

# COLOMBIA - Investigación de Educación Formal - 2015

**Investigación de Educación Formal - DIMPE - DANE**

report\_generated\_on: December 28, 2021

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

# Información general

## Identificación

NÚMERO DE ID  
DANE-DIMPE-EDUC-2015

## Información general

### RESUMEN

La educación constituye un proceso continuo de transformación, participación y cambio social que incrementa la libertad y el bienestar de las personas, en la medida en que contribuye al desarrollo de sus potencialidades, a la ampliación de sus capacidades y a la acumulación de capital humano. De esta manera, la educación configura una de las variables estratégicas más importantes de la política social que interviene en el avance de otras dimensiones fundamentales del bienestar y por ende en el desarrollo y crecimiento económico de un país.

En el campo de la educación, producir información estadística estratégica consiste en generar datos útiles y dinámicos que representen el panorama del sistema educativo en todas sus dimensiones, pero sobre todo, que permitan evaluar la prestación del servicio educativo y de esta manera, orientar el diseño de políticas públicas efectivas que contribuyan a su mejoramiento.

La investigación de educación formal (EDUC) es una operación estadística de tipo censo, de periodicidad anual. Es efectuada en establecimientos de educación formal legalmente constituidos, del sector oficial y no oficial, que ofrecen los niveles de preescolar, básica primaria, básica secundaria y media, localizados en el área urbana, y rural dentro del territorio nacional, de acuerdo con las variables estructurales estudiadas: matrícula; docentes, y situación académica de los estudiantes al terminar el año lectivo anterior (aprobación, deserción, reprobación).

Su principal objetivo es generar y difundir información estadística básica sobre la educación formal en Colombia, en los niveles de preescolar, básica primaria, básica secundaria y educación media, base para la formulación y seguimiento de la política pública sectorial, y el ejercicio de planeación y administración del servicio público de educación según las categorías político-administrativas del país.

La información en el período de referencia 2015, se obtiene directamente del reporte de las fuentes primarias (sedes educativas), por auto diligenciamiento del formulario único censal C600. El instrumento se presenta en medio físico y en medio electrónico.

El universo y la unidad de observación de la operación estadística de educación formal está conformada por todas las sedes educativas (establecimientos) de educación formal, legalmente constituidas, del sector oficial y del sector no oficial, que ofrecen los niveles de preescolar, básica primaria, básica secundaria y media, localizados en el área urbana y en el área rural dentro del territorio nacional.

La cobertura es nacional, con desagregación geográfica departamental y municipal. También se registra la desagregación por entidad territorial certificada de educación. La periodicidad de recolección y difusión de la información es anual.

Los resultados y las cifras oficiales se difunde a través de los diferentes medios establecidos por DANE:

- i) Difusión a través de página web.
- ii) A través de bancos de datos.
- iii) En el Archivo Nacional de Datos - ANDA en micro datos anonimizados.
- iv) Mediante consultas y procesamiento por demanda a través de la sala especializada.

El presente documento contiene de manera detallada la metodología utilizada para desarrollar la investigación de educación formal durante el período de referencia 2015.

TIPO DE DATO

Censos (cen)

## UNIDAD DE ANÁLISIS UNIDAD DE OBSERVACIÓN

Sedes educativas de educación formal, legalmente constituidas, del sector oficial y del sector no oficial, que ofrecen los niveles de preescolar, básica primaria, básica secundaria y media, localizados en el área urbana y en el área rural dentro del territorio nacional.

## UNIDAD DE ANÁLISIS

Sedes educativas de educación formal, legalmente constituidas, del sector oficial y del sector no oficial, que ofrecen los niveles de preescolar, básica primaria, básica secundaria y media, localizados en el área urbana y en el área rural dentro del territorio nacional. Teniendo como fuente primaria las sedes-jornadas dentro de las unidades de análisis.

# Ámbito

## NOTAS DISEÑO DE LA OPERACIÓN ESTADÍSTICA

## NECESIDADES DE LA INFORMACIÓN

En este contexto el DANE desde 1978 viene ejecutando la operación estadística de educación formal, definida como un censo escolar que llega a todos los establecimientos legalmente constituidos y que ofrecen los niveles de enseñanza desde preescolar hasta media. Provee datos sobre matrícula, docentes y jornadas; y otras variables sobre la situación de la población al finalizar el año lectivo escolar a partir de las cuales se sustentan los indicadores de eficiencia interna base para el estudio de los fenómenos de promoción y deserción escolar entre otros.

## OBJETIVOS

### \* Objetivo general

Generar y difundir información estadística básica sobre la educación formal en Colombia, en los niveles de preescolar, básica primaria, básica secundaria y educación media, como base para la formulación y el seguimiento de la política pública sectorial, y el ejercicio de planeación y administración del servicio público de educación según las categorías político-administrativas del país.

### \* Objetivos específicos

- Generar y difundir información estadística básica sobre la educación formal en Colombia en los niveles de preescolar, básica primaria, básica secundaria y educación media, de acuerdo con las variables estructurales estudiadas (matrícula, docentes, jornadas) y de la situación académica de los estudiantes al terminar el año lectivo anterior (aprobación, deserción, reprobación).
- Consolidar la información sobre los establecimientos educativos, sedes educativas y sedes jornadas, para el sector oficial y no oficial según las categorías político-administrativas del país.
- Producir las estadísticas estructurales sectoriales y las variables para la construcción del indicador de eficiencia interna.

## ALCANCE

La EDUC está dirigida a los establecimientos de educación formal legalmente constituidos, del sector oficial y no oficial, que ofrecen los niveles de preescolar, básica primaria, básica secundaria y media, localizados en el área urbana, y rural dentro del territorio nacional, de acuerdo con las variables estructurales estudiadas: matrícula; docentes, y sobre la situación académica de los estudiantes al terminar el año lectivo anterior (aprobación, deserción, reprobación).

Se obtiene dos productos sustantivos derivados de esta operación, por una parte el acervo estadístico primario de la educación básica formal, en sus diferentes niveles de desagregación espacial y temática, y el marco estadístico de establecimientos educativos.

## MARCO DE REFERENCIA

### - Marco teórico

La consolidación de los sistemas estadísticos nacionales, se explica en gran medida por el desarrollo y la evolución de la regulación y la planificación orientada a la producción de información básica oficial estratégica, acervo éste necesario para el cumplimiento de los fines del Estado, siendo uno de ellos la educación, bien superior, público y meritório.

### - Marco conceptual

Los principales conceptos que deben ser considerados para abordar el entendimiento de la investigación de educación formal.

### - Marco legal

## DISEÑO DE INDICADORES

Los indicadores educativos comprenden información primaria seleccionada para conocer aspectos relevantes de las variables que identifican el sector educativo.

### - Matrícula nivel preescolar

### - Matrícula nivel básica primaria

### - Matrícula nivel secundaria

### - Matrícula nivel media

### - Tasa de alumnos/docente

### - Tasa de aprobación

### - Tasa de reprobación

### - Tasa de deserción

### - Tasa de transferidos

## PLAN DE RESULTADOS

### Diseño de cuadros de salida o de resultados

La EDUC produce cuadros de salida que presentan los resultados estadísticos de la investigación, con una periodicidad anual y una desagregación a nivel departamental y municipal. Por demanda se alistan cuadros de salida por entidad territorial certificada de educación. A continuación se presenta un listado de los cuadros de salida de la investigación:

### Módulo I: Caratulo única. Sede Educativa.

#### - Sedes educativas por Naturaleza Jurídica.

#### - Sedes Educativas por Modalidad de Prestación del Servicio Educativo.

### Módulo II: Organización institucional para la prestación del servicio educativo.

#### - Relación entre sedes educativas.

### Módulo III: Información sobre personal ocupado en la sede educativa

#### - Personal ocupado en la sede educativa.

#### - Docentes ocupados en la sede educativa, según estatuto docente y vinculación laboral.

- Docentes escalonados en el sector oficial y no oficial según estatuto docente.
- Último nivel educativo alcanzado por el docente, según nivel de enseñanza y modelo educativo, por sexo.
- Docentes ocupados por la sede educativa, por sexo según carácter educativo.

Módulo IV: Información de matrícula, niveles de enseñanza, CLEI, Modelos educativos en la sede educativa.

- Matrícula según Nivel de enseñanza y tipo de matrícula.
- Matrícula según zona y tipo.
- Matrícula según sector y tipo.

Módulo V: Información sobre matrícula y docentes por Jornada.

- Matrícula en Preescolar por rangos de edad.
- Matrícula en primaria por rangos de edad.
- Matrícula en secundaria por rangos de edad.
- Matrícula en media por rangos de edad.
- Matrícula en ciclos lectivos integrados por rangos de edad.
- Matrícula en preescolar por grados.
- Matrícula primaria por grados.
- Matrícula secundaria por grados.
- Matrícula media por grados.
- Matrícula CLEI por grados.
- Matrícula Otros modelos educativos por grados.
- Matrícula en preescolar por secciones del país y municipios.
- Matrícula en primaria por secciones del país y municipios.
- Matrícula en secundaria por secciones del país y municipios.
- Matrícula en media por secciones del país y municipios.
- Matrícula en CLEI por secciones del país y municipios.
- Matrícula en Otros modelos educativos para jóvenes y adultos por secciones del país y municipios.

Módulo VII: promoción escolar del año lectivo anterior.

- Promoción del año lectivo anterior en Preescolar según grados.
- Promoción del año lectivo anterior en Primaria según grados.
- Promoción del año lectivo anterior en Secundaria según grados.
- Promoción del año lectivo anterior en Media según grados.
- Promoción del año lectivo anterior en CLEI según grados.

- Promoción del año lectivo anterior en Otros modelos educativos para jóvenes y adultos según grados.

#### Módulo VIII. Información sobre tecnologías de la información y las comunicaciones.

- Uso y acceso de dispositivos electrónicos.
- Equipos de cómputo.
- Uso de los equipos de cómputo según actividad.
- Frecuencia del uso de los computadores, por niveles de enseñanza.
- Acceso y uso de los equipos de las sedes educativas.

#### Módulo IX: Información sobre ingresos, costos y gastos de la sede educativa.

- Ingresos percibidos por la prestación del servicio educativo /Urbano Rural.
- Gastos para la prestación del servicio educativo.
- Gastos y costos generales para la prestación del servicio educativo.
- Tarifa anual del sector oficial en Preescolar por grados escolares, ofrecidos por la sede educativa.
- Tarifa anual del sector oficial en Primaria por grados escolares, ofrecidos por la sede educativa.
- Tarifa anual del sector oficial en Secundaria por grados escolares, ofrecidos por la sede educativa.
- Tarifa anual del sector oficial en Media por grados escolares, ofrecidos por la sede educativa.
- Tarifa anual del sector oficial en CLEI por grados escolares, ofrecidos por la sede educativa.
- Tarifa anual del sector oficial en Otros modelos educativos para jóvenes y adultos por grados escolares, ofrecidos por la sede educativa.

#### DISEÑO DEL CUESTONARIO

Los formularios censales son el medio de recolección anual de la información de la investigación de educación formal. Estos instrumentos son diligenciados para cada una de las jornadas escolares ofrecidas a nivel de sede educativa. Su diseño responde a la dinámica de las necesidades de información oficial enmarcadas en los principales hitos normativos y de regulación sectorial.

#### TOPICS

Topic	Vocabulary	URI
Profesionales de la enseñanza [6.6]	CESSDA	
Educación obligatoria y preescolar. [6.2]	CESSDA	
Política educativa [6.3]	CESSDA	
Formación continua [6.4]	CESSDA	
Enseñanza post-obligatoria [6.5]	CESSDA	
Formación profesional [6.7]	CESSDA	

#### KEYWORDS

Docente, Educación Formal, Educación informal, Educación para el trabajo y el desarrollo humano, Educación no formal, Establecimiento educativo, Institución educativa, Jornada escolar, Jornada única, Matrícula, Niveles de educación formal, Sede educativa, Servicio educativo, Educación preescolar, Educación básica primaria, Educación básica secundaria, Educación media académica, Educación media técnica, Modelos educativos, Promoción escolar

## Cobertura

### COBERTURA GEOGRÁFICA

Total nacional, departamental y municipal tanto para la zona urbana como rural y para los sectores oficial y privado.

### GEOGRAPHIC UNIT

#### DESAGREGACIÓN GEOGRÁFICA

Total nacional, departamental y municipal tanto para la zona urbana como rural y para los sectores oficial y privado.

### UNIVERSO

Sedes educativas de educación formal, legalmente constituidas, del sector oficial y del sector no oficial, que ofrecen los niveles de preescolar, básica primaria, básica secundaria y media, localizados en el área urbana y en el área rural dentro del territorio nacional. El tamaño del universo es en el orden de 72.000 jornadas escolares, 58.000 sedes educativas y 25.000 instituciones o establecimientos (sedes principales).

## Productores y Patrocinadores

### INVESTIGADOR(ES) PRIMARIO(S)

Nombre	Affiliation
Investigación de Educación Formal	DIMPE - DANE

### OTROS PRODUCTORES

Nombre	Affiliation	Role
Ministerio de Educación Naciona		
Instituto Colombiano de fomento a la ecuación superior		
Departamento Administrativo Nacional de Estadística		

### FINANCIAMIENTO

Nombre	Abbreviation	Role
Departamento Nacional de Estadística	DANE	

## Producción de metadato

### METADATO PRODUCIDO POR

Nombre	Abbreviation	Affiliation	Role
Luisa Fernanda Suarez León	lfsuarezl@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización - DIRPEN	Coordinadora Regulación
Amelia Vargas Castiblanco	avargasc@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización - DIRPEN	Documentador DIMPE
Rafael Humberto Zorro Cubides	rhzorroc@dane.gov.co	Dirección de Regulación, Planeación, Estandarización y Normalización - DIRPEN	Actualizador DIRPEN
Marly Johanna Tellez Lopez	mjtellezl@dane.gov.co	Dirección de Información, Mercadeo y Cultura Estadística	Verificador DIMCE

### FECHA DE PRODUCCIÓN DEL METADATO

2016-07-07

### ID DEL DOCUMENTO IDD

COL-DANE-EDUC-2015

## Muestreo

No content available



# Formularios

## Información general

### DISEÑO DEL CUESTIONARIO

Los formularios censales son el medio de recolección anual de la información de la investigación de educación formal. Estos instrumentos son diligenciados para cada una de las jornadas escolares ofrecidas a nivel de sede educativa. Su diseño responde a la dinámica de las necesidades de información oficial enmarcadas en los principales hitos normativos y de regulación sectorial.

En el año de referencia 2015 la investigación de educación formal se sometió a los procesos de evaluación de la certificación del proceso estadístico, de la cual se derivaron las siguientes directrices y acotamiento técnico:

- La operación estadística por sus características se define como un censo de establecimientos educativos y no como una operación basada en registros administrativos.
- Se recomendó rediseñar los cuestionarios con el propósito de responder a las necesidades de información y reducir la carga de trabajo a la fuente informante, a partir de un único instrumento censal, teniendo como unidad de observación la sede educativa; cerrar el periodo de referencia y el periodo de recolección.

En atención a las recomendaciones derivadas del proceso de evaluación de la calidad y certificación y mas específicamente como un logo asociado al proceso de rediseño de la investigación de Educación Formal, se cuenta con un único formulario censal dispuesto desde 2014 para el reporte de la información en medio electrónico o en medio físico. La publicación de cifras para 2015 se realiza conforme a los resultados alcanzados con este último instrumento de recolección. El formulario único censal C600 está conformado por los siguientes módulos:

#### - MÓDULO I. Carátula única sede educativa

Hace referencia al conjunto de variables base para la actualización y mantenimiento del directorio estadístico.

##### \* Datos de identificación y ubicación de la sede educativa

Entre otros datos se recaba el nombre completo de la sede educativa, código DANE de identificación de la sede educativa, el número de identificación tributaria (NIT), Dirección de la sede educativa y las variables espaciales donde se localiza la sede educativa según las unidades político administrativas.

##### \* Naturaleza jurídica de la sede educativa:

La información hace referencia a la naturaleza jurídica de las sedes educativa y al sector prestador del servicio (oficial y no oficial), datos sobre el acto de reconocimiento (para sedes educativas del sector no oficial) o de creación o expedición de la licencia de funcionamiento (para sedes educativas del sector no oficial), el año en que inició operaciones y la clasificación de la sede según el régimen según la evaluación y aprobación de tarifas (solo para sector no oficial).

##### \* Modalidad de prestación del servicio educativo:

La información hace referencia al esquema administrativo que rige la prestación del servicio. Tiene por objetivo identificar si éste se presta de forma directa o si se registra a través de las diferentes modalidades de contratación que utiliza la entidad territorial certificada para garantizar el acceso y permanencia de la población en edad escolar dentro del sistema educativo. La orientación de este ítem está regido por la norma y la reglamentación asociada al Decreto 2355 de 2009 y Decreto 2500 de 2010.

En el módulo I de Caratula Única, es igualmente posible identificar la oferta educativa de la sede educativa por niveles /modelos educativos, las jornadas y el calendario. Así mismo si la población escolar esta interna, semi-externa o externa.

#### - MÓDULO II. Organización institucional para la prestación del servicio educativo

La información hace referencia a la organización que establece la entidad territorial certificada entre las sedes educativas

bajo su jurisdicción - ya sean estas del sector oficial o no oficial - para responder la demanda educativa en su territorio y garantizar la atención y retención de un estudiante en el sistema educativo desde preescolar hasta media. En el caso del sector oficial, la organización entre sedes principales y sedes adscritas, responde también a otra finalidad como es la de contar con una estructura institucional funcional que facilite la programación y ejecución del gasto público, por ejemplo asociado a los fondos educativos, y para promover otras acciones de política o de gestión sectorial como por ejemplo el nombramiento de las plantas de docentes y el escalafón.

#### - MÓDULO III. Información sobre personal ocupado por la sede educativa

La información solicitada en este numeral hace referencia al número de personas ocupadas en la sede educativa, durante el período de referencia, según la actividad principal que desempeña. Para el grupo de personas encargadas de las actividades de docencia se recaban datos sobre el tipo de vinculación laboral y el estatuto docente en el que se encuentran registrados o inscritos, el último nivel educativo alcanzado y el nivel de enseñanza donde tiene la mayor asignación o carga académica.

#### - MÓDULO IV. Información sobre matrícula, niveles de enseñanza, CLEI, modelos educativos en la sede educativa

A partir de las variables consultadas en este módulo es posible identificar la oferta educativa a nivel de la sede educativa desagregada por niveles /modelos educativos, jornadas y calendario. Es un módulo de control tipo línea de base para el año de referencia y de síntesis general para la investigación de educación formal. El registro de datos en este módulo se utiliza como marco de contraste y control de consistencia frente al reporte de datos en el módulo V que tiene como unidad de observación sede-jornada.

#### - MÓDULO V. Información sobre matrícula y docentes por jornada

En el módulo V la unidad de observación es la sede-jornada:

Los dominios de información que se recaban en este módulo están referidos a: i) matrícula, ii) docentes; iii) intensidad horaria.

La estructura de datos para la variable de matrícula se ordena por sexo y edades en años cumplidos, desagregada por niveles/ modelos educativos, acorde con el marco de la educación tradicional (preescolar, básica y media) y la educación no tradicional que da cabida a la enseñanza impartida a través de modelos educativos flexibles incluso de la educación para adultos (Decreto 3011 de 1997), y otras estrategias o programas - como por ejemplo la estrategia de aceleración del aprendizaje o la escuela nueva- que en suma están orientados a garantizar la formación básica formal y la permanencia en el sistema educativo. Acorde con la modalidad que rige la prestación del servicio público de educación, se clasifica la matrícula como propia o contratada, conservando para cada una de estas categorías la misma estructura de datos y desagregación temática y espacial.

La estructura de datos para la variable de intensidad horaria se ordena por carácter y asignatura según nivel de enseñanza/modelos educativos.

La estructura de datos para la variable de docentes se ordena por jornada según el último nivel educativo alcanzado y donde tenga la mayor asignación académica.

#### - MÓDULO VI. Poblaciones especiales en la jornada (Ley 115, título III).

La información solicitada en este módulo indaga sobre los estudiantes matriculados en la jornada, que se pueden clasificar en las siguientes tres categorías:

- \* Población matriculada con limitaciones físicas, discapacidad cognitiva y capacidades excepcionales;
- \* Población matriculada perteneciente a grupos étnicos;
- \* Población matriculada víctima del conflicto armado.

#### - MÓDULO VII. Promoción escolar del año lectivo anterior.

La información requerida en este módulo hace referencia al total de estudiantes aprobados, reprobados, en condición de deserción y transferidos/trasladados al finalizar el año lectivo anterior, por grado y sexo, en los niveles de enseñanza preescolar, básica primaria, básica secundaria, media, ciclos lectivos integrados CLEI (educación para adultos según decreto 3011/1997) y otros modelos educativos para jóvenes y adultos (no incluidos en los niveles de educación tradicional ni en el CLEI). Estos modelos hacen referencia a escuela nueva, pos primaria, círculos del aprendizaje, telesecundaria, caminar en secundaria, SAT, SER, CAFAM.

Para el calendario A, esta información corresponde a la situación académica reportada en el mes de noviembre del año

anterior y para calendario B a la reportada en el mes mayo del año calendario del año de referencia.

- MÓDULO VIII. Información sobre tecnologías de la información y las comunicaciones

El objetivo de este módulo es medir el uso que dan las sedes educativas a las tecnologías de información y comunicaciones (TIC).

- MÓDULO IX. Información sobre ingresos, costos y gastos de la sede educativa.

En este módulo se busca determinar la totalidad de los ingresos, costos y gastos operacionales asociados con la prestación del servicio educativo como actividad principal de la sede educativa, otros servicios secundarios o auxiliares tales como restaurante, transporte, manufacturas etc. Igualmente la tarifa anual por grados escolares y/o ciclos académicos ofrecidos por la sede educativa, que incluye el costo de la matrícula más el valor de las pensiones cobradas durante el año académico. En la tarifa no se incluyen los costos periódicos por concepto de transporte, alimentación, entre otros.

## Recolección de datos

### Fecha de recolección de datos

Start	End	Cycle
2015-05-26	2015-12-15	Anual calendario A y B

### Time Periods

Start	End	Cycle
2015		Anual

### Modo de recolección de datos

Autodiligenciamiento de formulario electrónico vía página web (libre; por ejemplo encuestas de satisfacción u opinión)

### Notas de recolección de datos

#### DISEÑO DE LA EJECUCIÓN

#### SISTEMA DE CAPACITACIÓN

##### - CAPACITACIÓN

El DANE Central a través del personal del componente de logística y temática, se encarga de capacitar a los asistentes técnicos de todas las sedes y subsedes (oficinas territoriales), sobre los lineamientos, conceptos y el correcto diligenciamiento del Formulario C-600. Éstos a su vez transfieren la información al personal operativo. DANE Central refuerza el entrenamiento en cada uno de las etapas del proceso operativo y de recolección a través de video conferencias y jornadas de asistencia y supervisión técnica presencial. Igualmente dispone en los diferentes medios de difusión los instructivos y manuales de trabajo (v.g crítica, validación y consistencia, diligenciamiento en el caso del instrumento de recolección, etc.). Los temas que se abordan en este proceso son:

\* Metodología de la investigación.

\* Conceptos y variables utilizados en el marco conceptual de la investigación.

\* Le da a conocer el manejo del formulario único censal C600 en físico y electrónico, con sus respectivos manuales de diligenciamiento, novedades y crítica.

\* Explicación de la normatividad sectorial

##### - SENSIBILIZACIÓN

Continuamente se diseñan e implementan actividades para sensibilizar a las sedes educativas sobre el reporte y entrega oportuna de la información, y a Secretarías de Educación en acciones relacionadas con el seguimiento y control de la cobertura. Parte de estas actividades se ejecutan a través jornadas de asistencia y supervisión técnica en sedes y subsedes del DANE o por video conferencia direccionada hacia el personal técnico de las secretarías de educación, los directores de núcleo, los rectores o informantes idóneos de las sedes educativas, y asistentes técnicos y personal operativo.

#### ACTIVIDADES PREPARATORIAS

##### - SELECCIÓN DE PERSONAL

El componente logístico de la investigación realiza los estudios previos y los instructivos de contratación del personal operativo, conforme a la programación y el manual de procedimientos para la selección del personal que establece el DANE. Posteriormente son enviados a las respectivas sedes y subsedes territoriales donde se ejecuta el proceso de contratación,

iniciando por la convocatoria - pública y/o por invitación directa- de acuerdo con los lineamientos establecidos por DANE Central.

El entrenamiento del personal operativo está a cargo de los asistentes técnicos responsables de la investigación de educación en las sedes y sub-sedes territoriales. Desde DANE central se envía con antelación el material definido para tal fin y se realiza el acompañamiento técnico necesario.

El equipo de trabajo en las sedes y subsedes está conformado por el responsable de la investigación educación formal, un supervisor y el grupo de críticos asignados. El rendimiento esperado por crítico es de 600 y 7000 formulario/mes.

#### - ALISTAMIENTO DEL DIRECTORIO ESTADÍSTICO

El directorio estadístico de educación, que administra el DANE, es un sistema donde se referencia todas las sedes educativas (unidad primaria de análisis), que prestan el servicio público de educación de básica formal en los niveles de enseñanza desde preescolar hasta media incluso educación por ciclos para adultos (CLEI - Decreto 3011/1994). Esta referenciación contiene los datos de identificación, ubicación, clasificación y tamaño.

Como resultado se obtiene el marco de lista que contiene la inclusión de todas las sedes activas que en el año de referencia anterior rindieron a la investigación de educación formal y aquellas que se encontraban inactivas. Se incluyen las sedes educativas que durante el año de referencia actual se han creado (nacimientos o altas). Se excluyen del alistamiento todas las sedes educativas que fueron liquidadas en el año de referencia anterior (cierres definitivos o bajas) y todas las sedes educativas que por su definición misional no pertenecen al universo de la investigación (p.ej centros de atención a primera infancia), y todas aquellas sedes que presenten cierre definitivo producto de procesos de absorción, fusión o integración institucional y que lleven a la desaparición jurídica y administrativa de la fuente informante.

Con antelación al inicio del operativo de recolección, el responsable de la investigación en cada sede y subsele DANE debe realizar el seguimiento a la notificación efectuada por el sistema a cada fuente informante para garantizar la cobertura censal.

#### - NOTIFICACIÓN SEDES EDUCATIVAS

A la sede o subsele DANE, se envía el directorio de las sedes educativas, para que el asistente técnico a su vez les notifique sobre la apertura de la plataforma de captura y diligenciamiento en forma electrónica y distribución de formularios físicos de la investigación de educación formal. La plataforma cuenta con un mecanismo automatizado para la autenticación de la fuente y seguidamente ingreso al formulario.

### RECOLECCIÓN DE LA INFORMACIÓN

#### - ESQUEMA OPERATIVO

El operativo de recolección, validación y análisis de consistencia de la Investigación de educación formal está a cargo de las sedes, subsedes DANE, bajo el direccionamiento del coordinador operativo y con la participación técnica y administrativa de las Secretarías de educación. El siguiente esquema representa la organización operativa:

#### - RECOLECCIÓN DE LA INFORMACIÓN

La información de la Investigación de educación formal, se recolecta en 2 fases:

- i) Fase I para las fuentes informantes de calendario A.
- ii) Fase II para las fuentes informantes de calendario B, CLEI (decreto 3011).

El método de recolección utilizado en la investigación es de auto-diligenciamiento en el formulario único censal (C600) físico y electrónico. El responsable del diligenciamiento de la información es el rector de la institución o sede educativa.

Una vez se cuenta con la información reportada por la fuente informante, ya sea a través de formularios físicos o a través del formulario electrónico, se da inicio a varios procesos:

- Crítica, validación y consistencia de la información y captura de datos: estos procesos son ejecutados directamente desde

la plataforma electrónica para cada una de las fuentes informantes. El sistema de captura está diseñado con reglas de validación internas que reducen la probabilidad de inconsistencias o registros errados. La tarea del crítico está centrada en la detección y el análisis de datos atípicos (que saltan en el formulario electrónico), que no son controlados por el sistema, y su ajuste o corrección después de confirmación por la fuente. Igualmente es labor del crítico, la pre- crítica y transcripción de la información reportada a través de los formularios físicos (en papel).

- Seguimiento al conjunto de fuentes informantes a fin de controlar la cobertura censal y la calidad de los datos ya reportados;

- Reentrenamiento, a cada uno de los eslabones de la cadena de valor, mediante jornadas de asistencia y supervisión técnica para el personal operativo en sedes y subsedes del DANE y para el grupo de trabajo de las secretarías de educación, directores de núcleo o sedes educativas;

- Asignación de novedades: a partir del proceso de confirmación del estado de la fuente informante (activas, inactivas, liquidadas, etc.), al cierre operativo. Este proceso se ejecute con la coparticipación de las secretarías de Educación certificadas.

## Formularios

### DISEÑO DEL CUESTIONARIO

Los formularios censales son el medio de recolección anual de la información de la investigación de educación formal. Estos instrumentos son diligenciados para cada una de las jornadas escolares ofrecidas a nivel de sede educativa. Su diseño responde a la dinámica de las necesidades de información oficial enmarcadas en los principales hitos normativos y de regulación sectorial.

En el año de referencia 2015 la investigación de educación formal se sometió a los procesos de evaluación de la certificación del proceso estadístico, de la cual se derivaron las siguientes directrices y acotamiento técnico:

- La operación estadística por sus características se define como un censo de establecimientos educativos y no como una operación basada en registros administrativos.

- Se recomendó rediseñar los cuestionarios con el propósito de responder a las necesidades de información y reducir la carga de trabajo a la fuente informante, a partir de un único instrumento censal, teniendo como unidad de observación la sede educativa; cerrar el periodo de referencia y el periodo de recolección.

En atención a las recomendaciones derivadas del proceso de evaluación de la calidad y certificación y más específicamente como un logo asociado al proceso de rediseño de la investigación de Educación Formal, se cuenta con un único formulario censal dispuesto desde 2014 para el reporte de la información en medio electrónico o en medio físico. La publicación de cifras para 2015 se realiza conforme a los resultados alcanzados con este último instrumento de recolección. El formulario único censal C600 está conformado por los siguientes módulos:

#### - MÓDULO I. Carátula única sede educativa

Hace referencia al conjunto de variables base para la actualización y mantenimiento del directorio estadístico.

##### \* Datos de identificación y ubicación de la sede educativa

Entre otros datos se recaba el nombre completo de la sede educativa, código DANE de identificación de la sede educativa, el número de identificación tributaria (NIT), Dirección de la sede educativa y las variables espaciales donde se localiza la sede educativa según las unidades político administrativas.

##### \* Naturaleza jurídica de la sede educativa:

La información hace referencia a la naturaleza jurídica de las sedes educativas y al sector prestador del servicio (oficial y no oficial), datos sobre el acto de reconocimiento (para sedes educativas del sector no oficial) o de creación o expedición de la licencia de funcionamiento (para sedes educativas del sector no oficial), el año en que inició operaciones y la clasificación de la sede según el régimen según la evaluación y aprobación de tarifas (solo para sector no oficial).

##### \* Modalidad de prestación del servicio educativo:

La información hace referencia al esquema administrativo que rige la prestación del servicio. Tiene por objetivo identificar si éste se presta de forma directa o si se registra a través de las diferentes modalidades de contratación que utiliza la entidad territorial certificada para garantizar el acceso y permanencia de la población en edad escolar dentro del sistema educativo. La orientación de este ítem está regido por la norma y la reglamentación asociada al Decreto 2355 de 2009 y Decreto 2500 de 2010.

En el módulo I de Caratula Única, es igualmente posible identificar la oferta educativa de la sede educativa por niveles /modelos educativos, las jornadas y el calendario. Así mismo si la población escolar esta interna, semi-externa o externa.

#### - MÓDULO II. Organización institucional para la prestación del servicio educativo

La información hace referencia a la organización que establece la entidad territorial certificada entre las sedes educativas bajo su jurisdicción - ya sean estas del sector oficial o no oficial - para responder la demanda educativa en su territorio y garantizar la atención y retención de un estudiante en el sistema educativo desde preescolar hasta media. En el caso del sector oficial, la organización entre sedes principales y sedes adscritas, responde también a otra finalidad como es la de contar con una estructura institucional funcional que facilite la programación y ejecución del gasto público, por ejemplo asociado a los fondos educativos, y para promover otras acciones de política o de gestión sectorial como por ejemplo el nombramiento de las plantas de docentes y el escalafón.

#### - MÓDULO III. Información sobre personal ocupado por la sede educativa

La información solicitada en este numeral hace referencia al número de personas ocupadas en la sede educativa, durante el período de referencia, según la actividad principal que desempeña. Para el grupo de personas encargadas de las actividades de docencia se recaban datos sobre el tipo de vinculación laboral y el estatuto docente en el que se encuentran registrados o inscritos, el último nivel educativo alcanzado y el nivel de enseñanza donde tiene la mayor asignación o carga académica.

#### - MÓDULO IV. Información sobre matrícula, niveles de enseñanza, CLEI, modelos educativos en la sede educativa

A partir de las variables consultadas en este módulo es posible identificar la oferta educativa a nivel de la sede educativa desagregada por niveles /modelos educativos, jornadas y calendario. Es un módulo de control tipo línea de base para el año de referencia y de síntesis general para la investigación de educación formal. El registro de datos en este módulo se utiliza como marco de contraste y control de consistencia frente al reporte de datos en el módulo V que tiene como unidad de observación sede-jornada.

#### - MÓDULO V. Información sobre matrícula y docentes por jornada

En el módulo V la unidad de observación es la sede-jornada:

Los dominios de información que se recaban en este módulo están referidos a: i) matrícula, ii) docentes; iii) intensidad horaria.

La estructura de datos para la variable de matrícula se ordena por sexo y edades en años cumplidos, desagregada por niveles/ modelos educativos, acorde con el marco de la educación tradicional (preescolar, básica y media) y la educación no tradicional que da cabida a la enseñanza impartida a través de modelos educativos flexibles incluso de la educación para adultos (Decreto 3011 de 1997), y otras estrategias o programas - como por ejemplo la estrategia de aceleración del aprendizaje o la escuela nueva- que en suma están orientados a garantizar la formación básica formal y la permanencia en el sistema educativo. Acorde con la modalidad que rige la prestación del servicio público de educación, se clasifica la matrícula como propia o contratada, conservando para cada una de éstas categorías la misma estructura de datos y desagregación temática y espacial.

La estructura de datos para la variable de intensidad horaria se ordena por carácter y asignatura según nivel de enseñanza/modelos educativos.

La estructura de datos para la variable de docentes se ordena por jornada según el último nivel educativo alcanzado y donde tenga la mayor asignación académica.

#### - MÓDULO VI. Poblaciones especiales en la jornada (Ley 115, título III).

La información solicitada en este módulo indaga sobre los estudiantes matriculados en la jornada, que se pueden clasificar en las siguientes tres categorías:

- \* Población matriculada con limitaciones físicas, discapacidad cognitiva y capacidades excepcionales;
- \* Población matriculada perteneciente a grupos étnicos;

\* Población matriculada víctima del conflicto armado.

- MÓDULO VII. Promoción escolar del año lectivo anterior.

La información requerida en este módulo hace referencia al total de estudiantes aprobados, reprobados, en condición de deserción y transferidos/trasladados al finalizar el año lectivo anterior, por grado y sexo, en los niveles de enseñanza preescolar, básica primaria, básica secundaria, media, ciclos lectivos integrados CLEI (educación para adultos según decreto 3011/1997) y otros modelos educativos para jóvenes y adultos (no incluidos en los niveles de educación tradicional ni en el CLEI). Estos modelos hacen referencia a escuela nueva, pos primaria, círculos del aprendizaje, telesecundaria, caminar en secundaria, SAT, SER, CAFAM.

Para el calendario A, esta información corresponde a la situación académica reportada en el mes de noviembre del año anterior y para calendario B a la reportada en el mes mayo del año calendario del año de referencia.

- MÓDULO VIII. Información sobre tecnologías de la información y las comunicaciones

El objetivo de este módulo es medir el uso que dan las sedes educativas a las tecnologías de información y comunicaciones (TIC).

- MÓDULO IX. Información sobre ingresos, costos y gastos de la sede educativa.

En este módulo se busca determinar la totalidad de los ingresos, costos y gastos operacionales asociados con la prestación del servicio educativo como actividad principal de la sede educativa, otros servicios secundarios o auxiliares tales como restaurante, transporte, manufacturas etc. Igualmente la tarifa anual por grados escolares y/o ciclos académicos ofrecidos por la sede educativa, que incluye el costo de la matrícula más el valor de las pensiones cobradas durante el año académico. En la tarifa no se incluyen los costos periódicos por concepto de transporte, alimentación, entre otros.

## Data Collectors

Nombre	Abbreviation	Affiliation
Investigación de educación formal	EDUC	DIMPE- DANE

## Supervisión

### DISEÑO DE MÉTODOS Y MECANISMOS PARA EL CONTROL DE CALIDAD.

En la investigación de educación formal, más allá de las rutinas y procesos sistemáticos propios del modelo de producción, se implementan varios procesos y mecanismos para el control de la calidad, que involucra la gestión técnica de los diferentes componentes de DANE (logística, temática, sistemas y metodología estadística), y otras fuentes externas auxiliares como es el caso de las secretarías de educación certificadas y el Ministerio de Educación Nacional.

- Métodos y mecanismos de control de calidad (datos y base de datos)

El sistema de captura está diseñado con arreglo a una serie de reglas de validación, completitud y de aceptabilidad o rechazo que operan intra e inter módulos durante la fase de captura o auto diligenciamiento de los datos, con lo cual se reduce las probabilidades de grabar módulos con inconsistencias o registros errados y se anula la ocurrencia de formularios incompletos. No obstante, es posible que el sistema deje pasar datos atípicos o que no tienen referente de validación, que son detectados en las fases posteriores de crítica, análisis de consistencia, depuración y corrección.

Para ejecutar estos procesos, la investigación de educación formal cuenta con varios instrumentos y acciones de control:

- i) Módulo de captura con sistema integral de consistencia.
- ii) Ficha de análisis.
- iii) Matriz de variaciones interanuales para las variables estructurales.
- iv) Monitoreo del avance del operativo de recolección o rendición de información.
- v) Modelo de control de calidad.

- Módulo de captura con sistema integral de consistencia



Este proceso es ejecutado directamente desde la plataforma electrónica para cada una de las fuentes informantes. El sistema de captura está diseñado con reglas de validación internas que reducen la probabilidad de inconsistencias o registros errados. La tarea del crítico está centrada en la detección y el análisis de datos atípicos (que saltan en el formulario electrónico), que no son controlados por el sistema, y su ajuste o corrección después de confirmación por la fuente. Igualmente es labor del crítico, la pre- crítica y transcripción de la información reportada a través de los formularios físicos (en papel).

#### - La ficha de análisis

Es una funcionalidad disponible dentro del sistema de educación formal (SIEF), a partir de la cual, después de la captura de los datos (ya sea a través de formulario físico o formulario electrónico), se hacen visibles las variables y los valores susceptibles de control o corrección, labor ejecutada por el crítico después de revisar la coherencia interna del formulario o como resultado de la confirmación que se establece mediante consulta directa a la fuente primaria.

#### - Matriz de variaciones interanuales para las variables estructurales

Se cuenta con un formato en Excel con reglas de validación que permiten detectar variaciones interanuales en las variables de docentes y matrícula, de tal manera que permite realizar los ajustes y/o correcciones necesarias en la información registrada en aquellas sedes educativas que presentan inconsistencia y obtener una información con calidad. Estos ajustes se realizan consultando a las sedes educativas directamente quienes son la fuente primaria de la investigación.

#### - Modelo de control de calidad

El componente de temática corre un modelo de control de calidad, que funciona como un tablero de control (variaciones interanuales, promedios históricos, media histórica, razones, totales anuales etc.), con información histórica de la investigación de educación formal, información del MEN tomada a través de servicios web al sistema de matrícula - SIMAT y al Directorio único de establecimiento - DUE, e información de contraste DANE /MEN. Como resultado se obtiene la lista las fuentes con datos inconsistentes, atípicos, datos nulos etc., con desagregación a su distribución por entidad territorial, entidad territorial certificada de educación y por la segmentación geográfica o por otros criterios que se definan desde los componentes de la investigación.

La acción correctiva puede derivar en la devolución del formulario al componente de logística para que mediante la gestión ante la fuente primaria se alcance la consistencia necesaria, o también en una solicitud al componente de sistemas, cuando las inconsistencias no se originan en la fase de reporte, captura o crítica, sino que obedecen a fallas en la funcionalidad del módulo de captura o de control operativo en SIEF.

#### - Métodos y mecanismos de control consolidación base de datos

Para garantizar la calidad, integridad y completos de la base de datos se sigue una serie de rutinas que se implementan después del primer cierre operativo. Con éste se entrega la base de recolección preliminar que ya trae incorporada los ajustes, depuración y consistencia alcanzada durante la etapa de recolección. El control de calidad de la base de recolección se realiza conforme a las siguientes rutinas:

Se verifica la cobertura alcanzada para las variables estructurales de la investigación. El análisis se realiza agregando desde el nivel básico a nivel de sede- jornada, pasando a sede educativa, agregando a municipio y cerrando hasta departamento y entidad territorial certificada de educación. Para ello se utiliza un tablero de control que contiene:

- i) La variación interanual y los promedios históricos de los últimos tres años registrados en C600.
- ii) El registro de matrícula de SIMAT y de DUE, para el año de referencia, que se obtiene del servicio web que dispone el MEN al servicio de DANE.
- iii) El contraste de cifras en ordenes de magnitud y valores porcentuales DANE - MEN.

Desde el administrador del sistema SIEF se corre una serie de rutinas para detectar:

- i) Municipios sin información.
- ii) Fuentes en deuda.
- iii) Fuentes con deuda parcial.
- iv) Variables con datos nulos.
- v) Valores y atributos de las variables inconsistentes entre e inter módulos.
- vi) Ordenes de magnitud atípicos con variaciones o valores absolutos por fuera de los rangos de aceptabilidad.

Después de contar con la base de recolección definitiva (después de controles de calidad), se cierra para evitar alguna

modificación o afectación del registro y se procede a la preparación de los archivos para ejecutar el proceso de imputación. De la base cerrada se obtiene la lista de candidatos a imputar por deuda total o deuda parcial. El porcentaje de imputación siempre es menor que el porcentaje de deuda (total + parcial).

La base de datos definitiva se obtiene después de imputación, Los procesos de control en esta etapa se ajustan a:

- 1) Las rutinas de calibración, completitud e integridad que corre el grupo de metodología estadística con el objeto de garantizar el correcto ensamble o articulación de los datos imputados en la estructura de tablas de la base de datos y la coherencia de la información.
- 2) Se vuelven a correr los procesos indicados para blindar la base de recolección.

Antes de generar los cuadros de salida, anexos y la publicación de cifras y resultados se revisan los órdenes de magnitud y la tendencia en cada nivel de análisis y desagregación geográfica, por entidad territorial certificada de educación, por sector, por zona y variables centrales (matrícula, docentes, jornadas).

## - MÉTODOS Y MECANISMOS DE CONTROL OPERATIVO

- Monitoreo al avance del operativo de recolección o rendición de información

Es una funcionalidad disponible dentro del sistema de educación formal -SIEF (control Operativo -cobertura, a partir de la cual se realiza el seguimiento y control a las fuentes informantes, según el estado registrado con respecto al reporte de datos al reporte o ingreso de la información o diligenciamiento del formulario (sin digitar, digitado, en digitación, análisis de verificación y verificados), y según la novedad asignada (deuda, liquidadas, inactivas etc.).

La funcionalidad está diseñada para el monitoreo operativo a nivel de:

- i) Fuente primaria de información.
- ii) Sede y subsele territorial de DANE.
- iii) Entidad territorial certificada (secretaría de educación).

## DISEÑO DE PRUEBAS

En la investigación se ejecutan pruebas de funcionalidad en el ambiente de pruebas que dispone el SIEF. Para ello es necesario establecer previamente el requerimiento y las especificaciones para el desarrollo requerido para luego probar su funcionalidad y paso a producción.

El propósito de un caso de prueba es especificar la manera de probar el sistema, incluyendo las entradas, los resultados esperados. Los requerimientos son la fuente principal para obtener los casos de prueba. Se debe preparar la matriz de pruebas para poder controlar el proceso de pruebas de aceptación y determinar cuál es la duración de las pruebas del sistema.

El área de plataforma es la responsable de verificar que la solicitud de paso a producción sea viable de cumplir. El grupo de plataforma verifica los requerimientos de software, hardware, bases de datos y demás requisitos para la puesta en producción del software. Cuando la solicitud ha sido verificada por el responsable del grupo de plataforma, se procede con la autorización para salida a producción.

# Procesamiento de datos

## Edición de datos

---

### PROCESAMIENTO Y CONSOLIDACIÓN BASE DE DATOS

Los procesos de captura, crítica, validación y consistencia de la información: son ejecutados directamente desde la plataforma electrónica para cada una de las fuentes informantes. El sistema de captura está diseñado con reglas de validación internas que reducen la probabilidad de inconsistencias o registros errados y con funcionalidades que permiten visibilizar datos atípicos o sin referentes de validación durante el proceso de recolección y hasta el cierre operativo y entrega de la base de datos para los procesos finales de imputación y generación de cuadros de salida y anexos para difusión.

Todo el acervo de datos es almacenado directamente en la base de datos Oracle. Las tablas de la base de datos tienen relación entre sí, de acuerdo a las reglas de normalización, además de criterios como: llave primaria, llave foránea e índices, que permiten una mejor interrelación, desempeño y funcionalidad en la base de datos.

Adicionalmente se han creado vistas de base de datos para que los componentes de logística y temática puedan consultar el micro dato en cualquier momento y realizar cruces de información para detectar posibles inconsistencias en la información recolectada, las cuales son corregidas en la ficha de análisis del sistema. Estas vistas son consultadas por medio de la herramienta SAS y también son usadas para la generación de boletines.

### MECANISMOS DE SEGURIDAD

La información almacenada en la base de datos solo es accesible a los perfiles por medio del sistema de información, esto ayuda a garantizar la integridad de la información de la base de datos. Para el backup de la base de datos se siguen los lineamientos de la Oficina de Sistemas, por consiguiente se realiza un backup incremental diariamente y completo semanalmente.

## Otros procesamientos

---

### INDICADORES DE RESULTADO CONTROL DE CALIDAD

Al finalizar el operativo censal, se realiza el cálculo de los indicadores de calidad, que ofrecen una medida sobre el resultado alcanzado durante el año de referencia relacionado con la gestión operativa, cobertura censal y calidad de la información recabada. El compendio de indicadores se obtienen en diferentes niveles de desagregación conforme a su distribución por entidad territorial, entidad territorial certificada de educación y por la segmentación geográfica o por otros criterios que se definan desde los componentes de la investigación (temática, sistemas, logística, metodología estadística).

#### - Indicador de cobertura estadística

Permite conocer el porcentaje de sedes educativas que rindieron información del total de sedes contenidas en el directorio

#### - Indicador de no respuesta verificada

Presenta el porcentaje sedes educativas que no reportaron información en el operativo y/o que fueron identificadas como deuda por ejercicios posteriores a partir de información auxiliar, análisis panel o de intermitencia, del total de sedes contenidas en el directorio.

#### - Indicador de no respuesta operativa

Presenta el porcentaje de sedes educativas que no reportaron información en el operativo del total de sedes contenidas en el directorio.

- Indicador de calidad

El indicador de calidad permite conocer el porcentaje de sedes educativas cuya información recolectada tuvo que ser sometida a correcciones por presentar alarmas en las validaciones de control de calidad de la investigación.

## Estimación de datos



## Estimación de error de la muestra

### DISEÑO DEL ANÁLISIS DE RESULTADOS

Como resultado de las fases precedentes, anualmente se obtiene una base de datos consolidada, corregida, ajustada y actualizada, la cual ingresa a una fase de procesamiento para generar los cuadros de salida, los anexos y la publicación de cifras y resultados. Dentro del procesamiento se revisan los órdenes de magnitud y la tendencia en cada nivel de análisis, así como por desagregación geográfica, por entidad territorial certificada de educación, por sector y zona.

### ANÁLISIS ESTADÍSTICO

#### - ANÁLISIS DE COHERENCIA

Después de contar con la base de recolección definitiva (después de controles de calidad), se cierra para evitar alguna modificación o afectación del registro y se procede a la preparación de los archivos para ejecutar el proceso de imputación. De la base cerrada se obtiene la lista de candidatos a imputar por deuda total o deuda parcial. El porcentaje de imputación siempre es menor que el porcentaje de deuda (total + parcial).

La base de datos definitiva se obtiene después de imputación, Los procesos de control en esta etapa se ajustan a:

- 1) Las rutinas de calibración, completitud e integridad que corre el grupo de metodología estadística con el objeto de garantizar el correcto ensamble o articulación de los datos imputados en la estructura de tablas de la base de datos y la coherencia de la información.
- 2) Se vuelven a correr los procesos indicados para blindar la base de recolección realizando un ejercicio de contraste y panel con las sedes que se reportan en el SIMTA y DUE del MEN.

Antes de generar los cuadros de salida, anexos y la publicación de cifras y resultados se revisan los órdenes de magnitud y la tendencia en cada nivel de análisis y desagregación geográfica, por entidad territorial certificada de educación, por sector, por zona y variables centrales (matrícula, docentes, jornadas).

#### - ANÁLISIS DE ESTADÍSTICAS PRODUCIDAS

Se revisan una a una las variables de los directorios de instituciones educativas, sedes y jornadas para detectar faltantes o valores perdidos. Este mismo procedimiento se realiza con la información de matrícula y docentes.

Con la ayuda de la serie nacional, departamental y municipal para la información de matrícula se revisa la consistencia de las participaciones dentro del total de cada nivel educativo, sector, zona, género, grados y edades.

Para el caso de los docentes, se revisa la consistencia de las participaciones dentro del total de cada nivel educativo, género y último nivel educativo aprobado por el docente. Y para el caso de los establecimientos educativos, se revisa la consistencia de las participaciones dentro del total del sector y nivel educativo. Durante este proceso, se calculan las tasas de crecimiento de la información de matrícula, docentes y establecimientos, a nivel municipal, departamental y nacional.

#### - ANÁLISIS DE CALIDAD DE RESULTADOS

Consolidada y depurada la base de datos de recolección, se determinan las instituciones educativas, sedes y jornadas que se encuentran en deuda. Con base en la información de matrícula y docentes del año inmediatamente anterior, se cuantifica el total de matrícula, por nivel educativo, zona y sector; así como el total de docentes que se encuentra representado en dicha deuda.

Se utiliza el medio gráfico para verificar si la composición de la matrícula por sector, zona, grados, género y edades se mantiene. Este mismo análisis se realiza para revisar la estructura de la información de docentes y establecimientos educativos.

#### - ANÁLISIS DE CONTEXTO

Se verifica que los totales de matrícula por nivel educativo de cada uno de los municipios y departamentos se encuentren dentro de los rangos establecidos por la serie. Si no es así, se confirma si dicho comportamiento se encuentra justificado por la deuda o por algún otro efecto.

La revisión de los temas especiales como: matrícula de aceleración del aprendizaje, jóvenes en extra edad o adultos, alumnos con discapacidades o capacidades excepcionales, grupos étnicos, víctimas del conflicto, por modelos pedagógico, información de recurso humano, docentes por grados en escalafón, por áreas de enseñanza y docentes en comisión, se hace comparando los totales actuales con los totales de la serie nacional y departamental.

Para facilitar el análisis y comprobar la coherencia, consistencia y comparabilidad de la información, se grafican las series de matrícula, docentes y establecimientos, incluyendo el nuevo dato analizado. Finalmente, se elabora la presentación y la ficha técnica de la información, la cual debe ser expuesta anualmente al Comité de expertos.

#### - COMITÉS DE EXPERTOS

Antes de publicar la cifras de investigación, se presentan los resultados alcanzada en tres espacios de carácter técnico:

- i) Comité técnico de la investigación de educación formal con el equipo asesor de la Dirección de Metodología y Producción Estadística.
- ii) Comité interno con la participación de las diferentes direcciones técnicas del DANE y el componente asesor de la Dirección y Subdirección del Departamento.
- iii) Comité externo con participación de los integrantes del comité interno y delegados técnicos del Ministerio de Educación Nacional, el ICFES y el Departamento Nacional de Planeación (DNP).

## Otras formas de estimar datos

### IMPUTACIÓN

Los procesos de imputación son métodos con sustento estadístico que se aplican para el tratamiento de la no respuesta de las fuentes informantes dentro del periodo de recolección. En la investigación de educación formal hay dos tipos de no respuesta o deuda: total y parcial. La primera está conformada por las fuentes que no rindieron información, pero que se encontraban en el directorio estadístico del año anterior, y la no respuesta parcial que se refiere a la ausencia de información a nivel de los diferentes módulos o variables observados a través del instrumento de recolección (formulario único censal C600).

La inclusión o selección de las fuentes en deuda sujetas a imputación se realiza conforme a varios criterios:

- Trazabilidad histórica de la fuente y consistencia del registro.
- Fuentes que rindieron información en el periodo inmediatamente anterior al periodo a imputar y que quedaron en deuda en el año actual de referencia.
- Fuentes en deuda parcial.
- Parámetros de inclusión forzosa.

El proceso de imputación se aplica para las variables contenidas en el cuestionario y que en términos generales incluyen información de Identificación de la sede - jornada, conteo de jornadas, relación entre sedes (principales y adscritas), número de estudiantes por sede - jornada, para cada nivel y modelo educativo. Información del personal ocupado e información de promoción de la sede educativa.

El proceso de imputación de datos (año de referencia 2014) combinó el comportamiento histórico de la sede educativa a imputar con el comportamiento de las sedes educativas que comparten características similares a ella (donantes).

Se parte de los valores observados en el periodo inmediatamente anterior en la sede a imputar, las dos características principales en el proceso de imputación son: el histórico de la sede educativa y la tasa de crecimiento de la matrícula para unidades de observación (sedes educativas) con características similares. Las sedes educativas a imputar corresponden a aquellas sedes que rindieron información en el periodo inmediatamente anterior al periodo a imputar y/o aquellas de las que se conoce información auxiliar confiable y actualizada.

La caracterización de los donantes es la siguiente:

- Sedes que se encuentren en el mismo municipio que la sede a imputar.
- Sedes que pertenezcan a la misma zona (rural/urbano).
- Sedes que pertenezcan al mismo sector (oficial/no oficial).
- Sedes que sean de un tamaño similar a la sede a imputar (que el número de matriculados no sea superior o inferior al 20% de la sede imputada según la información histórica).

Para evitar comportamientos atípicos en la información, se exime de las sedes donantes a las sedes educativas que hayan tenido un comportamiento de rendición intermitente (de más de 5 ausencias de información), a partir del año 2005. Teniendo en cuenta los ejercicios de simulación y la consideración de un comité de expertos, se concluye que la imputación debe realizarse sobre las sedes educativas que rindieron el periodo inmediatamente anterior. La metodología construida es óptima para tasas de imputación bajas (inferiores al 5%), en otro caso deberá validarse la aplicación del modelo de imputación.

Las limitaciones de la metodología de imputación y en donde se decide imputar replicando la información observada en el periodo inmediatamente anterior son:

- Municipios con menos de 5 sedes educativas.
- Sedes donde no se encuentren donantes con sus mismas características (municipio, zona, sector y rango de matrícula).
- Al calcular la diferencia entre el valor imputado y el valor observado en el periodo inmediatamente anterior es superior al 20% del valor de la matrícula.

El proceso de imputación se realiza en el software estadístico SAS. El código en primera instancia prepara los históricos de matrícula y docentes de las sedes educativas para el periodo de imputación y los periodos inmediatamente anteriores según el módulo tratado. Luego para cada sede educativa en deuda elige el grupo de sedes educativas donantes y calcula la tasa de crecimiento de la matrícula de los donantes y la multiplica por la matrícula histórica de la sede educativa en deuda, obteniéndose así la matrícula imputada. Finalmente se distribuye la matrícula imputada a la sede según la estructura histórica de la sede educativa y la estructura del módulo tratado.

Se entregan las tablas imputadas según la estructura de los diferentes módulos de la investigación con los registros imputados, para que sean cargados por el componente de sistemas a las bases de datos de la investigación y sean sometidos a posteriores análisis de consistencia por los diferentes componentes.

Como consideración general, debe revisarse que los totales generales correspondan a la suma de los totales segregados, debido a que el algoritmo de imputación redondea los valores de matrícula y docentes lo que puede generar diferencias que deben corregirse sumando o restando el valor de ésta diferencia para garantizar trazabilidad entre los módulos.

Dado que se establecen supuestos para el buen comportamiento del modelo, éste debe estar bajo supervisión constante del equipo de metodología estadística, con el fin de evaluar su desempeño, comparando la distancia de los valores reales con los imputados en el conjunto de registros con información completa. En el evento que se identifiquen cambios en el comportamiento general de las fuentes, deberán ajustarse los parámetros del modelo, así mismo, debe considerarse el alcance del uso de la serie histórica de la sede educativa para complementar la metodología.