponibles. La documentación describe las características de los programas de captura para asegurar el registro de datos coherentes, respetando los valores válidos de respuesta, lógicos o aceptables, y reglas lógicas de consistencia de las variables. Dependiendo del método de recolección o acopio definido, se determina la forma en que los datos suministrados por las fuentes serán almacenados.

Los programas o aplicaciones varían de acuerdo con los objetivos, el cronograma y los recursos disponibles para la operación estadística. Pueden ser aplicaciones básicas, como hojas de cálculo, o herramientas sofisticadas que llevan a capturas en línea (digitación, dispositivos móviles de captura, cuestionario electrónico, reconocimiento automático de datos y caracteres, reconocimiento de imágenes y caracteres ópticos, lecturas de marcos ópticos, entre otros).

Si el proceso de recolección de la operación estadística se desarrolla con cuestionarios impresos (en papel), se debe incorporar y documentar el desarrollo de mecanismos de crítica estadística y de un instrumento de grabación en medios electrónicos.

2.3.10. Transmisión de datos

Documenta el procedimiento para organizar los archivos de datos recolectados y revisados en campo, como enviarlos a la base de datos central donde se almacenarán y el mecanismo de verificación para corroborar que la totalidad de registros enviados se ha almacenado. De esta manera se conforma la base de datos originales recolectados.

2.4. DISEÑO DEL PROCESAMIENTO

Describe el diseño de las herramientas tecnológicas, de software y hardware, que va a implementar para el procesamiento de los datos (software y hardware), los programas requeridos para la grabación, la consolidación y el almacenamiento. Así como, los protocolos y las herramientas definidas para garantizar la seguridad de la información.

Para el caso de estadísticas derivadas u operaciones a partir de registros administrativos, adicionalmente se contempla la depuración y tanto ajustes, agregaciones como desagregaciones de los datos provenientes de las diferentes fuentes.

2.4.1. Consolidación de archivos de datos

Documenta la forma como se van a organizar los archivos de datos transmitidos (por tipo, nivel de información, división o cualquier otro criterio) con el fin de conformar un archivo homogéneo para el procesamiento.

2.4.2. Codificación

Describe el procedimiento mediante el cual se establecen las correspondencias entre las respuestas de variables abiertas y las categorías predefinidas para las variables establecidas en el cuestionario, para ello se emplean las clasificaciones y nomenclaturas.

Para las variables que han sido previamente codificadas en el diseño, se incorporan en los aplicativos informáticos y durante la recolección se seleccionan las categorías (alternativas) correspondientes, con lo que en realidad se capturan los códigos. Igualmente, se especifica el procedimiento de codificación de variables cuando la recolección se lleva a cabo empleando cuestionarios impresos y cuando se generan nuevas variables.

2.4.3. Diccionario de datos

Documenta las características de las variables que son (o serán) utilizadas en la operación estadística, en el diccionario se definen todos los datos que serán incluidos en el procesamiento (código de la variable, tipo, extensión, valores válidos, especificación de obligatoriedad o no de la presencia del valor de la variable). Este diccionario facilitará tareas como la configuración de la base de datos, las reglas de edición (validación y consistencia) e imputación, y la definición de las especificaciones de estimación (para operaciones estadísticas por muestreo).

2.4.4. 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.

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

Describe los programas o aplicativos informáticos desarrollados para la edición e imputación automatizada, que permitirán generar una base de datos depurada de la operación estadística, con base en las reglas o especificaciones que elaboran los investigadores (temáticos y metodólogos), acompañada de los respectivos manuales del sistema y del usuario, para asegurar el correcto uso y manejo de las herramientas informáticas.

Además, se precisa la revisión y el tratamiento de los registros vacíos o aquellas preguntas que no tuvieron respuesta, es decir, la no respuesta (total o parcial). La no respuesta total ocurre cuando no se ha podido medir ninguna de las variables de interés. Por su parte, la no respuesta parcial se da cuando sólo se ha podido medir un subgrupo de variables de interés.

Posteriormente, se describe los programas o aplicativos informáticos desarrollados para el procesamiento estadístico de datos; es decir, las herramientas que permiten contrastar, estimar o calcular los parámetros de interés (según sea el tipo de operación estadística), organizar los resultados en los cuadros de salida predeterminados e indicar el error estándar de cada estimación (para el caso de operaciones estadísticas por muestreo) y la generación de reportes.

2.4.6. Diseño para la generación de cuadros de resultados

Describe los programas o los aplicativos informáticos desarrollados para el llenado de los cuadros de resultados, 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.

Documente las nuevas variables o unidades que se puedan requerir para el cálculo de indicadores (que se generan con las demás variables existentes en el archivo o base de datos) y las especificaciones para el llenado de los cuadros de salida (o de resultados), teniendo en cuenta el diseño de los cuadros, las

ponderaciones (factores de expansión y/o factores de ajuste), con lo cual se desarrollarán los programas informáticos apropiados.

2.5. DISEÑO DEL ANÁLISIS

Define y documenta los métodos, las técnicas y los procedimientos que permitan verificar la coherencia y calidad de la información estadística, generar los cuadros de resultados y determinar los métodos de análisis a aplicar para interpretar adecuadamente dichos resultados, teniendo en cuenta el tipo de operación estadística. Para ello, igualmente es necesario establecer los programas estadísticos que se emplearán para el efecto o el desarrollo de los aplicativos propios que sean necesarios.

2.5.1. Métodos de análisis de resultados

Describa los métodos establecidos para el análisis de los resultados de la operación estadística, entre los cuales se pueden considerar:

- Análisis de coherencia

Documente los métodos, técnicas o procedimientos definidos para identificar tendencias básicas, presencia de valores válidos, atípicos, la consistencia entre variables, aplicando técnicas estadísticas para llevar a cabo análisis de distribución, posibles correlaciones o formulación de hipótesis, análisis multivariados, entre otros. Estos análisis utilizan técnicas estadísticas que permiten tener una visión más amplia y robusta del comportamiento del fenómeno estudiado.

Documente los procedimientos que se aplicarán para comparar los resultados históricos, con otras fuentes de información similares y con los hechos sucedidos en el universo de estudio sobre la temática estudiada que den explicación de la información estadística obtenida.

Con el propósito de contrastar la coherencia y la consistencia de los resultados y realimentar los análisis realizados, se emplea los estándares estadísticos y se documentan los procedimientos para la comparabilidad con otras operaciones estadísticas similares que se produzcan en el SEN, en otros países, regiones del mundo o de organismos estadísticos internacionales.

- **Análisis estadístico**

El análisis implica organizar, procesar, validar, resumir e interpretar un conjunto de datos, con el propósito de producir conocimiento, tomar decisiones y evaluar programas o políticas implementadas.

Para las entidades que incorporen el análisis estadístico, en esta fase del proceso estadístico se define y documenta los métodos o los procedimientos que permiten organizar los datos recolectados/acopiados con el fin de explicar e interpretar las posibles relaciones existentes entre las variables estudiadas, dependiendo de la cantidad de variables que se analizan y se aplican técnicas de análisis que pueden ser univariadas, bivariadas o multivariadas.

- **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**

El análisis de datos bivariado es una forma evolucionada de análisis estadístico en el cual se cuantifica a nivel descriptivo e inferencial el nivel de covarianza entre dos variables y de esta forma se da cuenta de la relación entre dos variables. La cuantificación de la covarianza consiste en la construcción de coeficientes que permitan integrar en un valor estimado, información con respecto a la varianza conjunta entre dos variables y tiene como objetivo fundamental definir la magnitud y el sentido de la relación entre las variables (Sulbarán, 2012, pág. 3).

- **Análisis multivariado**

La parte del análisis estadístico que identifica la relación existente entre un conjunto de variables dependientes, valiéndose de métodos por medio de los cuales es posible el análisis simultáneo y la caracterización de las relaciones entre variables correlacionadas.

El análisis multivariado es un instrumento de análisis cuando existen múltiples variables, en un conjunto de datos; además, en algunos estudios se requiere de la combinación de varias preguntas para poder interpretar, comprender o analizar un concepto o fenómeno, por lo que la aplicación de una técnica univariada podría distorsionar la idea que en conjunto es preciso combinar (por ejemplo, la clase social se puede representar mejor por un conjunto de preguntas que incluyan el ingreso, la educación y la ocupación). En tales situaciones no es conveniente realizar análisis estadísticos univariados, porque pueden ofrecerse conclusiones inadecuadas o incluso erróneas de los resultados, o no extraerse toda la información del conjunto de datos, dado que se pasa por alto la interdependencia que pueda existir entre las variables objeto de estudio.

Todo análisis posible debe ser definido y documentado desde el diseño de la operación estadística. La documentación permite implementar los métodos o procedimientos establecidos para el análisis.

2.5.2. Anonimización de microdatos

Con el fin de asegurar la confidencialidad de los datos suministrados por las fuentes, es necesario, describir los protocolos de seguridad de la información que contribuyen a salvaguardar la identidad de las fuentes, documentando los procedimientos para el acceso a las bases de datos, los acuerdos o cláusulas de confidencialidad que suscribe el personal encargado del tratamiento de las bases, los procedimientos de almacenamiento de las bases, entre otros aspectos. De igual manera es importante definir, diseñar, desarrollar, implementar y documentar las técnicas de anonimización pertinentes de acuerdo con el tipo de variables que contiene la base de datos, se debe tener en cuenta que el fin último de la información son los usuarios, por eso es necesario proteger la privacidad de las fuentes y preservar el aprovechamiento de los datos.

2.5.3. Verificación de la anonimización de microdatos

Como parte del diseño del análisis de los resultados de la operación estadística es necesario documentar el procedimiento para verificar la aplicación de las técnicas definidas con el fin de eliminar el riesgo de identificación de las fuentes, es decir a personas naturales o jurídicas que suministran los datos para generar la información estadística, por ejemplo que los resultados a nivel de celdas, en los cuadros de resultados, no permita identificar a que fuente corresponde.

2.5.4. Comités de expertos

Describe el desarrollo de conversaciones internas y externas frente a los resultados con expertos internos de la entidad, nacionales e internacionales, en cada uno de los sectores económico-sociales, al igual que con directivos de la entidad productora, con el fin de contextualizar, analizar, validar y/o ajustar los resultados, para efectuar su difusión.

De igual manera, se establecen y documentan los protocolos o los procedimientos para la conformación y la convocatoria de los comités sectoriales o de expertos.

2.6. DISEÑO DE LA DIFUSIÓN Y COMUNICACIÓN

Hace referencia a los mecanismos, canales y medios previstos para la difusión y el acceso a la información producida por la operación estadística. Incluye las estrategias para promover el uso y la comprensión de la información estadística para responder a las necesidades de los grupos de interés.

2.6.1. Diseño de los sistemas de salida

Describe las diferentes actividades definidas para ensamblar y dar formato a microdatos, datos agregados y metadatos y a los canales y medios a través de los cuales se entrega la información estadística. Precise el lugar de almacenamiento, los mecanismos para su mantenimiento y actualización.

2.6.2. Diseño de productos de comunicación y difusión

Relacione los entregables que permiten la visualización de los resultados producidos por la operación estadística y el medio por el cual se disponen para consulta de los grupos de interés tales como cuadros de salida y de resultados, anuarios, calendarios de difusión, boletines, comunicados de prensa, series históricas con sus notas explicativas y bases de datos autónomas, multimedia o dinámicas indicando brevemente la forma de acceso (por ejemplo, mediante el portal web). Igualmente, los productos que se definen para la entrega de dichos resultados, publicaciones impresas, comunicados de prensa y sitios web, sistemas de consulta, visores geoestadísticos, mapas temáticos, geoportales, conjuntos de microdatos de uso público y documentos descargables, entre otros.

2.6.3. Entrega de productos

Presente la forma definida para comunicar y difundir los productos generados por la operación estadística canales, virtuales o presenciales y las estrategias de comunicación y difusión. La entrega de productos puede realizarse mediante rueda de prensa, eventos, a través de correos electrónicos o publicación en portal web, entre otros. Garantice la oportunidad y el cumplimiento en la difusión de resultados de las operaciones estadísticas, de acuerdo con las fechas y los horarios establecidos en el calendario anual de publicaciones.

2.6.4. Estrategia de servicio

Describa el método definido para brindar orientación y soporte a las dudas e inquietudes de los usuarios y dar respuesta oportuna, ya sea por canales virtuales, presenciales o telefónicos. La revisión periódica de las consultas y las solicitudes proporciona un insumo para el Sistema de Gestión de la Calidad y permite conocer nuevas necesidades o el cambio en las necesidades identificadas anteriormente.

2.7. DISEÑO DE LA EVALUACIÓN DE LAS FASES DEL PROCESO

Este subproceso documenta los criterios metodológicos para evaluar el desarrollo de cada una de las fases y al final del proceso estadístico, y determinar en qué medida se logra el cumplimiento de los objetivos planteados en la operación estadística, en contraste con las necesidades de información de los usuarios y con los resultados obtenidos.

Describe los criterios de evaluación para cada una las fases y la final del proceso estadístico. Los criterios pueden incluir aspectos como, por ejemplo, el uso, la funcionalidad, la descripción documental, la claridad, la calidad de los insumos, facilidad de acceso, aplicabilidad, exhaustividad, entre otros aspectos y dependiendo el elemento de evaluación.

En desarrollo de las fases se puede optar por preparar listas de chequeo para verificar el cumplimiento de criterios, de objetivos o la exhaustividad de los aspectos a incluir, de la documentación a elaborar, entre otros.

La documentación identifica, además de los criterios, los puntos clave de evaluación o de control, el procedimiento, el responsable, quien ejecuta, quien evalúa, quien toma las decisiones y quien implementa las acciones, con base en los indicadores que sean formulados para tales fines.

Los hallazgos de la evaluación deben verse reflejados en un plan de mejora que establecerá las acciones a implementar en las siguientes iteraciones de la operación estadística.

2.8. DISEÑO DE LOS SISTEMAS DE PRODUCCION Y FLUJOS DE TRABAJO

Este subproceso documenta la arquitectura, los procesos y los subprocesos de la operación estadística, sus interrelaciones, los equipos de trabajo que intervienen, los productos intermedios y finales que se generan, y los riesgos asociados al desarrollo de las diferentes fases.

3. DOCUMENTACIÓN RELACIONADA

Consiste en presentar un listado de los documentos técnicos y metodológicos utilizados en la generación de las estadísticas, como pueden ser: manuales, instructivos, cuestionarios, metodología de cálculo de los indicadores, etc., código de los programas tanto de recolección como de procesamiento y su documentación, así como la descripción de sus contenidos y utilidad. Debe relacionarse el vínculo o el sitio donde pueden consultarse en caso necesario.

GLOSARIO

Corresponde a la relación de los conceptos estandarizados utilizados en una operación estadística con el fin de garantizar la comparabilidad de la de la misma.

BIBLIOGRAFÍA

Presenta la totalidad del material bibliográfico consultado durante el desarrollo de la operación estadística tipo registro administrativo, con la finalidad de documentar el desarrollo del trabajo. Se organiza por orden alfabético, según el apellido del autor y para los casos en los cuales no se cuenta con información sobre el autor o es anónima se organiza de acuerdo con el título del documento.

El DANE utiliza para la elaboración de documentos el Manual de Estilo del DANE. Para las demás entidades se recomienda emplear la norma APA (Asociación Americana de Psicología).

ANEXOS

Se muestra la información adicional a la presentada en el cuerpo del documento con el fin de soportar, sustentar o ampliar dicha información como, por ejemplo: tablas, gráficos, formatos, cuestionarios, cuadros, esquemas funcionales, etc.

Bibliografía

- Comisión Económica de las Naciones Unidas para Europa (UNECE). (2000) "Terminology on Statistical Metadata", Conferencia de estadísticos europeos sobre estudios y estándares estadísticos, No. 53, Génova. Recuperado de http://ec.europa.eu/eurostat/ramon/coded_files/UNECE_TERMINOLOGY_STAT_METADATA_2000_EN.pdf
- Comisión Económica de las Naciones Unidas para Europa (UNECE). (2016) Modelo genérico del proceso estadístico (GSBPM). V 1.0 en español. (v 5.0 en inglés 2013). Recuperado de <https://statswiki.unece.org/download/attachments/97356247/GSBPM%205.0%20-%20SPANISH%20LANGUAGE%20VERSION.docx?version=1&modificationDate=1483543725840&api=v2>.
- Artículo 155 de la Ley 1955 de 2019 y se adiciona el título 3 a la parte 2° del libro 2 del Decreto 1170 de 2015 Único del Sector Administrativo de Información Estadística"
- Decreto 1743 de 2016. 1 de noviembre. Por el cual se reglamenta el artículo 160 de la Ley 1753 de 2015.
- Departamento Administrativo Nacional de Estadística DANE. (2017) Lineamientos para el proceso estadístico en el Sistema Estadístico Nacional.
- Departamento Administrativo Nacional de Estadística DANE. (2019) Sistema de consulta de conceptos estandarizados.
- DANE. (2012b). Regulación estadística en el marco del sistema estadístico nacional. Bogotá: DANE. Recuperado de <https://www.dane.gov.co/files/sen/PlegableSEN.pdf>
- DANE. (2012c). Resolución 1503 de 2011. Bogotá: DANE.
- DANE. (2012d). Resolución 691 de 2011. Bogotá: DANE.

- DANE. (2012e). Sistema de Nomenclaturas y clasificaciones. Bogotá: DANE.
- DANE - Instituto Colombiano de Normas Técnicas y Certificación – ICONTEC. (2017) Norma técnica de la calidad del proceso estadístico.
- Instituto Nacional de Estadística y Geografía (INEGI). (2006). Elaboración de Síntesis metodológicas. Serie: Lineamientos para la generación de estadística básica. México: INEGI.
- INEGI. (2012). Proceso estándar para el aprovechamiento de registros administrativos. México: INEGI.
- INEGI. (2010a). Captación en registros administrativos. México: INEGI.
- INEGI. (2010b). Captación de encuestas por muestreo. Serie: Documentos para la generación de estadística básica. México: INEGI.
- Niño, V. (2011). Metodología de la investigación. Diseño y ejecución. Bogotá: Ediciones dela U. p. 79.
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (s.f.). Glosario de términos estadísticos. Recuperado el 26 enero de 2012 de <http://stats.oecd.org/glossary/detail.asp?ID=532>