

Departamento Administrativo Nacional de Estadística



Diseño
DSO

Dirección de Metodología y Producción
Estadística-DIMPE

**Manual Anonimización de las Bases de
la Encuesta Sobre Ambiente y
Desempeño Institucional Departamental
EDID 2021**

Mayo 2022

TABLA DE CONTENIDO

<u>INTRODUCCIÓN.....</u>	<u>2</u>
<u>MARCO CONCEPTUAL</u>	<u>3</u>
<u>1.1 OBJETIVOS DEL DOCUMENTO</u>	<u>3</u>
<u>1.2 CARACTERÍSTICAS DE LA ENCUESTA.....</u>	<u>3</u>
<u>Objetivos De La Encuesta</u>	<u>3</u>
<u>Objetivo General</u>	<u>3</u>
<u>OBJETIVOS ESPECÍFICOS.....</u>	<u>3</u>
<u>¿PERIODO DE RECOLECCIÓN?</u>	<u>4</u>
<u>¿A QUIÉNES SE LES APLICA LA ENCUESTA?.....</u>	<u>4</u>
<u>1.3 ESTRUCTURA CUESTIONARIO</u>	<u>4</u>
<u>1.4 PROCESAMIENTO DE DATOS</u>	<u>6</u>
<u>1.5 DISPOSICION DE DATOS</u>	<u>10</u>
<u>1.6 APLICACIÓN DE LINEAMIENTOS PARA ANONIMIZACIÓN</u>	<u>10</u>
<u>1.7 DICCIONARIO BASE DE DATOS.....</u>	<u>11</u>

INTRODUCCIÓN

Los procedimientos para asegurar la reserva estadística o preservar la confidencialidad de la información entregada por las fuentes de información, a las agencias estadísticas nacionales, entre ellas el DANE, a través de censos, encuestas por muestreo o registros administrativos, y permitir el uso de los microdatos para estudios de carácter especializado, se dividen en dos:

El primero, y más aplicado, es el de suprimir o quitar variables, el cual consiste en una revisión de las variables que hacen parte de la base de datos de las encuestas que se realizan y que se quiere entregar para uso público con el fin de suprimir de la misma, aquellas que pudieran revelar la identidad, la ubicación o usarse para determinar la posición relativa de la fuente de información.

El segundo, menos usado, consiste en transformar mediante funciones de naturaleza matemática o de otro tipo, las variables de las encuestas que se quieren entregar para uso público, sin distorsionar las relaciones técnicas entre las variables que se encuentra en la forma original y la transformada.

En el caso del DANE, el procedimiento que se implementará es el de suprimir variables, para poner a disposición del público las bases de datos que preserven la reserva estadística y aseguren la confidencialidad de la información, pero que permitan su uso para estudios más amplios que los que tradicionalmente entrega la institución.

MARCO CONCEPTUAL

1.1 OBJETIVOS DEL DOCUMENTO

- **Objetivo General**

Definir los procedimientos de carácter técnico para preservar la reserva estadística en los microdatos de la base de datos insumo de la Encuesta sobre ambiente y desempeño institucional departamental EDID 2021, que el DANE entregará para uso público.

- **Objetivos Específicos**

Determinar las variables, a nivel de microdato, producto de la recolección de información que realiza el DANE a través de la encuesta sobre ambiente y desempeño institucional departamental a servidores(as) públicos de 32 gobernaciones, 30 alcaldías de ciudades capital de departamento y de Bogotá Distrito Capital (Alcaldía y 13 Secretarías Distritales); las cuales serán de uso público, bajo licencia y en sitio especializado, preservando la reserva estadística, el principio de confidencialidad y siguiendo las recomendaciones internacionales.

1.2 CARACTERÍSTICAS DE LA ENCUESTA

OBJETIVOS DE LA ENCUESTA

Objetivo General

Conocer la percepción que tienen las y los servidores públicos sobre el ambiente y el desempeño de las entidades a las que prestan sus servicios, a partir de su experiencia laboral y personal durante los últimos doce meses.

Objetivos Específicos

- Obtener información sobre la percepción de las y los servidores públicos respecto al ambiente institucional de las entidades, a partir de su experiencia personal de trabajo en la entidad y su percepción sobre la gestión del talento humano en la organización.
- Recoger información sobre la percepción de las y los servidores públicos respecto al desempeño institucional de las entidades, a través de su opinión sobre diferentes políticas organizacionales relacionadas con la gestión de la entidad, el uso de tecnologías e información, la rendición de cuentas y la prevención de prácticas irregulares.
- Generar indicadores de desarrollo de la administración pública territorial, que permitan clasificar las organizaciones burocráticas en un momento dado y comparar su evolución a lo largo del tiempo.

- **¿A quién se dirige la encuesta?**

A las gobernaciones, alcaldías de capital de departamento y el Distrito Capital (Bogotá D.C.).

- **¿Periodo de recolección?**

La recolección de los datos se realizó del 22 de julio al 5 de octubre de 2021.

- **¿A quiénes se les aplica la encuesta?**

Servidores(as) públicos(as) que laboran al nivel central de las gobernaciones, alcaldías de capital de departamento y del Distrito Capital (Bogotá D.C.), los cuales son clasificados de la siguiente forma:

- Nivel 1. Directivo/ asesor.
- Nivel 2. Profesional/ técnico.
- Nivel 3. Asistencial.

- **¿Cómo se recolecta la información?**

La recolección de la encuesta se realiza mediante dos métodos:

- Principalmente por auto-diligenciamiento de un formulario electrónico a través de la página web del DANE.
- Auto diligenciamiento a través de formulario físico.

1.3 ESTRUCTURA CUESTIONARIO

El cuestionario consta de 220 literales que permiten profundizar aspectos relacionados con el ambiente y el desempeño institucional de las entidades públicas del orden nacional, a través de las percepciones de sus servidores. El siguiente cuadro resume la estructura del cuestionario.

Tabla1. Estructura de cuestionario

Nombre	Descripción	Preguntas (P), literales (L) y capítulos (CAP)
INSTRUCCIONES GENERALES	Contiene la información básica de la encuesta, a partir de seis ayudas dirigidas a conocer el objetivo general, el periodo de referencia, los beneficios que se obtienen con la información recolectada, su carácter confidencial, el momento de publicación de los resultados y la disponibilidad de los resultados de versiones anteriores.	
MÓDULO INFORMACIÓN SOCIODEMOGRÁFIC A.	I. Contiene toda la información relacionada con la ubicación laboral del encuestado.	(P) = 7
UBICACIÓN LABORAL		
MÓDULO AMBIENTE INSTITUCIONAL	II. Los capítulos que contiene buscan conocer la percepción de las y los servidores públicos respecto a la experiencia laboral que han vivido durante los últimos doce meses en la organización pública a la cual se encuentran vinculados.	CAP = 2
CAPÍTULO EXPERIENCIA PERSONAL TRABAJO	C. Las preguntas buscan conocer su percepción como servidor(a) público sobre las experiencias laborales que los encuestados han vivido durante los últimos doce meses en la entidad	(P) = 24 (L) = 100
CAPÍTULO GESTIÓN TALENTO HUMANO	D. Los enunciados buscan conocer su percepción sobre la gestión del talento humano realizada por la organización en la cual trabajan.	(P) = 13 (L) = 36
MÓDULO DESEMPEÑO INSTITUCIONAL	III. Los capítulos que contiene buscan conocer la percepción de las y los servidores públicos respecto a su experiencia frente a la gestión de la entidad y ante temas de política pública relevantes para las organizaciones del Estado.	CAP = 4
CAPÍTULO E. USO DE TECNOLOGÍA INFORMACIÓN	E. Las preguntas del este capítulo indagan por las herramientas, tecnologías e información utilizadas por los servidores/as para desempeñar sus funciones.	(P) = 9 (L) = 21
CAPÍTULO GESTIÓN ENTIDAD	F. Los enunciados buscan conocer la percepción de las y los servidores con respecto a las prácticas al interior de la organización orientadas a fortalecer la gestión institucional.	(P) = 6 (L) = 31

CAPÍTULO RENDICIÓN CUENTAS	G. DE	Los enunciados indagan por el proceso de rendición de cuentas en la entidad y el acceso a la información pública.	(P) = 7 (L) = 19
CAPÍTULO PREVENCIÓN PRÁCTICAS IRREGULARES	H. DE	Los enunciados indagan sobre prácticas y comportamientos esperados en las organizaciones públicas que contribuyen a la prevención de prácticas irregulares.	(P) =6 (L) =6
TOTAL EDID 2021			(P) = 72 (L) = 220

1.4 PROCESAMIENTO DE DATOS

El procedimiento para anonimizar o preservar la reserva estadística en las bases de microdatos de la Encuesta sobre ambiente y desempeño institucional EDID 2021 que el DANE entregará para uso público es el de suprimir variables de esas bases.

Con este propósito, la base de la EDID 2021 tiene la siguiente estructura:

1. Variables de identificación: corresponde al conjunto de variables que identifican la información relacionada con la información suministrada por las oficinas de recursos humanos de cada una de las entidades objeto de estudio. (Capítulo B: Ubicación laboral).
2. Variables observadas: corresponde al conjunto de información que se toma directamente en las fuentes como parte del operativo de recolección. (Modulo I a módulo II; capítulo C a capítulo H)
3. Variables de proceso: estas variables están relacionadas con el procesamiento posterior que sufre la información con el fin de generar el índice.

En ese sentido, el proceso para entregar una base de microdatos anonimizados a partir de la información insumo de la EDID que preserve la reserva estadística consiste en suprimir bloques de variables, o variables dentro de los bloques mencionados.

Las exclusiones se establecen puntualmente en las variables de identificación, específicamente para los datos relacionadas con el nivel jerárquico, el tiempo de servicio en la entidad, nivel educativo y edad. Estas variables no se incluyen en la base con el propósito de eliminar la posibilidad de identificar una fuente, especialmente cuando se consulta los resultados de las entidades que cuentan con menor tamaño de muestra.

Como parte del proceso de anonimización, en el [Archivo Nacional de Datos - ANDA](#) se pone a disposición del público los microdatos en dos archivos:

- La información marcada como EDID_2021_REGION, contiene los datos agregados a nivel de regiones (no se incluyen datos por entidad). de acuerdo con las regiones definidas por el DANE para la desagregación de información regional, las cuales se conforman de la siguiente manera:
-
- - **Bogotá:** se asume Bogotá D.C como región, agrupando a la Alcaldía Mayor de Bogotá y las Secretarías Distritales.
- - **Atlántica e Insular:** las gobernaciones de Atlántico, Bolívar, Cesar, Córdoba, La Guajira, Magdalena, Sucre y San Andrés, Providencia y Santa Catalina (8 departamentos) y las alcaldías de Barranquilla, Cartagena, Valledupar, Montería, Riohacha, Santa Marta y Sincelejo (7 alcaldías).
- - **Oriental:** Boyacá, Cundinamarca, Meta, Norte de Santander y Santander (5 departamentos, excluye Bogotá D.C) y las alcaldías de Tunja, Villavicencio, Cúcuta y Bucaramanga (4 alcaldías).
- - **Central:** Antioquia, Caldas, Caquetá, Huila, Quindío, Risaralda y Tolima (7 departamentos) y las alcaldías de Medellín, Manizales, Florencia, Neiva, Armenia, Pereira e Ibagué (7 alcaldías).
- - **Pacífica:** Cauca, Chocó, Nariño y Valle del Cauca (4 departamentos) y las alcaldías de Popayán, Quibdó, Pasto y Cali (4 alcaldías).
- - **Amazonía / Orinoquía:** Arauca, Casanare, Vichada, Putumayo, Amazonas, Guainía, Guaviare y Vaupés (8 departamentos). Se publica como “archivos de uso público” (información muy anonimizada; menor nivel de seguridad) y las alcaldías de Arauca, Yopal, Puerto Carreño, Mocoa, Leticia, Inírida, San José del Guaviare y Mitú (8 alcaldías).
- La información marcada como EDID_2021_ENTIDAD contiene los datos desagregados a nivel de entidad y únicamente se entrega con la variable de identificación SEXO. Se publica como “archivos de datos con licencia de acceso” (información menos anonimizada; mayor nivel de seguridad).

La documentación relacionada está disponible libremente para consulta de los usuarios.

Se autoriza el uso de la información contenida en este portal, siempre y cuando se haga la siguiente cita textual:

"Fuente: Departamento Administrativo Nacional de Estadística: www.dane.gov.co . Queda en cambio prohibida la copia o reproducción de los datos en cualquier medio electrónico (redes, bases de datos, cd rom, diskettes) que permita la disponibilidad de esta información a múltiples usuarios sin el previo visto bueno del DANE por medio escrito.

Para los usuarios que requieran consultar los microdatos con las variables del capítulo de ubicación laboral a nivel de entidad, se dispone una base con este conjunto de datos en la sala de procesamiento de datos especializada del DANE, al cual pueden acceder de acuerdo con los términos establecidos por la entidad para consultas a través de este medio."

El DANE no se responsabiliza por el uso e interpretación realizada por terceros.

1.4.1 VARIABLES DE IDENTIFICACIÓN

Variables de identificación: selección muestral y operativo en campo

Para la construcción del marco muestral, a partir del cual se realiza la selección de la muestra, las oficinas de recursos humanos de cada una de las entidades objeto de estudio, suministran la información de contacto de las y los servidores públicos que laboran en la sede principal de la entidad.

A partir de esta información se realiza la selección de las y los servidores públicos encuestados, considerando que estas variables sólo son utilizadas durante la recolección de información para acceder a la entidad y contactar a las y los servidores públicos encuestados. En algunos casos, cuando se realizan verificaciones del operativo de campo, se contacta a los informantes con el propósito de validar la información básica y tener controles operativos sobre la realización de la encuesta.

Las variables de identificación son:

Entidad, área, cédula, nombre, cargo, nivel jerárquico, sexo, municipio, sede, teléfono, correo electrónico.

En cumplimiento del artículo 5 de la Ley 79 de 1993, estas variables son absolutamente confidenciales, están protegidas por la reserva estadística y por ningún motivo se incluye en las bases de datos de acceso público; así como tampoco se encuentran dentro de las bases de datos utilizadas al interior del DANE para la generación de estadística básica.

En consecuencia, todas las variables relacionadas con identificación de las fuentes son eliminadas de las bases de datos.

Variables de identificación Capítulo B identificación laboral

B1 - Codent: Entidad: corresponde al nombre completo de la entidad en la cual trabaja el servidor(a) público. En el formato web esta pregunta se encuentra prediligenciada, mientras que, en el cuestionario físico, debe marcarse durante el proceso de recolección.

A cada una de las 201 entidades objeto de estudio, se le asigna un código número de identificación. (Esta variable solo está disponible en la base EDID_2021_ENTIDAD)

B2 Cargo que desempeña dentro de la entidad: la investigación establece de manera general para todas las entidades objeto de estudio, tres niveles jerárquicos para clasificar las y los servidores públicos que diligencian la encuesta, de la siguiente manera:

Nivel 1: directivo/ asesor

Nivel 2: profesional/técnico

Nivel 3: asistencial

A partir de 2016, los niveles jerárquicos se presentan en la base de datos de manera desagregada en cinco categorías de la siguiente manera:

- Directivo
- Asesor
- Profesional
- Técnico
- Asistencial

(Esta variable solo está disponible en la base EDID_2021_REGION y en la base dispuesta en sala de procesamiento especializado)

B3 Tiempo de servicio en la entidad: a partir de esta pregunta las y los servidores inician el diligenciamiento. Esta pregunta permite conocer el tiempo de permanencia del encuestado(a) en la entidad y, por ende, tener un indicio sobre su experiencia acerca de ella.

(Esta variable solo está disponible en la base EDID_2021_REGION y en la base dispuesta en sala de procesamiento especializado)

B4 Sexo: esta pregunta determina el sexo del encuestado(a). El encuestado(a) solo puede seleccionar una opción de respuesta.

B5 Nivel educativo: esta pregunta determina el nivel educativo más alto alcanzado por el encuestado(a).

Esta variable se presenta con los siguientes niveles de agregación:

1. 0 a Media: incluye los niveles ninguno, preescolar, básica primaria (1° -5°), básica secundaria (6° - 9°) y media (10° - 13°).
2. Técnico: incluye los niveles técnico sin título y técnico con título.
3. Universitario: incluye los niveles universitario sin título y universitario con Título
4. Posgrado: incluye los niveles posgrado sin título y posgrado con título.

Pregunta B6. Edad: esta pregunta determina la edad del encuestado(a). La pregunta hace explícita la solicitud de informar sobre los años cumplidos al momento de diligenciar la encuesta.

Para garantizar la confidencialidad de las fuentes, esta variable se presenta agregada en los siguientes rangos:

1. 18 a 28 años.
2. 29 a 40 años.
3. 41 a 50 años.
4. 51 a 60 años.
5. 61 años y más.

Variables de identificación EDID_2021_REGION

Para prevenir la posible identificación de algunas de las fuentes, de este conjunto de variables se suprime la variable que identifica a cada una de las entidades. Sin embargo, con propósitos de consulta se crearon una nueva variables: 1) Región (REGIÓN) en la cual se agrupan las entidades participantes de acuerdo las regiones definidas por el DANE.

La codificación de estas variables se encuentra en el archivo Diccionario de base de datos EDID 2021.

Variables de identificación EDID_2021_ENTIDAD

En caso de que la solicitud de información requiera un nivel de desagregación por entidad, es posible acceder a una base anonimizada adicional, en la cual, para prevenir la posible identificación de algunas de las fuentes, se suprimen las variables que identifican: el cargo que desempeña en la entidad (B2), el tiempo de servicio en la entidad (B3), el nivel educativo (B5), edad (B6) y funciones de jefatura o supervisión de personal (B07). Las únicas variables de identificación que se entregan para consulta con este nivel de desagregación son entidad (CODENT) y sexo (B4).

1.4.2 VARIABLES OBSERVADAS

Al tratarse de una encuesta de percepción con respuestas cerradas es posible presentar los resultados para cada una de las variables de la encuesta abarcando el contenido del cuestionario desde el Módulo II -Capítulo C – C1a hasta el Módulo III – Capítulo H –H6. De esta manera no se realizan exclusiones a este conjunto de variables.

1.4.3 VARIABLES DE PROCESO

Este conjunto de variables identifica los procesos que sigue la información una vez se ha recolectado, el objetivo principal de este componente es preparar la información insumo para que pueda emplearse en el cálculo de los resultados y agregaciones. Así dentro de este componente se incluye los códigos por región (REGIÓN) y el factor de expansión final (FEXPFINAL).

1.5 DISPOSICION DE DATOS

Una vez surtido el proceso mencionado anteriormente, por el cual se preserva la reserva estadística, para la base de datos considerada a partir de la información insumo de la EDID 2021 que serán de consulta para uso público, se cuenta con un anexo que contiene el diccionario de variables que se tienen disponibles en la base de datos en cuestión con lo cual se facilite su uso al público en general, adicionalmente se cuenta con la ficha metodológica que describe las características de la operación estadística.

1.6 APLICACIÓN DE LINEAMIENTOS PARA ANONIMIZACIÓN

La Oficina de Sistemas será la responsable de ejecutar las indicaciones dadas en el documento para la generación de las bases de datos anonimizadas que se dispondrán al público mediante consulta en sitio especializado en el DANE.

1.7 DICCIONARIO BASE DE DATOS

En el archivo Diccionario de base de datos EDID 2021 se incluye la información correspondiente a las variables, características y valores incluidos en la base de datos para uso público.