

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

COORDINACIÓN DE REGULACIÓN ESTADÍSTICA

GUÍA PARA LA ELABORACIÓN DEL DOCUMENTO METODOLÓGICO DE OPERACIONES ESTADÍSTICAS

2023

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

B. Piedad Urdinola Contreras
Directora

Leonardo Trujillo Oyola
Subdirector

María Fernanda De La Ossa Archila
Secretaría General

Directores técnicos:

Javier Sebastián Ruiz Santacruz
Dirección de Censos y Demografía (DCD)

Álvaro Enrique Duque Soto
**Dirección de Difusión, Mercadeo
y Cultura Estadística (DICE)**

Sandra Liliana Moreno Mayorga
Dirección de Geoestadística (DIG)

Andrea Ramírez Pisco
**Dirección de Metodología y Producción
Estadística (DIMPE)**

Andrés Felipe Ortiz Rico
Dirección de Recolección y Acopio (DRA)

Julieth Alejandra Solano Villa
**Dirección de Regulación, Planeación,
Estandarización y Normalización (DIRPEN)**

Juan Pablo Cardoso Torres
**Dirección de Síntesis y Cuentas
Nacionales (DSCN)**

Iván Rolando Castillo Prieto
Coordinador de Regulación Estadística

Geovanny Hernández Gaitán
Claudia Milena Sánchez Cruz
Aura María Moreno Gamba
Coordinación de Regulación Estadística

Sonia Marcela Naranjo Morales
Corrección de estilo

© DANE, 2023

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

Contenido

Introducción	6
Estructura del documento metodológico	7
Estructura del documento metodológico	7
INTRODUCCIÓN	9
1. ANTECEDENTES.....	9
2. DISEÑO DE LA OPERACIÓN ESTADISTICA	10
2.1. DISEÑO TEMÁTICO.....	10
2.1.1. Necesidades de información	10
2.1.2. Formulación de objetivos	11
2.1.3. Alcance.....	12
2.1.4. Marco de referencia	12
2.1.5. Definición de variables y construcción de indicadores estadísticos	13
2.1.6. Resultados estadísticos	14
2.1.7. Estándares estadísticos utilizados	14
2.1.8. Diseño del cuestionario	16
2.1.9. Normas, especificaciones o reglas de edición e imputación de datos.....	16
2.2. DISEÑO ESTADÍSTICO	17
2.2.1. Universo de estudio.....	17
2.2.2. Población objetivo	18
2.2.3. Cobertura geográfica.....	18
2.2.4. Desagregación geográfica.....	18
2.2.5. Desagregación temática	18
2.2.6. Fuentes de datos	18
2.2.7. Unidades estadísticas	19
2.2.8. Período de referencia	20
2.2.9. Período de recolección/acopio y frecuencia	20
2.2.10. Marco estadístico (censal o muestral, aplica para operaciones estadísticas por censo o por muestreo, en los demás casos omite este apartado).....	20

2.2.11.	Diseño muestral (aplica para operaciones estadísticas por muestreo, en los demás casos omite este apartado).....	20
2.2.12.	Ajustes de cobertura (o ajuste de cobertura por no respuesta, aplica para operaciones estadísticas por censo o por muestreo, en los demás casos omite este apartado) 23	
2.2.13.	Especificaciones de ponderadores (aplica para operaciones estadísticas por muestreo probabilístico, en los demás casos omite este apartado)	23
2.3.	DISEÑO DE LA RECOLECCIÓN/ACOPIO	23
2.3.1.	Métodos y estrategias de recolección o acopio de datos	24
2.3.2.	Estructura organizacional del operativo y definición del equipo requerido.....	25
2.3.3.	Esquema de entrenamiento del personal	25
2.3.4.	Conformación del equipo	26
2.3.5.	Proceso de sensibilización y acuerdos de intercambio	26
2.3.6.	Elaboración de manuales	27
2.3.7.	Diseño de la estrategia de comunicación y plan de contingencias.....	27
2.3.8.	Diseño de la estrategia de seguimiento y control	27
2.3.9.	Diseño de sistemas para la obtención de datos.....	28
2.3.10.	Transmisión de datos.....	28
2.4.	DISEÑO DEL PROCESAMIENTO	29
2.4.1.	Consolidación de archivos de datos	29
2.4.2.	Codificación	29
2.4.3.	Diccionario de datos	30
2.4.4.	Revisión y validación.....	30
2.4.5.	Diseño de instrumentos de edición (validación y consistencia) e imputación de datos 30	
2.4.6.	Diseño para la generación de cuadros de resultados	31
2.5.	DISEÑO DEL ANÁLISIS	31
2.5.1.	Métodos de análisis de resultados	31
2.5.2.	Anonimización de microdatos	33
2.5.3.	Verificación de la anonimización de microdatos.....	33
2.5.4.	Comités de expertos.....	33
2.6.	DISEÑO DE LA DIFUSIÓN Y COMUNICACIÓN	34

2.6.1. Diseño de los sistemas de salida	34
2.6.2. Diseño de productos de difusión y comunicación	34
2.6.3. Entrega de productos	34
2.6.4. Estrategia de servicio.....	35
2.7. DISEÑO DE LA EVALUACIÓN DE LAS FASES DEL PROCESO	35
2.8. DISEÑO DE LOS SISTEMAS DE PRODUCCION Y FLUJOS DE TRABAJO	35
3. DOCUMENTACIÓN RELACIONADA	36
GLOSARIO	36
BIBLIOGRAFÍA	36
ANEXOS.....	36
Bibliografía.....	38

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 1000:2020 (DANE, 2020a) y los Lineamientos para el Proceso Estadístico para el SEN (DANE, 2020b), 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 y brindando cumplimiento a los diferentes atributos de la calidad estadística. El DANE como ente rector actualiza regularmente esta guía¹ según recomendaciones internacionales, por lo que esta se constituye en el metadato principal de la operación estadística, dado que precisa los criterios técnicos estadísticos para su desarrollo, 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 del proceso estadístico, luego de llevar a cabo las pruebas² 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 las pautas generales relativas a la estructura, la presentación y el contenido del documento metodológico, empleando criterios estandarizados y armonizados³; y referentes internacionales.

Esta guía se estructura en tres partes: en la primera se presentan las generalidades de las operaciones estadísticas; en la segunda se hace una descripción de los elementos que hacen parte del documento metodológico, y en la tercera se relaciona la bibliografía de referencia útil sobre los tópicos tratados. Cabe resaltar que el documento metodológico es un estándar que permite la conservación de la memoria de las operaciones estadísticas.

¹ Por esta razón se encuentra publicada tanto la versión 2020 como su actualización 2023.

² Para el desarrollo de las diferentes pruebas se recomienda consultar Formato para el Plan de Pruebas de las operaciones estadísticas, disponible en <https://www.dane.gov.co/files/sen/lineamientos/Formato-Plan-de-Pruebas.xlsx>

³ 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.

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, 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, define el diseño de manera organizada teniendo en cuenta los subprocesos del proceso estadístico, con el fin de brindar todos los elementos técnicos y metodológicos que enmarcan el desarrollo de la operación en las demás fases. Si algún literal no es llevado a cabo por las particularidades propias de la operación estadística, este debe diligenciarse con un No aplica.

Estructura del documento metodológico

INTRODUCCIÓN

1. ANTECEDENTES

2. DISEÑO DE LA OPERACIÓN ESTADISTICA

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. Resultados estadísticos
- 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 y frecuencia
- 2.2.10. Diseño muestral
- 2.2.11. Especificaciones de ponderadores

2.3. DISEÑO DE LA RECOLECCIÓN/ACOPIO

- 2.3.1. Métodos y estrategias de recolección o acopio de datos
- 2.3.2. Estructura organizacional del operativo y definición del equipo requerido
- 2.3.3. Esquema de entrenamiento del personal
- 2.3.4. Conformación del equipo
- 2.3.5. Proceso de sensibilización y acuerdos de intercambio
- 2.3.6. Elaboración de manuales
- 2.3.7. Diseño de la estrategia de comunicación y plan de contingencias
- 2.3.8. Diseño de la estrategia de seguimiento y control
- 2.3.9. Diseño de sistemas para la obtención de datos
- 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 de resultados
- 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 COMUNICACIÓN

- 2.6.1. Diseño de los sistemas de salida

2.6.2. Diseño de productos de difusión y comunicación

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 PRODUCCIÓN Y FLUJOS DE TRABAJO

3. DOCUMENTACIÓN RELACIONADA

GLOSARIO

BIBLIOGRAFÍA

ANEXOS

INTRODUCCIÓN

Es un resumen claro y sencillo que describe el contexto temático de la operación estadística. En este apartado 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 la fuente de datos y se describen la relevancia⁴ de la operación estadística y su relación con la política pública.
- 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, una vez estos hayan sido aprobados.
- La estructura del documento; se aconseja realizar una descripción mínima de todos los capítulos o las secciones que componen la metodología.

1. ANTECEDENTES

Presente 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, 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

⁴ NTCPE 1000:2020, página 27. NOTA 1: la relevancia puede evidenciarse a través de leyes, decretos, resoluciones, el Plan Nacional de Desarrollo, documentos del Consejo Nacional de Política Económica y Social (CONPES), Planes de Desarrollo Territoriales, acuerdos internacionales, Plan Estadístico Nacional, documentos de política pública, entre otros.

NOTA 2: la relevancia puede evidenciarse en la importancia de la información estadística para el país.

análisis de los resultados.

2. DISEÑO DE LA OPERACIÓN ESTADISTICA

Determine los métodos y los procedimientos con los que se construyen los instrumentos, procesos y actividades, se recolectan/acopian, procesan y analizan los datos, se difunde la información y evalúan las fases del proceso estadístico para cumplir los objetivos y satisfacer las necesidades de información.

Los diseños se validan y ajustan de acuerdo con los resultados de las pruebas que se desarrollan en la fase 3 (construcción), de la misma manera los diseños se ajustan cuando se requiere una actualización que no modifica los objetivos o el alcance de la operación o cuando surge la necesidad de un rediseño⁵ de la operación estadística a partir de nuevas necesidades de información, el cambio de las existentes o los resultados obtenidos de la evaluación final que se realiza en la fase 8 (evaluación).

2.1. DISEÑO TEMÁTICO

Documente las necesidades de información que se han consultado, analizado y confirmado, las cuales han permitido formular el objetivo general y los objetivos específicos 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 las variables, los 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

Documente 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 y mecanismos utilizados para la priorización de las necesidades y resaltar los principales resultados y conclusiones obtenidas en dicha

⁵ Actividad que incluye una o más modificaciones en las fases de Detección y Análisis de Necesidades y Diseño que implican cambios en el alcance y los objetivos, por ende, en las demás fases del proceso estadístico.

aplicación.

2.1.2. Formulación de objetivos

Delimite de manera clara y concreta las necesidades de información confirmadas de la operación estadística. Los objetivos no requieren de preámbulos o explicaciones, se expresan de manera concreta con el fin de evitar distorsiones e interpretaciones incorrectas.

Este apartado describa el objetivo general y los objetivos específicos de la operación estadística. Estos deben ser congruentes con las necesidades de información que se pretenden satisfacer.

- a) **Objetivo general:** describa el logro o el propósito por alcanzar de la operación estadística. También puede considerarse como la representación global del resultado esperado y que debe estar acorde con la pregunta central de investigación planteada, en términos del tipo de análisis, objeto, la población de estudio y delimitación geográfica.
- b) **Objetivos específicos:** describa los propósitos que se pretenden lograr de manera particular dando respuesta o solución al objetivo general. Estos exponen de manera clara y precisa el fenómeno de estudio, las unidades sobre las que se van a hacer las mediciones y el alcance de la operación estadística.

Los objetivos deben cumplir con las siguientes características:

- Ser precisos, medibles, evaluables, viables y seguir un diseño metodológico.
- Ser claros y estar muy bien definidos de manera que facilite la comprensión al equipo de la operación estadística y a los usuarios
- 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 contempla la siguiente estructura:

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

Que responde a las preguntas:

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

Por 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”.

El cual, de acuerdo con las preguntas, se desglosaría así:

- *qué*: proporcionar información básica;
- *sobre qué*: fuerza de trabajo;
- *para qué*: determinar el tamaño y la estructura de la fuerza de trabajo;
- *dónde*: Colombia.

2.1.3. Alcance

Describir 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 directamente relacionada con los objetivos 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 partir de la revisión, documentación y análisis de estos referentes se debe especificar cuáles son considerados dentro del alcance de la operación estadística.

- a) Marco teórico:** corresponde a la literatura que sustenta la operación estadística, desde la perspectiva temática y/o teórica, así como el resumen de otros enfoques analizados.
- 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:** describir la normatividad (La Constitución y la Ley, Decretos, resoluciones y otros actos administrativos, Jurisprudencia, Convenios y acuerdos) 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:** describir los principales trabajos y las recomendaciones realizadas por organizaciones internacionales, especificando las razones por las cuales fueron seleccionados y sus principales aportes. Tales organizaciones pueden ser, por ejemplo: 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 e influyeron en el diseño, especificando las razones por las cuales fueron seleccionados y sus principales aportes.

2.1.5. Definición de variables y construcción de indicadores estadísticos

Enuncie y describa las variables⁶ a medir con la operación estadística y a emplear en el diseño de los indicadores⁷ 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.

Además, relacione y describa los indicadores que se calculan o estiman en la operación estadística, sus objetivos, fórmulas, la metodología de cálculo y la información de la manera

⁶ 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).

⁷ Expresión cuantitativa observable y verificable, que permite describir características, comportamiento o fenómenos de la realidad, a través de la medición de una variable o relación entre variables. Tomado del Sistema de Consulta de conceptos estandarizados del DANE.

cómo se estiman, analizan e interpretan.

Relacione y describa con precisión las variables necesarias para el cálculo de los indicadores, estas inciden 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, valores absolutos, totales o indicadores especiales⁸.

2.1.6. Resultados estadísticos

Describa de manera general los resultados esperados de la operación estadística, el tipo de reportes o de informes, cuadros de salida, gráficos, mapas temáticos, sistemas geoespaciales, atlas estadísticos, infografías, anuarios 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 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 información. 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.

2.1.7. Estándares estadísticos utilizados

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

⁸ También denominados *indicadores compuestos o sintéticos* corresponden a la agregación de indicadores parciales de un determinado componente en un único indicador.

de la información estadística.

Los estándares pueden ser:

- *Clasificaciones estadísticas*⁹.
- *Conceptos estandarizados*¹⁰.
- *Nomenclaturas*¹¹.
- *Variables*¹².
- *Unidades estadísticas*¹³.
- *Data Documentation Initiative DDI y Dublin Core*¹⁴.
- *Statistical Data and Metadata Exchange - SDMX*¹⁵.

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

⁹ 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.

¹⁰ 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.

¹¹ Son conjuntos de códigos numéricos o alfanuméricos que sirven para identificar diferentes elementos. Ejemplo: países, unidades territoriales, operaciones estadísticas, entidades estatales.

¹² 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.

¹³ Corresponde a las unidades de observación, las unidades de análisis y, cuando aplique, las unidades de muestreo.

¹⁴ 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.).

¹⁵ 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.

2.1.8. Diseño del cuestionario¹⁶

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 obtener la información estadística que se presentará en los cuadros de salida.

El cuestionario se estructura generalmente en bloques temáticos o capítulos conformados por grupos de preguntas que tienen una secuencia lógica, su diseño debe facilitar el diligenciamiento y minimizar el desgaste de la fuente.

Es necesario conocer y documentar los objetivos de cada uno de los módulos o los capítulos que lo componen, así como de las preguntas, lo cual se convierte en un recurso fundamental al momento de evaluar la modificación o la exclusión de una pregunta del cuestionario, 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.

Otro aspecto esencial es la definición de pruebas al cuestionario. Las pruebas buscan que el cuestionario sea coherente con el fenómeno de estudio, 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

Documente 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

¹⁶ Para el diseño del cuestionario se recomienda consultar la Guía para la elaboración, la revisión y el rediseño de cuestionarios para la recolección, disponible en <https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/normas-y-estandares/lineamientos>.

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 este apartado se describe, de forma general, el método, las reglas o las especificaciones seleccionadas o establecidas para la edición y la imputación de los datos, y en un documento específico se registra con todo detalle el conjunto de métodos, reglas o especificaciones. Así mismo, se debe relacionar la ubicación de los manuales de validación y consistencia y los manuales de imputación (cuando apliquen), con el fin de facilitar a los interesados su consulta.

2.2. DISEÑO ESTADÍSTICO

Consiste en la definición, la construcción y 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. Para operaciones por muestreo probabilístico 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 definir 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 individuos o elementos en los cuales pueden representarse determinadas características a ser estudiadas, 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 elementos de los que se desea obtener los datos y sobre los que se presentan conclusiones o resultados; está compuesta por unidades que comparten alguna característica, tienen una localización geográfica y un tiempo como periodo de referencia.

Puede ser igual al universo o un subconjunto de este. 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 que 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, 2022).

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, 2010).

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 la fuente de datos, la cual puede ser:

- **Fuente primaria:** es aquella que obtiene los datos directamente de las unidades estadísticas empleando instrumentos propios. Son utilizadas principalmente en operaciones estadísticas por censo y muestreo. En la documentación para este criterio

se debe proporcionar el número de fuentes, cuando sea posible.

- **Fuente secundaria:** corresponde al conjunto de datos, sobre hechos o fenómenos, que se obtienen a partir de la recolección realizada por otros¹⁷. Algunos ejemplos de fuente secundaria pueden ser las bases de datos de otras operaciones estadísticas o de registros administrativos (los cuales en algunos casos no responden a un objetivo estadístico en particular). Cuando este sea el caso debe registrarse la entidad o área productora, una descripción breve del conjunto o los conjuntos de datos obtenidos, año de ejecución y relacionarlas, así como las transformaciones que se realizan para el aprovechamiento estadístico.

2.2.7. Unidades estadísticas

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

- Unidad de observación:** elemento o conjunto de elementos sobre los que se hace medición de las diferentes variables en una 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.
- Unidad de análisis:** elemento de estudio sobre la que se presentan los resultados o las conclusiones de la operación estadística. 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:

- Unidad de muestreo:** describa 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.

¹⁷ Entiéndase por “otros” a otras entidades o procesos ajenos al desarrollo de la operación estadística.

2.2.8. Período de referencia

Documente el período de tiempo específico (día, semana, mes, año) o punto en el tiempo al que corresponde la observación realizada, o la variable de interés. 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 y frecuencia

Describa el intervalo de tiempo o la fecha en la cual se realiza el levantamiento o acopio de los datos de la 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, adicionalmente indica la frecuencia con que se lleva a cabo la recolección o el acopio de los datos (anual, semestral, trimestral, mensual).

2.2.10. Marco estadístico (censal o muestral, aplica para operaciones estadísticas por censo o por muestreo, en los demás casos omita 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 censal según las especificaciones de la operación estadística (censo o muestreo). 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 (aplica para operaciones estadísticas por muestreo, en los demás casos omita este apartado)

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 lo

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 afijación¹⁸, 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

¹⁸ 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.

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

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.

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, aplica para operaciones estadísticas por censo o por muestreo, en los demás casos omite este apartado)

Describa 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; para el caso de 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 (aplica para operaciones estadísticas por muestreo probabilístico, en los demás casos omite este apartado)

Luego de establecer las formulaciones matemáticas para el cálculo de los factores de expansión y las varianzas, 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/ACOPIO¹⁹

En esta sección describa la manera como se obtienen los datos necesarios en una operación estadística.

En la recolección se establecen aspectos como la estructura operativa y tecnológica, los roles y perfiles del personal idóneo para la recolección de la información, las cargas de trabajo, las rutas de recolección, la digitación, la transmisión o la digitalización de los datos, la validación de los datos, los transportes necesarios para los desplazamientos, la logística necesaria para la recolección, la cartografía actualizada, la georreferenciación de las unidades de observación, el entrenamiento al personal, la sensibilización de las fuentes, y el manejo de novedades operativas.

En el acopio se determina y gestiona el intercambio de los archivos de datos que contengan los microdatos.

Para las operaciones estadísticas que utilizan fuentes alternativas relacionadas con instrumentos tecnológicos de medición (por ejemplo: sensores remotos, pluviómetros,

¹⁹ 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.

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, cuestionario electrónico, asistido con Dispositivo Móvil de Captura - DMC, y con cuestionario impreso), entrevista por invitación pública, 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.

Adicionalmente, se pueden prever y documentar otras estrategias como establecer puntos fijos, el cual consiste en establecer una ubicación específica en donde las unidades de observación, por ejemplo, personas, se pueden acercar y brindar sus datos. Esta estrategia se puede emplear en la entrada de condominios combinada con la visita a las viviendas, o 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, siempre y cuando no se requiera georreferenciar las viviendas o tomar información por observación de viviendas u otros.

Para operaciones que usan fuentes de datos secundarias se describe el procedimiento y la plataforma tecnológica para el acopio de los datos (estructurados²⁰ o no estructurados²¹), 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.

²⁰ Son datos que tienen un modelo de datos y formato predefinido y que se ajustan a una forma de tablas de registros o filas con campos de significados fijos y relaciones o enlaces entre las tablas. 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

²¹ Son los datos que no tienen un modelo de datos predefinido, no están organizados de manera predefinida o su estructura no se ajusta perfectamente a una tabla de datos relacional

2.3.2. Estructura organizacional del operativo y definición del equipo requerido

Documente 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, 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 para la recolección, como es el caso de los 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 entrenamiento de personal; la planeación de las entrevistas; la organización de la supervisión; el monitoreo, y el control.

Documente 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

Describa 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. El entrenamiento debe contener aspectos relacionados con la confidencialidad estadística, el marco teórico y conceptual, el manejo de herramientas y equipos, la contextualización sobre la entidad productora de información, las actividades para el seguimiento y la supervisión de la recolección, los procesos de sensibilización aplicados a la fuente y los criterios para la evaluación de desempeño y calidad del trabajo.

2.3.4. Conformación del equipo

Documente de manera general el procedimiento para la conformación del equipo, teniendo en cuenta los perfiles requeridos para los diferentes roles, incluyendo los instructores cuando sea necesario. Puede considerarse, por ejemplo, que luego de la invitación pública 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, describa 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

Documente el conjunto de acciones encaminadas a generar conciencia en las fuentes, acerca de la importancia de la operación estadística. Describe los mecanismos, los textos, los medios y los instrumentos que serán utilizados en la sensibilización²², así como los actores que participarán en la comunicación con la unidad de información o de reporte. Incluye un plan para difundir 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.

Las operaciones estadísticas que utilizan fuentes de datos secundarias, deben describir las gestiones adelantadas para los acuerdos de intercambio con los proveedores de datos, como los servicios web, la extensión de intercambio, el canal de transmisión, las fechas periódicas de transmisión, el acuerdo a nivel de servicio que involucre aspectos de confidencialidad, uso de datos, responsabilidad de las partes, disponibilidad del servicio, restricciones, soporte, vigencia, etc., la validación de los datos. Los acuerdos deben documentarse atendiendo la guía para realizar intercambios de información SEN, con el fin

²² Proceso comunicativo dirigido a las fuentes y a los grupos de interés para facilitar la entrega de los datos requeridos por la operación estadística. (DANE, 2022)

de 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 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.

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 describe 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 mecanismos para el seguimiento de las actividades previstas en el diseño.

2.3.9. Diseño de sistemas para la obtención 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.

En esta sección, documente 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 captura de los datos en medios electrónicos.

2.3.10. Transmisión de datos

Documente el procedimiento para organizar los archivos de datos obtenidos, como transmitirlos a la base de datos central donde se almacenarán y el mecanismo de verificación para corroborar que la totalidad de registros transmitidos se almacenaron. De

esta manera se conforma la base de datos original.

2.4. DISEÑO DEL PROCESAMIENTO

Describa el diseño de las herramientas tecnológicas, de software y hardware, que va a implementar para el procesamiento de los datos, los programas requeridos para la captura o acopio, la consolidación y el almacenamiento. Así como, los protocolos y las herramientas definidas para garantizar la seguridad de la información.

Para el caso de las estadísticas derivadas u operaciones a partir de registros administrativos, adicionalmente se contemplan todas las transformaciones realizadas a los datos obtenidos para su aprovechamiento estadístico, estas consideran la depuración, imputación, así como ajustes, agregaciones o desagregaciones de los datos provenientes de las diferentes fuentes.

También describe la integración de los archivos de datos, la verificación de la completitud y calidad de los datos, las especificaciones de rutina para la codificación, la edición, la imputación, la estimación, la validación y consistencia de la información y la generación de resultados.

2.4.1. Consolidación de archivos de datos

Documente 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

Describa 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. Igualmente, se especifica el procedimiento de codificación de variables cuando la recolección se lleva a cabo empleando cuestionarios impresos, cuando se generan nuevas variables o se obtienen bases de datos sin codificar.

2.4.3. Diccionario de datos²³

Describe de manera general las características de las variables que son (o serán) utilizadas en la operación estadística y que fueron incluidas en el diccionario de datos. El diccionario define 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

Documente el procedimiento establecido para revisar y validar²⁴ 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

Describir los programas o aplicativos informáticos que serán utilizados para el desarrollo de 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, y 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, organizar los resultados en los cuadros de salida predeterminados e indicar el error estándar de cada estimación (para el caso de

²³ Para la elaboración del diccionario de datos se recomienda consultar la Guía para documentar el diccionario de datos, disponible en https://www.sen.gov.co/files/sen/lineamientos/Gu%C3%ADa_Diccionario_de_Datos.xlsx

²⁴ En este apartado la validación corresponde a la verificación de los datos directamente con la fuente (siempre que sea posible), cuando se presenten omisiones o inconsistencias, con el fin de evitar sesgos en los datos.

operaciones estadísticas por muestreo) y la generación de reportes.

2.4.6. Diseño para la generación de cuadros de resultados

Describa los programas o los aplicativos informáticos desarrollados para la conformación 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 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, el comportamiento del fenómeno objeto de estudio y los resultados obtenidos, este último con los resultados esperados. De igual manera, se verifica la consistencia de las variables, los datos atípicos, la georreferenciación de las fuentes y la información estadística obtenida con relación a las series y los indicadores históricos. 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.

- **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 definen y documentan 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, medidas de dispersión de la variable y análisis de estacionalidad. 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 ellas. 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 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

Es importante definir, diseñar, desarrollar, implementar y documentar las técnicas de anonimización utilizadas por la operación estadística 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. Es necesario, describir los protocolos de seguridad de la información que salvaguardan 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.

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 (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 sesiones de trabajo 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 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

Documente la estrategia de difusión que incluye las diferentes maneras para promover el uso y la comprensión de la información estadística que responde a las necesidades de los grupos de interés. Tenga en cuenta los cuadros de salida, los productos estadísticos, las actividades y el calendario para la difusión, la forma de presentación de la información estadística, las instrucciones para la lectura, interpretación y uso de la información, los canales y los medios de difusión, los metadatos de la operación estadística, y la asesoría a usuarios de la información estadística.

2.6.1. Diseño de los sistemas de salida

Describa las 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. Precisa el lugar de almacenamiento, los mecanismos para su mantenimiento y actualización.

2.6.2. Diseño de productos de difusión y comunicación

Relacione los entregables que permitan 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 los resultados, publicaciones impresas, 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. Se debe asegurar la oportunidad y el cumplimiento en la difusión de los 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

Documente los criterios metodológicos para evaluar el desarrollo de cada una de las fases 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.

Describa los criterios de evaluación para cada una de las fases y al 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 del 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.

En este apartado se identifica, 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 actividades de mejoramiento, estas deben ser incluidas en el informe final de evaluación, así como el reporte de su ejecución en la siguiente evaluación.

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

Documente la arquitectura, se identifican los procesos y los subprocesos correspondientes de la operación estadística, se establece las interrelaciones entre los diferentes procesos, se establecen los equipos de trabajo responsables que intervienen, se determinan los productos intermedios y finales requeridos y los riesgos asociados con sus controles para

evitar que se materialicen 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 utilizados en una operación estadística con el fin de garantizar la comparabilidad de esta.

Es importante que los conceptos que se utilicen en la operación estadística estén estandarizados o realicen el proceso de estandarización por el DANE para garantizar la comparabilidad, la integración y la interoperabilidad de la información estadística. En el glosario de términos se relaciona la definición y su fuente.

BIBLIOGRAFÍA

Relacione las publicaciones y el material bibliográfico empleado para desarrollar la operación estadística.

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. Si hay más de un documento por autor se organiza por fecha, desde la más reciente a la más antigua.

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.

ASPECTO RELEVANTE

Es importante consultar, para el desarrollo de actividades específicas del proceso estadístico, las guías y recomendaciones establecidas por el DANE en el siguiente enlace:
<https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/normas-y-estandares/lineamientos>

Bibliografía

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”

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.
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).
<https://statswiki.unece.org/download/attachments/97356247/GSBPM%205.0%20-%20SPANISH%20LANGUAGE%20VERSION.docx?version=1&modificationDate=1483543725840&api=v2>.

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). (2020a). Norma Técnica de Calidad del Proceso Estadístico NTCPE 1000:2020.

Departamento Administrativo Nacional de Estadística (DANE). (2020b). Lineamientos para el proceso estadístico en el Sistema Estadístico Nacional. Versión 2.0.

Departamento Administrativo Nacional de Estadística (DANE). (2022). Sistema de consulta de conceptos estandarizados.

Instituto Nacional de Estadística y Geografía (INEGI). (2006). Elaboración de Síntesis metodológicas. Serie: Lineamientos para la generación de estadística básica. México.

Instituto Nacional de Estadística y Geografía (INEGI). EGI. (2012). Proceso estándar para el aprovechamiento de registros administrativos. México.

Instituto Nacional de Estadística y Geografía (INEGI). (2010). Captación de encuestas por muestreo. Serie: Documentos para la generación de estadística básica. México.

Organización para la Cooperación y el Desarrollo Económico (OCDE). (s.f.). Glosario de términos estadísticos. <http://stats.oecd.org/glossary/detail.asp?ID=532>

Sulbaran, Dimas. 2012. Análisis Bivariado de Datos.
<http://caelum.ucv.ve/bitstream/10872/11400/1/An%C3%A1lisis%20bivariado%20de%20datos.pdf>