



PLAN ESTADÍSTICO DISTRITAL (PED)

Lineamientos para la anonimización de los microdatos que se producen y difunden en el Distrito



ALCALDÍA MAYOR
DE BOGOTÁ D.C.

SECRETARÍA DE
PLANEACIÓN



Alcaldía Mayor de Bogotá

Claudia Nayibe López Hernández

Secretaría Distrital de Planeación

María Mercedes Jaramillo Garcés

Subsecretaría de Información

Jorge Alberto Torres Vallejo

Dirección de Información y Estadísticas

Daniela Perez Otavo

EQUIPO TÉCNICO

María Teresa Rojas Linares

Elías José Mejía Fuentes

Plan Estadístico Distrital (PED)

**Lineamientos para la anonimización de los microdatos
que se producen y difunden en el Distrito**

Bogotá D.C., noviembre de 2022

INTRODUCCIÓN

La Constitución Política de Colombia establece en su Artículo 15: *"Todas las personas tienen derecho a su intimidad personal y familiar y a su buen nombre, y el Estado debe respetarlos y hacerlos respetar. En la recolección, tratamiento y circulación de datos se respetarán la libertad y demás garantías consagradas en la Constitución"*. En este sentido, la Ley Estatutaria 1266 de 2008 establece el Habeas Data que regula el manejo de la información contenida en bases de datos y establece disposiciones sobre la recolección, tratamiento y difusión de los datos personales. De la misma manera, la Ley 1712 de 2014, reglamenta el derecho de acceso a la información pública a través de una política de datos abiertos, propiciando la apertura y reutilización de los datos.

En materia de información estadística, el Código Nacional de Buenas Prácticas para las Estadísticas Oficiales¹, en el principio 5: Confidencialidad, se indica: "Las entidades pertenecientes al Sistema Estadístico Nacional -SEN- deben garantizar la protección y la confidencialidad de la información con la que se producen las estadísticas oficiales, así como evitar la identificación de las fuentes". Enfatiza la necesidad de contar con protocolos para la protección y seguridad de la información y para anonimizar los datos, para que se pueda dar acceso a microdatos anonimizados en donde se garantice la confidencialidad.

En este marco, con el objetivo de orientar a las entidades distritales productoras de información estadística sobre las técnicas de anonimización y los lineamientos para dar acceso a los usuarios de los microdatos, con la garantía de la protección y la confidencialidad de la información, en desarrollo de la estrategia relacionada con la promoción de la difusión y acceso de la información estadística del Plan Estadístico Distrital (PED) 2019-2024, la Dirección de Información y Estadística (DIE) de la Secretaría Distrital de Planeación (SDP) pone a disposición los lineamientos para la anonimización de los microdatos que se producen y difunden en el distrito, mediante operaciones estadísticas por muestreo, censos, registros administrativos, estadísticas derivadas cuya responsabilidad pertenece a cada entidad del Distrito.

Es importante indicar que las actividades de anonimización hacen parte del proceso de producción de información estadística y por lo tanto deben diseñarse y documentarse desde los análisis de riesgos, la aplicación de las técnicas adecuadas, hasta la evaluación de los resultados para la difusión de las bases anonimizadas.

¹ DANE. Código Nacional de Buenas Prácticas para las Estadísticas Oficiales. En: https://www.dane.gov.co/files/sen/bp/Codigo_naL_buenas_practicas.pdf

1. Algunos conceptos relacionados

Anonimización de microdatos: "proceso técnico que consiste en transformar los datos individuales de las unidades de observación, de tal modo que no sea posible identificar sujetos o características individuales de la fuente de información, preservando así las propiedades estadísticas en los resultados". (Decreto 1743 de 2016: Art. 2.2.3.1.1).

Microdato: "cada uno de los datos sobre las características de las unidades de estudio de una población (individuos, hogares, establecimientos, entre otras) que se encuentran consolidados en una base de datos" (Decreto 1743 de 2016: Art. 2.2.3.1.1).

Base de datos: "conjunto o colección de datos interrelacionados entre sí, que se utilizan para la obtención de información de acuerdo con el contexto de estos y que son almacenados sistemáticamente para su posterior uso". (Sistema de Consultas de Conceptos del DANE, 2018).

Datos sensibles: "aquellos que afectan la intimidad del Titular o cuyo uso indebido puede generar su discriminación, tales como aquellos que revelen el origen racial o étnico, la orientación política, las convicciones religiosas o filosóficas, la pertenencia a sindicatos, organizaciones sociales, de derechos humanos o que promueva intereses de cualquier partido político o que garanticen los derechos y garantías de partidos políticos de oposición así como los datos relativos a la salud, a la vida sexual y los datos biométricos". (Ley Estatutaria 1581 de 2012. Artículo 5°).

Datos personales: "información asociada a una persona y que permite su identificación. Por ejemplo: nombres, documento de identidad, lugar de nacimiento, estado civil, edad, lugar de residencia, trayectoria académica, laboral, o profesional". (Superintendencia de Industria y Comercio en: <https://www.sic.gov.co/sobre-la-proteccion-de-datos-personales>). También se le conoce como variables de identificación directa.

Pseudoidentificadores: son aquellas variables temáticas que pueden convertirse en identificadoras indirectas cuando se establecen relaciones funcionales entre ellas y pueden conducir a identificar algunas unidades de observación. Es el caso de la relación que puede existir entre las variables escolaridad (posgrado), con pertenencia a un grupo étnico en un municipio de Ciudad región Bogotá, que al aplicar procesos estadísticos como tablas de frecuencias que crucen las variables se presenten pocos casos que lleven a identificar a la fuente.

Escenario de riesgo: consiste en el análisis e identificación de las variables sensibles, de los datos personales y de aquella información confidencial que tiene posibilidad de ser individualizada a partir de las características de una variable o de la combinación de varias variables de la base de datos.

2. Propiedades y métodos de anonimización

Hace referencia a las características que deben conservar las bases de datos que han sido anonimizadas después de aplicar los métodos que modifican los datos, con el objeto de preservar la confidencialidad de la unidad de observación y garantizar el comportamiento, tendencia, nivel de desagregación y las relaciones entre las variables, para que estos puedan ser reutilizados.

2.1 Propiedades de la base de datos anonimizada

En concordancia con lo expuesto por el DANE *“el equipo de trabajo deberá establecer las propiedades estadísticas que se deben mantener en la base de datos anonimizada”*. Los responsables de aplicar las técnicas de anonimización de las bases de datos, deben garantizar que se conserven las siguientes características, para que de esta forma puedan ser reutilizadas por los usuarios:

- La tendencia o comportamiento en periodos de tiempo. Por ejemplo, si la base de datos de la variable *“evasión del Impuesto de Industria y Comercio ICA”* ha venido disminuyendo anualmente, se debe conservar esta tendencia.
- Las medidas estadísticas de variables definidas como “significativas”. Por ejemplo, los temáticos especifican que el atributo *“promedio del subsidio asignado por hogar”* a nivel de UPZ se debe conservar o máximo que tenga una variación del 10%, salvo que suceda un evento que explique una mayor variación (caso de su aumento por COVID).
- Las cifras para los niveles de desagregación geográfica o temática para garantizar a los usuarios los análisis a los niveles requeridos. Por ejemplo, el equipo requiere que se continúe presentando información sobre la variable *pertenencia a un grupo étnico* para cada categoría a nivel localidad.
- Las posibles relaciones existentes entre las variables. *Por ejemplo*, los temáticos han definido que los coeficientes de correlación entre las variables *nivel de concentración de sustancias contaminantes y las mediciones sobre la calidad del aire*, se mantengan con el fin de no distorsionar los análisis.

² DANE. Guía para la anonimización de bases de datos en el Sistema Estadístico Nacional. 2018. Página 24. En: http://www.dane.gov.co/_les/sen/lineamientos/DSO-020-LIN-08.pdf

2.2 Técnicas de anonimización

Se refieren a procedimientos que modifican en forma sistemática los datos, de modo que se minimiza el riesgo de identificar las unidades de observación. Consiste en incluir, suprimir y/o modificar algunos datos que se considera pueden llevar a deducir la persona, establecimiento o en general la unidad de observación. Con la aplicación de estas técnicas se producen supresiones parciales o totales en los resultados de los diferentes tipos de variables y reducen el nivel de detalle de la base de datos. Las técnicas de anonimización se dividen en: i) apoyadas en la no perturbación de datos. ii) apoyadas en la perturbación de datos.

2.2.1 Técnicas basadas en la no perturbación de datos

Estas técnicas utilizan supresiones parciales, reducción o recodificación de la información para minimizar el riesgo de identificación de las unidades de observación. Este tipo de técnicas son utilizadas para evitar que las variables de identificación y localización y los datos atípicos faciliten la identificación de las personas, establecimientos u otra unidad de observación. Dentro de este tipo de técnicas se encuentra:

- **Eliminación de variables:**

Consiste en eliminar una variable con toda su información, debido a que contiene información sensible o datos personales y con ella se puede identificar de forma directa la unidad de observación. Por ejemplo, variables de nombres, documento de identidad, correo, dirección, teléfono y fotografías, entre otras, que pueden dar cuenta de identificación y/ o ubicaciones.

- **Recodificación:**

Radica en combinar varias categorías por una categoría más general que contenga un mayor número de casos. Es decir, se recodifican los valores de la variable para tener una nueva categoría. Puede utilizarse en variables cuantitativas o en categóricas. Por ejemplo, en el caso de la variable "Titulo otorgado en doctorado" puede ser recodificado y agrupado en nuevas categorías: "Doctorado por áreas de conocimiento". Para variables cuantitativas se agrupan por medio de intervalos, es el caso de agrupar la variable "ingresos".

³ Las variables pueden ser cuantitativas (continuas o discretas) o cualitativas. Las variables cuantitativas se expresan mediante un número, en tanto que las cualitativas expresan características o cualidades, y no pueden ser medidas con números. Una variable categórica es aquella que permite clasificar o agrupar una serie de datos por medio de valores fijos asociados a una cualidad o categoría concreta.

⁴ Para mayor detalle se puede consultar: Hundepool et al.(2012). En: https://www-researchgate-net.translate.googleusercontent.com/publication/280156259_Statistical_Disclosure_Control

- **Codificación por externo superior y/o inferior:**

Consiste en asignar una codificación a las unidades de observación por la facilidad de identificarlas al tener en una variable valores extremos máximos o mínimos, para un nivel de desagregación (geográfico o temático). Por ejemplo, en una localidad de Bogotá es reducido el número de supermercados "grandes" y pueden ser fácilmente identificados por la variable "ventas mensuales", al presentar valores externos altos, por lo que se utiliza una codificación. Por el número reducido de ellos que tienen presencia; por lo que se asigna un código a estos supermercados.

- **Supresión local:**

En esta técnica se eliminan los valores de una variable y se reemplazan por valores "faltantes". Suelen identificarse las unidades de observación por relación, combinación o cruce de variables en tablas de contingencia con frecuencia de ocurrencia baja, que hace que estas variables sean consideradas riesgosas y pueda llegar a identificaciones. Se aplica para variables categóricas. Por ejemplo, al cruzar las variables "orientación sexual y "discapacidad", se obtiene una frecuencia de pocos individuos para alguna categoría, lo que hace fácil su identificación, por lo que se reemplazan estas categorías como "datos faltantes".

2.2.2 Técnicas basadas en la perturbación de datos

La aplicación de este tipo de técnicas significa transformar los datos de las variables consideradas sensibles, implica: incluir, modificar o eliminar los datos existentes, para lograr que las unidades de observación no sean identificables. Dentro de este tipo de técnicas se encuentra⁵:

- **Microagregación:**

Con esta técnica se reemplazan algunos datos de una variable, por el promedio de ese subconjunto de datos, debido a que, por presentar valores extremos o únicos para un nivel de desagregación geográfica o temática, resulta fácil la identificación de las unidades de observación. Es aplicable a datos cuantitativos. Por ejemplo: para una localidad de Bogotá, se encuentran varios predios de tamaño significativamente alto respecto a los de la misma localidad, por tanto, se agrupan y se reemplazan los valores de cada uno de ellos por el promedio (mts²).

⁵ Templ, Meindl, Kowarik y Chen. 2014. En:

<https://www.readkong.com/page/introduction-to-statistical-disclosure-control-sdc-8459007>

• Adición de ruido:

Consiste en modificar los valores de una variable sensible a partir de la suma o multiplicación de números aleatorios, lo que proporciona un conjunto de datos menos exactos, pero con la condición de que se conserven las mismas características estadísticas. Por ejemplo, en un registro administrativo que tiene variables que caracteriza los funcionarios del Distrito, una variable considerada sensible es el ingreso mensual, por lo que se decide generar una serie de números aleatorios para adicionar a cada uno de los valores de esta variable, a fin de presentar los valores con menor exactitud.

• Intercambio de datos:

Con esta técnica se perturban los datos a partir del intercambio aleatorio de los valores de un registro en una variable identificada con riesgo, por otro registro de la misma variable (considerada sin riesgo), de modo que se conserva el nivel de detalle y la distribución de los valores. Puede utilizarse en variables continuas o en categóricas.

• Redondeo:

Consiste en sustituir el valor de las variables que tienen valores decimales por valores redondeados, en aquellas unidades de observación considerada con riesgo. Se aplica en variables cuantitativas continuas. Suele aplicarse esta técnica luego de realizar la microagregación para las variables que no admiten valores con decimales. Es el caso del número de hijos que han sido agrupados por grupos quinquenales.

• Reajuste de los factores de expansión:

Para las operaciones estadísticas por muestreo, puede darse el caso que usuarios especializados, al conocer el diseño muestral, podrían revertir el proceso para identificar las unidades de observación; por tanto, la técnica consiste en modificar los pesos. En otro caso, se podrían excluir los factores de expansión en cada una de las etapas del diseño muestral, para presentar solo el factor de expansión consolidado.

Otra de las técnicas poco utilizadas en los procesos de anonimización de bases de datos, son los métodos de datos sintéticos, que no son considerados los más adecuados, ya que modifican todos los registros en la base de datos, aunque conservan las mismas tendencias y correlaciones. La técnica consiste en poner a disposición del usuario una base con datos simulados, a partir del diseño de algoritmos de simulación que utiliza modelos de regresión cuantílica, imputación adicional y datos combinados⁶.

⁶ Handbook on Statistical Disclosure Control. A Hundepool, J Domingo-Ferrer, L Franconi. - ESSnet. 2010. Eurostat, Centros de Excelencia (CENEX).

3. Proceso de anonimización de bases de datos

El proceso de anonimización de una base de datos comprende las siguientes etapas:

- 1) Requerimientos.
- 2) Análisis de riesgos de identificación de las unidades de observación.
- 3) Selección de las técnicas de anonimización más adecuadas.
- 4) Análisis de viabilidad.
- 5) Aplicación de las técnicas de anonimización seleccionadas.
- 6) Evaluación de resultados.

Etapa 1. Requerimientos

En esta etapa se constituye el equipo de trabajo; se dispone de la base de datos a anonimizar y los instrumentos que permiten interpretarla; se determina la infraestructura tecnológica a utilizar y se determina la seguridad y los medios para disponer la base de datos anonimizada.

Cuidando que se garantice el anonimato de las unidades de observación, el equipo de trabajo debe documentar el proceso, las técnicas aplicadas, el concepto de viabilidad y las limitaciones de uso de la base de datos anonimizada, en el marco de las evidencias de la aplicación de las fases del proceso estadístico en las etapas de diseño y difusión.

En cuanto al equipo de trabajo se recomienda este compuesto por:

- 1) Profesionales especializados que conozcan la temática de la operación estadística o del registro administrativo, ya que sobre ellos recae la identificación de las variables sensibles y con riesgo y, la determinación de la viabilidad de la anonimización.
- 2) Profesionales que manejen paquetes estadísticos, como R, SAS, SPSS, Stata, para que realicen los análisis exploratorios de las bases de datos y la aplicación de las técnicas de anonimización.

Con respecto a los responsables de la operación estadística o del registro administrativo, es fundamental que proporcionen la base de datos a anonimizar en su última versión y con la verificación de la consistencia y coherencia; debe proporcionarla al equipo de trabajo sin restricciones, pero con protocolos de acceso o acuerdos de confidencialidad. Así mismo, deben proporcionar el modelo entidad- relación⁷ y el respectivo Diccionario de datos⁸.

⁷ El modelo entidad - relación es un modelo conceptual para bases de datos relacionales que permite representar cualquier abstracción, percepción y conocimiento en un sistema de información formado por un conjunto de objetos denominados entidades y relaciones, a partir de una representación gráfica llamada diagrama entidad-relación, adoptando el enfoque más natural del mundo real. Página 5. En: <https://www.dane.gov.co/files/sen/lineamientos/Recomendaciones-para-elaborar-modelos-entidad-relacion.pdf>

⁸ En el Instructivo para la creación de Diccionario de Datos. IDECA. 2019, se presentan los elementos necesarios para definir los datos básicos del diccionario de datos. Se indica: El diccionario es una herramienta que contribuye al fortalecimiento de la gestión de datos, en la cual se encuentran las definiciones de los datos y sus propiedades (atributos, dominios, asociaciones y operaciones). Así mismo, permite identificar claramente el propósito, alcance y campo de aplicación, sirviendo a los usuarios como orientación del tipo de información que encontrará en su contenido. Página 9. En: <https://www.ideca.gov.co/sites/default/files/documentacion/instructivodd.pdf>

Por su parte, la entidad a realizar el proceso de anonimización debe disponer de los paquetes estadísticos con licencias, equipos con capacidad para el manejo de grandes volúmenes de información y tendrá definida la Infraestructura, las claves de acceso y la administración de la base anonimizada junto con las notas metodológicas de uso que se requieran. En este sentido, es aconsejable para cuando se carece de estos requerimientos, establecer acuerdos o convenios con entidades o dependencias que cuenten con los recursos y la infraestructura requerida.

Etapa 2. Análisis de riesgos

En esta etapa el equipo de trabajo realiza el análisis de riesgos, para poder identificar las unidades de observación que deben ser protegidas por el tipo o combinación de información que contienen (datos personales, datos sensibles, valores únicos, extremos, atípicos, valores combinados con poca participación), para ello:

- Identifica, caracteriza y evalúa, a partir de la base de datos y del diccionario de datos, las variables o registros que presentan riesgos, debido a que contienen información de la unidad de observación: **1)** con datos sensibles que afectan la intimidad de las personas y pueden causar discriminación; **2)** con datos personales que permite su identificación o ubicación; iii) con variables temáticas (pseudoidentificadores) que al combinarlas pueden ser riesgosas, por presentar baja frecuencia.
- Define para las variables con identificadores directos: i) las que deben ser eliminadas de la base de datos; ii) las que pueden ser agrupadas.
- Establece y realiza posibles combinaciones entre las variables directas y las variables temáticas para diferentes niveles de desagregación, a fin de verificar las unidades de observación susceptibles de identificar por los usuarios y que requieren ser protegidas aplicando las técnicas de anonimización.
- Define los niveles de desagregación temática y geográfica a los cuales se puede presentar la información sin que presente riesgos de identificación y sobre los cuales se deben conservar las propiedades globales.
- Define las variables "importantes" que deben conservar las propiedades estadísticas (porcentajes, promedios, tasas, porcentajes de variación, distribuciones) para desagregaciones en las bases de datos anonimizadas.
- Calcula y analiza la distribución de frecuencias para las variables categóricas consideradas "sensibles", para identificar las clases que presentan poca participación para un nivel de desagregación temático o geográfico y que pueden ser consideradas riesgosas.
- Propone recodificaciones de las variables con frecuencias bajas y analiza la utilidad para los usuarios de las bases de datos agrupadas a estos niveles.
- Calcula y analiza las medidas estadísticas descriptivas y de tendencias para las variables

cuantitativas consideradas "importantes" por los temáticos, como el promedio, varianza y cuartiles a fin de identificar las variables que presentan riesgo.

- Identifica bases de datos externas, que al combinar algunas de sus variables, con las variables de la base de datos a anonimizar, podrían conducir a un riesgo.
- Identifica y lista las unidades de observación con riesgo de identificación a partir de la realización de las anteriores actividades y cálculos.

Etapas 3. Selección de las técnicas de anonimización

En esta etapa el equipo evalúa y define las técnicas de anonimización que van a aplicar sobre las variables y registros identificados como riesgosos, para la selección debe tener en cuenta:

- El tipo de variable: categórica o cuantitativa.
- Que la técnica minimice la ocurrencia de los riesgos identificados.
- Que el procedimiento minimice la pérdida de información o de perturbación de los datos.
- Que se conserven las propiedades estadísticas.
- Que se respeten los resultados globales a los niveles de desagregación requeridos.
- Que no se pierda la utilidad de reutilizar de la información.

Planteadas las técnicas de anonimización, el equipo realiza el informe de anonimización que contiene:

- La categorización de los riesgos que se presentan con las variables involucradas.
- Las variables con identificadores directos que deben ser eliminadas.
- Las variables y los registros considerados riesgosos porque tienen categorías con bajo porcentaje a un nivel requerido.
- Las variables consideradas pseudoidentificadoras que al combinarlas con otras son consideradas riesgosas.
- La descripción del riesgo que se ocasiona sobre los registros en las variables.
- El porcentaje de registros que se ven involucrados en cada riesgo.
- La técnica de anonimización a aplicar.

Etapa 4. Análisis de viabilidad

A partir de los insumos anteriores el equipo elabora la propuesta de viabilidad de la anonimización, para lo cual analiza y determina que se cumplan los siguientes aspectos a fin de emitir un concepto favorable:

- Que las necesidades de los usuarios se puedan cumplir con la base de datos anonimizada a entregar.
- Que no existan regulaciones normativas, políticas en la entidad o aspectos temáticos sobre la entrega de la base de datos.
- Que las técnicas de anonimización a aplicar no permitan identificar unidades de observación.
- Que las propiedades estadísticas se conserven a los niveles de desagregación temáticas y geográficas requeridos por los temáticos.

Etapa 5. Aplicación de las técnicas de anonimización

Definidas las técnicas de anonimización que serán utilizadas en el proceso, se debe proceder al proceso de programación en el lenguaje informático que disponga la entidad. Se deben tener en cuenta los siguientes aspectos para la aplicación e implementación de las técnicas.

- Crear una base espejo en la cual se realizarán las distintas operaciones que sean necesarias para la aplicación e implementación de las técnicas seleccionadas.
- Resaltar las variables a las cuales se les aplicaran las técnicas de anonimización, con el fin de poder tener la trazabilidad y facilitar la etapa de evaluación y verificación de resultados.
- Definir un orden en las variables de acuerdo con el grado de importancia e impacto en la operación estadística o registro administrativo para la implementación de las técnicas, que le permitan al equipo llevar el control del proceso e ir validando los resultados obtenidos.
- Identificar y determinar las unidades de observación (registros) que serán de apoyo para el enmascaramiento y protección de las unidades de observación que han sido identificadas como riesgosas; con el propósito de perturbar lo menos posible la base de datos para garantizar su mayor grado de utilidad.

Etapa 6. Evaluación de resultados

La etapa del proceso de evaluación de la información es fundamental en el proceso, ya que se define la efectividad de las técnicas empleadas para garantizar la protección y confidencialidad de las fuentes de observación (personas, empresas, viviendas, hogares. etc).

Se deben tener en cuenta los siguientes aspectos en el proceso de evaluación de resultados:

- Se deben calcular nuevamente los escenarios de riesgos definidos en la etapa 2, con el fin de verificar si los registros que se habían identificado no persistan como riesgosos y la aparición de nuevos registros.
- Validar que las propiedades estadísticas de la base original se conserven, es decir se debe verificar que los totales, promedios, desviaciones de las variables sean iguales a los obtenidos en la etapa de revisión; en caso que no se cumplan estas propiedades se debe revisar las técnicas empleadas y las operaciones realizadas para garantizar el éxito del proceso.
- Si se identificaron nuevos registros riesgosos después de implementadas las técnicas se debe realizar el proceso de anonimización a estos registros y validar nuevamente que sean debidamente protegidos.
- Suplidos todos los pasos y verificadas las propiedades de la base de datos se puede disponer por los medios de difusión definidos por cada entidad para el acceso a los microdatos de las distintas operaciones estadísticas.

Bibliografía

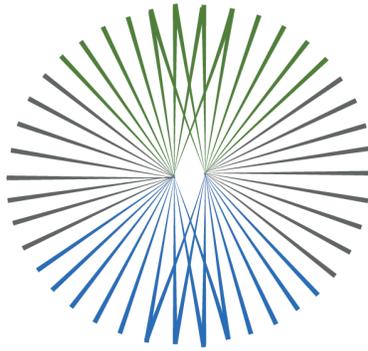
DANE. Código Nacional de Buenas Prácticas para las Estadísticas Oficiales. En: https://www.dane.gov.co/files/sen/bp/Codigo_nal_buenas_practicas.pdf.

DANE. DANE. Lineamientos para la Anonimización de Microdatos [Internet]. 2014. Recuperado a partir de: http://www.dane.gov.co/_les/sen/lineamientos/DSO-020-LIN-08.pdf.

Eurostat, Centros de Excelencia (CENEX). 2010. Handbook on Statistical Disclosure Control. A Hundepool, J Domingo-Ferrer , L Franconi... - ESSnet.

Ministerio de Salud. (s.f.). Lineamientos para la Anonimización de Datos del Sistema Nacional de Estudios y Encuestas Poblacionales para la Salud. Disponible en: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/GCFI/lineamientos-anonimizacion-sistema-encuestas.pdf>.

Templ, Meindl, Kowarik y Chen. 2014. En: <https://www.readkong.com/page/introduction-to-statistical-disclosure-control-sdc-8459007>



PED

Plan Estadístico
Distrital



ALCALDÍA MAYOR
DE BOGOTÁ D.C.

SECRETARÍA DE
PLANEACIÓN

