



---

# **Guía de indicadores para la evaluación de la calidad de los registros administrativos**

---

**SUBDEPARTAMENTO DE CALIDAD Y ESTÁNDARES  
DEPARTAMENTO DE METODOLOGÍAS E INNOVACIÓN ESTADÍSTICA**

**INSTITUTO NACIONAL DE ESTADÍSTICAS  
Marzo/2024**

**SUBDEPARTAMENTO DE CALIDAD Y ESTÁNDARES**  
**DEPARTAMENTO DE METODOLOGÍAS E INNOVACIÓN ESTADÍSTICA**  
**DIRECCIÓN NACIONAL**

---

Instituto Nacional de Estadísticas

Marzo/2024

## ÍNDICE

ANTECEDENTES.....	1
1. CONSIDERACIONES METODOLÓGICAS .....	2
1.1. Aseguramiento de la calidad estadística en el uso de registros administrativos .....	2
1.2. Ciclo de evaluación de la calidad del registro .....	4
1.3. Alcances de los indicadores .....	5
1.4. Formulación de los indicadores.....	6
1.5. Conceptos relevantes.....	6
2. BATERÍA DE INDICADORES .....	8
2.1 Retraso en la entrega/disponibilidad de los RRAA para su aprovechamiento estadístico ...	8
2.2 Actualidad del registro .....	9
2.3 Cumplimiento en la entrega de metadatos del registro .....	10
2.4 Porcentaje de errores por variable del conjunto de datos .....	11
2.5 Porcentaje de variables con errores en la base de datos.....	12
2.6 Porcentaje de observaciones duplicadas .....	12
2.7 Tasa de completitud de las variables .....	13
2.8 Tasa de subcobertura .....	14
2.9 Tasa de sobrecobertura.....	15
2.10 Tasa de edición .....	16
2.11 Tasa de imputación.....	17
2.12 Tasa de codificación.....	18
2.13 Tasa de unidades no vinculables .....	19
2.14 Tasa de vinculación.....	20
2.15 Tasa de robustez de valores atípicos.....	21
2.16 Utilización de métodos de Anonimización .....	21
2.17 Grado de coherencia con otras Fuentes.....	22
2.18 Tiempo o lapso de desfase en la entrega de resultados finales.....	23
2.19 Puntualidad en la publicación de resultados.....	24
3. GLOSARIO.....	25
4. REFERENCIAS.....	28

## HISTORIAL DE VERSIONES

<b>Versión</b>	<b>Fecha</b>	<b>Unidad Responsable</b>	<b>Descripción de cambios</b>
1.0	12-03-2024	Departamento de Metodologías e Innovación Estadística - Subdepartamento de Calidad y Estándares - Sección de Calidad Estadística.	No aplica, primera versión oficial.
1.1	13-03-2024	Departamento de Metodologías e Innovación Estadística - Subdepartamento de Calidad y Estándares - Sección de Calidad Estadística.	Se rectifica la tabla 1.

## ANTECEDENTES

El Subdepartamento de Calidad y Estándares (SCE) tiene dentro de sus funciones asegurar la implementación y cumplimiento de estándares y normas metodológicas, así como también velar por la evaluación de la calidad en la producción de estadísticas de la Institución. De acuerdo con esto, el SCE ha trazado dentro de sus objetivos de corto y mediano plazo aportar a la evaluación de la calidad de los registros administrativos, entregando orientaciones metodológicas en su uso y elaborando herramientas concretas para la evaluación de la calidad de estos, alineado a los objetivos estratégicos de la institución para el periodo 2022-2026<sup>1</sup>.

La relevancia de evaluar la calidad de los registros administrativos obedece a su amplio e intensivo uso como fuente de información estadística en la actualidad, siendo considerada como una fuente de información para la generación de estadísticas oficiales en complemento a encuestas, censos y otros tipos de procesos estadísticos. No obstante, la incorporación y aprovechamiento estadístico de nuevas fuentes de información, como los registros administrativos, en las estadísticas oficiales, deben someterse a un proceso de análisis y tratamiento que permita resguardar criterios mínimos de calidad, aun cuando se hayan capturado con diversos controles y validaciones de carácter administrativo.

En ese contexto, organismos internacionales y oficinas estadísticas han desarrollado y fortalecido metodológicamente las recomendaciones para el uso de registros administrativos en la producción de estadísticas oficiales, promoviendo herramientas concretas para la evaluación de la calidad, dentro de las que destacan los cuestionarios de autoevaluación e indicadores de calidad para medir distintos aspectos de la calidad de los registros. En particular, este documento expone una batería de indicadores que permite evaluar la calidad de los registros administrativos que utilizará la institución y otros organismos públicos que lo requieran, orientada a ser aplicada durante las etapas de procesamiento y análisis de cada registro para permitir su adecuado uso estadístico.

El trabajo de investigación por parte del SCE comenzó el 2020, instancia en la cual se revisaron experiencias y recomendaciones internacionales en materia de indicadores para la evaluación de los registros administrativos, lo que dio como resultado una propuesta basal, la cual fue revisada e implementada por diversos equipos dentro de la Institución, así como por determinados servicios públicos. Como resultado de ese proceso fue posible identificar fortalezas y oportunidades de mejora, que permitieron construir esta guía, recogiendo las principales necesidades y requerimientos para una adecuada implementación y generando una nueva versión para las y los usuarios que requieran analizar los registros administrativos desde etapas tempranas de evaluación para su uso estadístico.

El documento se estructura en dos partes, primero se entregan consideraciones metodológicas para la utilización de los indicadores, dando cuenta de dos elementos: (i) el marco general que promueve la batería de indicadores y su relación con las buenas prácticas y aseguramiento de la calidad en materia estadística; y, (ii) las etapas de aplicación de los indicadores según su uso estadístico, así como también aquellas actividades a las cuales se pueden asociar en la producción estadística. En la segunda parte, se presenta de forma detallada cada uno de los indicadores con el objeto de entregar la mayor cantidad de orientaciones y lineamientos para su uso e implementación práctica.

---

<sup>1</sup> Para más detalle ver: <https://www.ine.gob.cl/institucional/plan-estrategico-2022-2026/objetivos-y-lineas-estrategicas>

## 1. CONSIDERACIONES METODOLÓGICAS

### 1.1. Aseguramiento de la calidad estadística en el uso de registros administrativos

La batería de indicadores propuestos se ancla en los lineamientos contenidos en el **Marco de Aseguramiento de la Calidad Estadística** (Naciones Unidas, 2019), el **Código Regional de Buenas Prácticas** (CRBP) (Cepal, 2023), así como, a las recomendaciones de indicadores emanadas desde la Oficina Europea de Estadística (Eurostat) y de la Comisión Económica de las Naciones Unidas para Europa (UNECE), entre otros organismos internacionales. En lo particular, la referencia del Marco de Aseguramiento de la Calidad Estadística provee de recomendaciones para asegurar la calidad de las estadísticas oficiales en el **Sistema Estadístico Nacional** (SEN), formado por la **Oficina Nacional de Estadísticas** (ONE) y otros organismos productores del país.

Estas recomendaciones se estructuran a través de una herramienta que estipula que el aseguramiento de la calidad debe aplicarse en cuatro niveles: el **Nivel A** corresponde a la *gestión del sistema estadístico*; el **Nivel B**, a la *gestión del entorno institucional*; el **Nivel C**, a la *gestión del proceso estadístico*; y por último, el **Nivel D**, a la *gestión de los productos estadísticos*. A cada uno de estos niveles se asocian un conjunto de requisitos que deben resguardarse, los cuales corresponden a un listado de acciones y/o buenas prácticas que deben desarrollarse para la generación de estadísticas oficiales para cualquiera de las fuentes de información<sup>2</sup> consideradas (censos, encuestas, estadísticas basadas en registros administrativos, y procesos que integran múltiples fuentes).

Sumado a lo anterior, dentro del esquema de aseguramiento de la calidad se mencionan los principales métodos y herramientas para la evaluación de la calidad, entre estos destacan los indicadores de calidad (tanto para productos como para procesos), informes de calidad, encuestas a usuarios, autoevaluaciones, auditorías externas (revisiones de calidad internas o externas), así como las revisiones por pares.

Respecto a las recomendaciones que se generan para las distintas fuentes de información, se plantean ventajas<sup>3</sup>, consideraciones, limitaciones y aspectos críticos en el uso de fuentes de registros administrativos. Respecto de este último, se incluyen aspectos relacionados con la gobernanza y marco legal de intercambio de registros entre instituciones del sistema estadístico; elementos vinculados a la relevancia y coherencia conceptual entre la información que captura el registro (cobertura temática) y el propósito del estudio; a la integración de datos provenientes de múltiples registros, la combinación entre registros con datos de fuentes estadísticas (encuestas y censos, por ejemplo), entre otros aspectos.

---

<sup>2</sup> “Las fuentes de datos estadísticos incluyen encuestas por muestreo, censos y registros estadísticos. Existen diferentes tipos de censos, como censos de población y vivienda, censos de empresas o económicos, y censos agrícolas, entre otros. Las encuestas por muestreo y los registros estadísticos pueden cubrir diferentes unidades de observación, por ejemplo, individuos/personas, hogares y empresas. Los registros estadísticos pueden provenir de diferentes fuentes, por ejemplo, los registros estadísticos de empresas suelen estar basados en fuentes administrativas” (Naciones Unidas, 2019, pág. 61).

<sup>3</sup> “El uso de fuentes administrativas ofrece muchas ventajas potenciales, como la rentabilidad, la reducción de la carga a los encuestados, mejora de la puntualidad, pertinencia, precisión y fiabilidad debido a su capacidad para obtener datos muy desagregados. En muchos dominios estadísticos, los datos administrativos son indispensables para la producción de Estadísticas” (Naciones Unidas, 2019, pág. 64).

De acuerdo con los lineamientos para el aseguramiento de la calidad de los registros administrativos, se hace especial énfasis en las siguientes problemáticas en cuanto a su uso estadístico, las cuales deben ser consideradas como elementos relevantes para la evaluación de la calidad y en la metodología misma del proceso estadístico (Naciones Unidas, 2019, pág. 64):

- Cooperación insuficiente con los proveedores o administradores de datos, una frecuente falta de claridad con respecto a los roles y responsabilidades de las diferentes partes interesadas, así como desafíos legales para obtener acceso y establecer vínculos de largo plazo con las oficinas administrativas que poseen registros (elementos relacionados al Principio 1: coordinación el sistema nacional de estadística, y el Principio 2: gestión de las relaciones con usuarios de datos, proveedores de datos y otras partes interesadas);
- Uso incoherente o falta de uso de conceptos, definiciones y clasificaciones estandarizadas (elementos relacionados al principio 3: gestión de estándares estadísticos);
- Los conceptos subyacentes de los datos administrativos no reflejan los conceptos utilizados para medir los fenómenos (elementos relacionados al Principio 14: asegurar Relevancia);
- Dificultades metodológicas y técnicas para el acceso a fuentes de datos administrativos, vinculación de registros, integración de datos provenientes de múltiples fuentes, transmisión de datos y combinación con datos de fuentes estadísticas (elementos relacionados al Principio 12: asegurar procedimientos estadísticos apropiados);
- Subcobertura o sobrecobertura de la población objetivo, desalineación con el período de referencia estadístico y/o sesgo inherente (o no cuantificado) como resultado del propósito original del conjunto de datos administrativos (elementos relacionados al Principio 15: asegurar la precisión y confiabilidad).

Por su parte, el CRBP (Cepal, 2023) en concordancia con lo establecido en el **Marco de Aseguramiento de la Calidad Estadística**, incluye en su actualización el fortalecimiento -desde el punto de vista metodológico- del uso de información generada con fines administrativos, y la generación y desarrollo de estándares, mecanismos de coordinación con los proveedores de datos dentro del sistema estadístico. Específicamente en lo que refiere al uso de registros administrativos, se incluye dentro de la actualización del CRBP 2023 un principio nuevo el cual es denominado como “**Aprovechamiento de fuentes secundarias de datos**”, su objetivo es resguardar que la oficina nacional de estadística y los demás miembros del sistema estadístico nacional promueven el aprovechamiento de fuentes secundarias de datos como insumo en la producción de información estadística, para mejorar la eficiencia y reducir la carga del informante. Entre las prácticas que forman parte de este principio, se encuentran:

- La promoción del intercambio de datos y metadatos administrativos, así como de otras fuentes, entre las diferentes áreas de la Oficina Nacional de Estadística y entre los miembros del Sistema Estadístico Nacional. El objetivo es minimizar la carga del encuestado, mejorar la relación costo-eficiencia, incorporar nuevas temáticas y evitar duplicidades en la producción de información estadística.
- La promoción de la mejora continua y la documentación de los registros administrativos por parte de las entidades proveedoras de datos.

- El establecimiento de mecanismos que permitan la transferencia de información asociada a los registros administrativos para que puedan ser empleados en la producción estadística.
- La documentación de los criterios y evaluación de la calidad de los registros administrativos, para su uso con fines estadísticos.

## 1.2. Ciclo de evaluación de la calidad del registro

La batería de indicadores se ancla en un ciclo de evaluación y gestión de la calidad para los registros administrativos. En primer lugar, este ciclo de evaluación ideal consiste en aplicar el **Instrumento de Medición de la Calidad de los Registros Administrativos (IMCRA)** (INE, 2022) por parte de la Oficina Administrativa y la Oficina Estadística, el cual permite iniciar una aproximación a los principales atributos de interés, desde el punto de vista de la calidad, en cuanto a la producción y uso del registro.

El IMCRA permite identificar para quienes utilicen un registro administrativo con fines estadísticos, en las principales características de este, por ejemplo, su población objetivo, cobertura, desagregación de los datos, etc. Su propósito es obtener información precisa sobre la situación de los datos, evaluar la conformidad con los requerimientos estadísticos y el procesamiento de datos. Una vez aplicado el IMCRA y, por ende, identificadas las principales características del registro y oportunidades de mejora para su uso estadístico, se debe aplicar la batería de indicadores de calidad al registro administrativo (o aquellas métricas priorizadas por el área responsable de la producción estadística).

Luego de esta primera etapa de evaluación, una vez aplicada la batería de indicadores definida, es posible que las personas usuarias se encuentren ante dos situaciones: (i) que sea la primera vez que la Oficina Estadística (el INE, por ejemplo) esté utilizando el registro y, por ende, se encuentre en una fase de exploración del conjunto de datos (primera aproximación), en cuyo caso, se debe resguardar que la aplicación de los indicadores sea pertinente a dicha etapa; o (ii) puede ocurrir que el registro ya sea explotado estadísticamente por la Oficina Estadística, en cuyo caso se recomienda aplicar la batería completa de indicadores (los atingentes), dado que, estos cubren todo el proceso de producción estadístico.

Es importante considerar que, para los integrantes del sistema estadístico (ministerios, servicios públicos o municipalidades, por ejemplo) que apliquen los indicadores para evaluar registros administrativos, así como también el IMCRA, tendrán la oportunidad de profundizar en aspectos del registro y, por ende, constituye una línea base para efectuar cambios y/o ajustes en función de los resultados obtenidos. De la misma forma, a través de esta medición se establecen directrices de cómo proyectar la documentación necesaria para todo el ciclo de vida del registro administrativo en cuestión.

Los indicadores de calidad de los registros administrativos se dirigen a evaluar atributos de cada registro y variable de interés; la batería consta de **19 indicadores dispuestos según la etapa en que se analiza y utiliza el registro**, es decir, a la evaluación de aquellos atributos relacionados con su generación en la Oficina Administrativa y los datos disponibles para uso estadístico, y, por otra parte, a la evaluación de los elementos propios del procesamiento y generación de resultados en su uso estadístico. En razón a lo anterior, la primera parte de la batería se enfoca en evaluar la información de entrada, del registro administrativo y de los metadatos que acompañan el conjunto de datos, y la



segunda aborda aspectos relacionados con el procesamiento y los resultados del uso del registro administrativo.

De acuerdo con lo anterior, los indicadores están enfocados en evaluar los siguientes aspectos:

- **Calidad de entrada:** este conjunto de indicadores se focaliza en evaluar la base de datos (o conjunto de datos) proveniente de la oficina administrativa, así como de los metadatos que la acompañan. La evaluación de los indicadores debe utilizarse para tomar acciones correctivas según el estado de los datos y documentarlos para explicitar las decisiones en esta etapa inicial y que afectarán el procesamiento y análisis. A la evaluación de la calidad de la entrada se le puede denominar también, “análisis o evaluación” de la fuente, pues se hace referencia a la medición del estado original en el que se encuentran disponibles los datos provenientes de la entidad administrativa.
- **Calidad del proceso:** conjunto de indicadores que tiene por objetivo evaluar la calidad de los datos que resultan del tratamiento (o también conocido como procesamiento) y/o de la integración de los datos administrativos en su estudio con fines de uso estadístico. Está orientado a evaluar las acciones más relevantes que se realizan durante el procesamiento y análisis, apuntando a entregar información respecto a la precisión de la información producida y potenciales acciones de edición de datos, así como su relación con las definiciones metodológicas de medición.

Es fundamental que, en la aplicación de los indicadores se considere la complejidad del propósito estadístico del uso del registro para evaluar la pertinencia del uso de un indicador específico. Según el propósito o uso estadístico del registro, algunos indicadores pueden resultar atingentes de aplicar, mientras que otros no. Por ejemplo, si el objetivo es generar tabulados con estadísticas descriptivas, existen indicadores que no serán pertinentes porque ciertos procesos no se realizan. Por otro lado, si se tiene como objetivo la elaboración de un Registro Estadístico de Población o un Registro Estadístico de Empresas, será pertinente evaluar ciertas acciones relacionadas con labores de procesamiento e integración, y vinculación entre varios registros administrativos, que, en el primer ejemplo no eran necesarias.

En ese sentido, en la formulación de cada indicador, se señala explícitamente como recomendación, para qué tarea está dirigido cada uno de ellos, permitiendo a los usuarios tener mayor orientación sobre los procesos realizados en el uso estadístico del registro.

### 1.3. Alcances de los indicadores

Dado el carácter reciente de la discusión de evaluación de la calidad de registros administrativos y ante la emergente documentación de referencia, es importante avanzar en la generación de evidencia que permita establecer niveles de aceptabilidad respecto al uso de la información contenida. En este sentido, cada equipo de trabajo deberá identificar aquellos indicadores factibles de implementar, de acuerdo con sus requerimientos y necesidades. Además, se deberán definir los distintos niveles de cumplimiento o categorías de resultados que arroje cada indicador (por ejemplo, su aceptabilidad o no), de acuerdo con la experiencia y evidencia de resultados que tenga cada equipo con un indicador específico (construcción de umbrales tolerables).

Otra de las características de este conjunto de indicadores es que se orienta a evaluar aspectos cuantitativos de la base de datos, por tanto, aunque existen indicadores cualitativos en la batería respecto al cumplimiento de tareas o insumos dentro del uso del registro, estas actividades son más oportunas de evaluar mediante otras herramientas que permitan desarrollar con mayor profundidad las bondades o limitantes del uso del registro, por lo que su inclusión se limita a aspectos puntuales de la evaluación. Por ejemplo, si se requiere evaluar la pertinencia entre conceptos y variables presentes en el registro para su uso estadístico, es más recomendable desarrollar un análisis o un ejercicio de operacionalización conceptual en sustitución a la aplicación directa de un indicador que mostrará si satisface criterios de pertinencia o precisión desde el punto de vista de una medición puntual.

#### 1.4. Formulación de los indicadores

Para los fines de esta guía, cada indicador se expone con información de referencia para entregar orientaciones respecto a su aplicación práctica. Cada indicador propuesto se presenta en el formato de ficha la cual contiene los siguientes elementos:

- **Nombre del indicador:** contiene el nombre y su sigla/abreviación en caso corresponda.
- **Definición:** expresa el significado del indicador en cuanto a su formulación.
- **Etapas de evaluación:** puede estar referida a la calidad de entrada o a la calidad del proceso y generación de resultados.
- **Principio de calidad asociado:** según los principales principios de calidad definidas en el Marco de aseguramiento de calidad, se define a cuál de ellas se asocia cada indicador, a saber: accesibilidad, claridad, credibilidad, coherencia, comparabilidad, oportunidad, precisión, puntualidad y relevancia.
- **Uso e interpretación:** se indica para que se ocupa, como también en otras ocasiones, indica cuando es recomendable calcular este indicador.
- **Fórmula de cálculo:** se entrega la fórmula de cálculo con el detalle de sus componentes. La fórmula puede ser cualitativa o cuantitativa.
- **Referencias:** bibliografía sobre la cual el indicador fue adaptado.

#### 1.5. Conceptos relevantes

Con el fin de contribuir a la correcta aplicación de la batería de indicadores, se presentan los principales conceptos que transversalmente son incluidos en la formulación de cada indicador y que hacen referencia al ordenamiento de un conjunto de datos para ser analizado.

- **Variable:** es una característica o atributo medible de una unidad estadística la cual puede asumir un conjunto de valores los cuales pueden ser cuantitativos o cualitativos. Para el análisis de los indicadores de la presente batería, cada variable representará una columna del conjunto de datos y en caso de no ser presentada de esta forma (p.ej. en casos en los que una variable esté contenida en más de una columna dentro del registro administrativo), la edición de los datos debe garantizarla.

- **Observación:** es un conjunto de mediciones realizadas en condiciones similares sobre un mismo objeto o unidad estadística. En este caso cada observación corresponde a una fila dentro del conjunto de datos, reconociendo que una misma unidad de análisis podría tener más de una fila dentro de un conjunto de datos<sup>4</sup>.
- **Valor:** es el estado o resultado de variable en el momento en que fue medida para una unidad determinada. Los valores ocupan un campo o una celda dentro de un conjunto de datos, los cuales corresponden a la intersección entre una fila (observación) y una columna (disposición de la variable).

En la figura n°1 se presenta de manera gráfica, la forma en que deben ser interpretados los conceptos mencionados anteriormente para los procedimientos del cálculo de cada indicador.

Imagen 1. Conceptos relevantes para la aplicación de batería de indicadores



Fuente: imagen extraída desde <https://jkunst.com/computacion-estadistica-EPG3308/slides/02-AED-Viz-ggplot2.html#4>

Por ejemplo, en el caso en que se requiera analizar el porcentaje de errores que tiene una variable, se deben contabilizar todos los valores que presentan errores (de acuerdo con las especificaciones definidas) para el total de valores de dicha variable. En la figura, cada variable tiene un total de 5 valores, por tanto, si se requiere analizar el porcentaje de errores para cada una de ellas se debiese contabilizar los valores que presentan errores del total.

<sup>4</sup> Un ejemplo muy gráfico se encuentra en bases de datos de salud y reproducción, en donde, para una misma persona pueden registrarse distintos eventos médicos de distinta naturaleza, los cuales serán representados en distintas filas de una base de datos, aun cuando representan a un mismo RUN.

## 2. BATERÍA DE INDICADORES

A continuación, se presenta el listado completo de indicadores de acuerdo con la etapa de evaluación para la cual van dirigidos. Para evaluar la calidad de entrada del registro se presentan 7 indicadores y para la calidad del proceso 12 indicadores, respectivamente.

Tabla 1. Batería de indicadores para la evaluación de RRAA

Etapa de evaluación	Indicadores
<b>Calidad de Entrada</b>	<ol style="list-style-type: none"> <li>1. Retraso en la entrega/disponibilidad de los RRAA</li> <li>2. Actualidad del registro</li> <li>3. Cumplimiento en la entrega de metadatos del registro</li> <li>4. Porcentaje de errores por variable del conjunto de datos</li> <li>5. Porcentaje de variables con errores en la base de datos</li> <li>6. Porcentaje de observaciones duplicadas</li> <li>7. Tasa de completitud de las variables</li> </ol>
<b>Calidad del proceso y resultados</b>	<ol style="list-style-type: none"> <li>8. Porcentaje de subcobertura</li> <li>9. Porcentaje de sobrecobertura</li> <li>10. Tasa de edición</li> <li>11. Tasa de imputación</li> <li>12. Tasa de codificación</li> <li>13. Tasa de unidades no vinculables</li> <li>14. Tasa de vinculación</li> <li>15. Tasa de robustez de valores atípicos</li> <li>16. Uso de métodos de anonimización</li> <li>17. Grado de coherencia con otras fuentes</li> <li>18. Tiempo (lapso) de desfase en la entrega de resultados finales</li> <li>19. Puntualidad de la publicación de los resultados</li> </ol>

### 2.1 Retraso en la entrega/disponibilidad de los RRAA para su aprovechamiento estadístico

Nombre del Indicador	Retraso en la entrega/disponibilidad de los RRAA para su aprovechamiento estadístico.
Definición	Diferencia en días entre la fecha programada de entrega de los registros administrativos para su aprovechamiento estadístico y la fecha real de entrega a la Oficina Estadística.
Etapa de evaluación	Calidad de entrada
Principio de calidad asociado	Puntualidad
Uso e interpretación	Su uso está orientado a evaluar la calidad de la entrada del conjunto de datos proveniente de la Oficina Administrativa, en este caso, el objetivo es medir el retraso en la entrega de la base de datos que contiene la información del o los registros.

	<p>Este indicador permite cuantificar un retraso en la entrega pactada<sup>5</sup> entre servicios y la disponibilidad de los RRAA, lo cual puede generar inconvenientes en la planificación de la Oficina Estadística y desajustes en la ejecución de procesos de cálculo predefinidos. Además, puede incidir directamente en la calidad de los análisis y resultados si el retraso tiene algún impacto en la oportunidad de su uso, es decir, un desfase entre el momento en que ocurre el fenómeno estudiado y la publicación de la información estadística o en el tiempo disponible para procesar y analizar adecuadamente los datos, previa publicación oficial.</p> <p>Si existe diferencia entre la fecha programada y la fecha real de recepción del registro, hay que evaluar acciones que permitan ajustar los convenios y/o arreglos entre instituciones para solucionar los retrasos e impacto en las actividades planificadas en el uso estadístico del registro.</p>
Fórmula de cálculo	<p><b>T1 = F1 – F2</b></p> <p>F1: Fecha programada de entrega de los RRAA para su aprovechamiento estadístico. F2: Fecha real de entrega de los RRAA a la Oficina Estadística.</p> <p>Unidades de medida: días hábiles (puede convertirse en semanas o meses si se estima conveniente).</p>
Referencias	Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 44).

## 2.2 Actualidad del registro

Nombre del Indicador	Actualidad del registro
Definición	Diferencia en días entre la fecha de referencia de los datos del registro y la fecha de disponibilidad para su uso estadístico.
Etapas de evaluación	Calidad de entrada
Principio de calidad asociado	Puntualidad
Uso e interpretación	<p>Su uso está orientado a evaluar la calidad de la entrada del conjunto de datos proveniente de la oficina administrativa, en este caso, el objetivo es medir el tiempo de diferencia entre la fecha al cual hacen referencia los datos del registro y la fecha en la cual se encuentra disponible para su uso estadístico (al momento de entrada de los datos para la oficina estadística).</p> <p>Este indicador busca identificar el grado de desfase del uso del registro, producto de la brecha temporal y que puedan incidir en la desactualización de la información, por ejemplo, considerando unidades que si bien en el momento de captación del registro eran parte de la población de interés, en el presente ya no lo son; también puede que las características estructurales de las unidades hayan cambiado en el</p>

<sup>5</sup> Dicho acuerdo entre servicios públicos puede concretarse mediante actos administrativos tales como convenios de colaboración, acuerdos de transferencia de datos vía oficios, resoluciones, entre otros.

	<p>tiempo o que pueda afectar la integración y/o comparabilidad con otros registros debido a la brecha temporal.</p> <p>La fecha a la cual hacen referencia los datos puede ser un periodo o tramo de tiempo que, por ejemplo, abarque entre el primer día y el último día del año, en cuyo caso se debe tomar como fecha de cálculo, el último día del año, pues a esa fecha se encontraban vigentes los datos del registro. En caso de que el registro sea extraído en un momento dado del tiempo (una “foto”) la fecha de cálculo debe ser aquella que corresponda al día en que se obtuvo dicha extracción u obtención del registro.</p>
Fórmula de cálculo	<p><b>T1 = F1 – F2</b></p> <p>F1: Fecha al cual hacen referencia los datos. F2: Fecha del registro disponible para su uso estadístico.</p> <p>Unidades de medida: días corridos (puede convertirse en semanas o meses si se estima conveniente).</p>
Referencias	United Nations Guidelines on Statistical Business Registers (UNECE, 2020, pág. 247).

### 2.3 Cumplimiento en la entrega de metadatos del registro

Nombre del Indicador	Cumplimiento en la entrega de metadatos del registro
Definición	Nivel de cumplimiento de entrega del libro de códigos o diccionario de variables de la base de datos del registro administrativo.
Etapas de evaluación	Calidad de entrada
Principio de calidad asociado	Gestión de metadatos
Uso e interpretación	<p>Busca conocer desde una perspectiva cualitativa si se cumplieron con los metadatos requeridos para la correcta interpretación y procesamiento de la base de datos del registro, particularmente el libro de códigos o diccionario de todas las variables que componen el registro. Este indicador se debiese controlar (verificar) inmediatamente tras la recepción de los archivos de datos del registro.</p> <p>El libro de códigos o diccionario de variables es un metadato que proporciona la descripción de las variables, particularmente respecto al nombre de la variable, la etiqueta tipo de variable, los valores que pueden asumir la variable, valores perdidos (omisiones) y la etiqueta de los valores. En ese sentido, el diccionario de variables o libro de códigos es una herramienta básica y fundamental para la revisión de la calidad del registro, así como también entrega información respecto cómo tratar a cada variable que compone la base o conjunto de datos.</p> <p>En caso de no existir este metadato, es una de las primeras tareas que se deben realizar para proseguir con el análisis de calidad, siendo importante considerar que muchas veces, la actividad de generar un diccionario de variables se deba realizar de forma conjunta entre la Oficina Estadística y la institución que provee el registro.</p>

Fórmula de cálculo	<ul style="list-style-type: none"> <li>• <b>Cumplimiento ideal:</b> el proveedor del registro administrativo entrega un libro de códigos o diccionario de variables a la Oficina Estadística y este se encuentra en condiciones óptimas para la interpretación y análisis del conjunto de datos.</li> <li>• <b>Cumplimiento intermedio:</b> el proveedor del registro administrativo entrega un libro de códigos o diccionario de variables a la Oficina Estadística, pero este no se encuentra en condiciones óptimas para la interpretación adecuada del conjunto de datos. Por ejemplo, se encuentra incompleto, presenta impresiones, corresponde a información sin trazabilidad o presenta errores.</li> <li>• <b>No cumplimiento:</b> el proveedor no dispone de un libro de códigos o diccionario de variables a la Oficina Estadística.</li> </ul>
Referencias	Fortalecimiento de los registros administrativos para su uso estadístico (DANE, 2023). <sup>6</sup>

## 2.4 Porcentaje de errores por variable del conjunto de datos

Nombre del Indicador	Porcentaje de errores por variable
Definición	Razón entre la cantidad de errores por variable y el total de valores de la variable, expresada en porcentaje.
Etapas de evaluación	Calidad de entrada
Principio de calidad asociado	Precisión
Uso e interpretación	<p>Consiste en verificar el cumplimiento de los datos para su uso estadístico, validando que la información cumple los rangos, longitud y dominio previstos para la variable en su descripción.</p> <p>Cuantificar los errores o valores no válidos del conjunto de datos, conforme a lo definido en el libro de códigos o diccionario de variables, permite corregir y tomar medidas metodológicas para un adecuado uso de los datos. Además, se recomienda retroalimentar a la Oficina Administrativa de los hallazgos, permitiendo el conocimiento y corrección de potenciales fuentes de error.</p> <p>Según las necesidades de los equipos que analicen la base de datos, convenga especificar este indicador para cada error específico, por ejemplo, para el rango, longitud y dominio por separado, y construir el indicador agregado del total de errores para la variable. Adicionalmente, es preciso indicar que en su aplicación podrían existir subgrupos para los cuales deba distinguirse su “no medición” por definición de la forma en que se construye la variable versus un error presente en la observación analizada.</p>
Fórmula de cálculo	$PEV = \frac{\text{Número de valores con errores de la variable}}{\text{Número total de valores de la variable}} \times 100$ <p>El denominador refiere al total de valores pertinentes de la variable para el momento en el tiempo en que analiza.</p>

<sup>6</sup> Corresponde a un curso dictado por el Departamento Administrativo Nacional de Estadística (DANE) el año 2023, en el cual participaron funcionarios del INE. Para más información: <https://www.dane.gov.co/index.php/actualidad-dane/5664-inscribase-a-los-cursos-virtuales-gratuitos-del-sistema-estadistico-nacional-sen-2023>.

	Solo se deben considerar dentro del denominador aquellos valores en los que la presencia de información sea un requisito definido.
Referencias	Fortalecimiento de los registros administrativos para su uso estadístico (DANE, 2023).

## 2.5 Porcentaje de variables con errores en la base de datos

Nombre del Indicador	Porcentaje de variables con errores
Definición	Razón entre la cantidad de variables errores y al total de variables del conjunto de datos, expresada en porcentaje.
Etapas de evaluación	Calidad de entrada
Principio de calidad asociado	Precisión
Uso e interpretación	<p>Permite tener una panorámica general del conjunto de variables del registro que presentan problemas, con el fin de transmitir a la Oficina Administrativa los problemas del conjunto de datos y evaluar posibles correcciones a realizar.</p> <p>Es importante contar con la información respecto al tipo de variable, longitud, dominio y rango para determinar cuándo un valor presenta algún error. Por lo tanto, este indicador se construye después de un análisis detallado de cada variable de forma independiente (de acuerdo con el indicador anterior n°3).</p> <p>La información que provee este indicador es importante comunicarla de forma oportuna con la Oficina Administrativa para poder corregir aspectos relacionados con la entrega de información, así como también con las posibles correcciones (ediciones) que se le deban realizar a la información.</p>
Fórmula de cálculo	$TVE = \frac{\text{Número de variables con errores}}{\text{Número total de variables}} \times 100$ <p>El numerador refiere a todas las variables que presentaron errores dentro del conjunto de datos (la cantidad de errores va medida de 1 a “n”), es decir, aquellas variables que presentaron a lo menos un error deben ser consideradas.</p> <p>Por su parte, el numerador corresponde al total de variables que contiene el conjunto de datos.</p>
Referencias	Fortalecimiento de los registros administrativos para su uso estadístico (DANE, 2023).

## 2.6 Porcentaje de observaciones duplicadas

Nombre del Indicador	Porcentaje de observaciones duplicadas
Definición	Razón entre cantidad de observaciones duplicadas y total de observaciones del conjunto de datos, expresada en porcentaje.
Etapas de evaluación	Calidad de entrada



Principio de calidad asociado	Precisión
Uso e interpretación	<p>Se utiliza para saber cuántas observaciones pertenecientes al conjunto de datos se detectan como casos duplicados. Para este indicador se toma como definición de un caso duplicado cuándo todos los valores en dos o más registros son idénticos, es decir, cada campo o columna en los registros tiene el mismo valor. La detección de casos duplicados es relevante dentro del procesamiento del registro administrativo puesto que permite identificar casos en que, de acuerdo con los criterios metodológicos adoptados, pueden ser eliminados o no tomados en cuenta para cálculos o análisis, debido a la alteración que podría afectar al análisis de los datos.</p> <p>Junto a lo anterior, la existencia de casos duplicados puede ser indicio de problemas en fases o procesos de gestión del registro por parte de la oficina administrativa, en ese sentido, en caso de detectar este tipo de casos es importante poder informar a los organismos pertinentes.</p> <p>Cabe señalar que, para la evaluación de la calidad de la base de datos, quedan fuera de la conceptualización de caso duplicado, a aquellos casos en que por medio del análisis de una variable identificadora, una unidad se encuentre presente más de una vez en una base de datos (o conjunto de datos), puesto que dichos casos pueden corresponder a diferentes eventos en las que la unidad participa del acto administrativo, por tanto, los valores para ciertas variables van a ser los mismos, pero para otras no. Por ejemplo, una persona usuaria del sistema público de salud, que acudió en diversas oportunidades a realizar exámenes médicos o controles, no corresponde a un caso duplicado en tanto cada prestación médica corresponde a un evento o acto totalmente diferente.</p> <p>Por último, este indicador también se recomienda calcular al momento de integrar o vincular dos o más registros diferentes para detectar posibles duplicados que surjan al combinar conjuntos de datos diferentes.</p>
Fórmula de cálculo	$Pod = \frac{\text{Número de observaciones duplicadas de la base A}}{\text{Número total de observaciones de la base A}} \times 100$ <p>El numerador corresponde al número total de observaciones duplicadas en el registro, las que se pueden identificar mediante la comparación y análisis de identificadores únicos u otras variables relevantes.</p>
Referencias	<p>Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 30).</p> <p>Fortalecimiento de los registros administrativos para su uso estadístico (DANE, 2023).</p>

## 2.7 Tasa de completitud de las variables

Nombre del Indicador	Tasa de completitud de variables (TCV)
Definición	Razón entre el número de valores con datos y el número de valores total requeridos de alguna variable específica.
Etapas de evaluación	Calidad de entrada

Principio de calidad asociado	Precisión
Uso e interpretación	<p>Este indicador es fundamental en el análisis de la calidad de los registros administrativos puesto, que a mayor nivel de completitud de los datos que forman parte de una variable, mayor es su precisión conforme se tendrá toda la información disponible para analizarse. Por contrapartida, en caso de que existan datos incompletos en mayor proporción, indica que el uso de esa variable tendrá problemas para los posteriores análisis que se deban realizar, sobre todo si es que la ausencia de valores se concentra en alguna subpoblación específica. Este indicador permite comunicar a la oficina administrativa los problemas relacionados con la falta de información y a tomar acciones para a corregir dichos problemas de captación de los datos y/o procesamiento.</p> <p>Este indicador debe ser calculado cuando el conjunto de datos llega a la oficina estadística (calidad de la entrada). Por su parte, se puede considerar el cálculo de este indicador como parte de la entrega de resultados (calidad de los resultados) que puede considerar, por ejemplo, las acciones de imputación.</p> <p>Junto a lo anterior, tener información complementaria sobre las razones de que los datos se encuentren incompletos puede ayudar a tomar decisiones metodológicas en el tratamiento de la información, como, por ejemplo, en las acciones de edición, así como en los planes para mejorar la completitud en el futuro del registro.</p>
Fórmula de cálculo	$TCV = \frac{\text{Número de valores con datos}}{\text{Número de valores de datos pertinentes}} \times 100$ <p>El numerador corresponde al conjunto de valores con información para la variable de análisis.</p> <p>El denominador refiere al total de valores pertinentes de la variable (solo se deben considerar dentro del denominador aquellos valores en los que la presencia de información sea un requisito).</p>
Referencias	ESS Quality and Performance Indicators (UNECE, 2014, pág. 3).

## 2.8 Tasa de subcobertura

Nombre del Indicador	Tasa de Subcobertura
Definición	Razón entre la población que contiene el registro y el total de unidades de la población objetivo.
Etapas de evaluación	Calidad de proceso
Principio de calidad asociado	Precisión
Uso e interpretación	Este indicador indica falta de cobertura del registro referida a la población objetivo, evidenciando las limitaciones del registro en cuanto no contiene a la totalidad de la población que se requiere analizar. Para realizar esta evaluación es importante contar con un conocimiento general de la población objetivo que conforma el registro, así como

el periodo de referencia con el que se realizará la comparación. Esto asume que para el cálculo del indicador se disponga de una fuente auxiliar o un archivo de referencia.

La subcobertura de un registro administrativo, puede llegar a ser una limitación frecuente o común cuando se lo pretenda usar con fines estadísticos, sin embargo, no constituye un impedimento para el uso del registro, sino más bien, corresponde a un potencial alcance (y sesgo) que puede ser atenuado con el uso de otros registros administrativos para satisfacer los requerimientos estadísticos en cuanto a la completitud de la población objetivo.

Por ejemplo, en el marco de la conformación de un Registro Estadístico de Población, se puede disponer de una base de datos de los trabajadores asalariados del país. Dicho registro no considera a personas que puedan trabajar en el sector informal o que no estén trabajando en un periodo determinado, ni a personas menores de una edad determinada (menos de 15 años, por ejemplo). Dichas limitaciones sobre la población objetivo no implican desechar el uso del registro, sino considerar los alcances en el uso de la información, definir y acotarlo conceptualmente y recurrir a otras fuentes que permitan complementar las observaciones presentes en el registro si se estima pertinente.

Si no es posible estimar el total de la población objetivo, por ejemplo, con otro registro administrativo de referencia o auxiliar que provea los stocks poblacionales, inicialmente el indicador se puede expresar indirectamente documentando y señalando a la población que —teóricamente— no se considera y estableciendo estas limitaciones/exclusiones como un alcance dentro del uso estadístico del registro.

Mientras más se acerque al 100% el resultado de este indicador dará cuenta de que el registro tiene problemas de subcobertura respecto a la población objetivo. De ser necesario este indicador debe ser calculado para subpoblaciones de interés definidos en el diseño del procesamiento y análisis. Idealmente su valor observado debe acercarse a 0.

$$T_{Sub} = \frac{\text{Total de unidades del registro de referencia no contenidos en el registro}}{\text{Total unidades del registro de referencia}} \times 100$$

**Fórmula de cálculo** El numerador corresponde al total de unidades del registro administrativo de referencia que no se encuentran contenidas en el conjunto de datos del registro que se está evaluando. Por su parte el denominador contiene todas las unidades del registro de referencia. Se asume que las unidades del registro de referencia contienen el total de unidades que conforman la población objetivo.

**Referencias** Fortalecimiento de los registros administrativos para su uso estadístico (DANE, 2023).

## 2.9 Tasa de sobrecobertura

Nombre del Indicador	Tasa de sobrecobertura
Definición	Razón entre el total de unidades que refiere a la población que contiene el registro y el total de unidades de la población objetivo.

Etapa de evaluación	Calidad de proceso
Principio de calidad asociado	Precisión
Uso e interpretación	<p>La sobrecobertura en registros administrativos se refiere a la inclusión de ciertos grupos o entidades en los datos, para los cuales no corresponde su observación bajo la definición específica de población objetivo del registro con fines de uso estadístico. Algunos registros pueden incluir entidades que no son relevantes para el análisis en cuestión, por ejemplo, en los registros de empresas, puede haber inclusiones de entidades que ya no se encuentren activas o que correspondan a secciones de un clasificador estadístico que no corresponde medir o caracterizar y, por tanto, fuera de la población de interés estadístico.</p> <p>Para realizar esta evaluación es importante contar con un conocimiento general de la población objetivo que conforma el registro (su proceso generador), así como el periodo de referencia con el que se realizará la comparación. En particular, es importante señalar que se requiere tener disponible una fuente de información auxiliar que permita conocer y definir valores poblaciones totales (stock) que faciliten la referencia/comparación.</p> <p>La sobrecobertura puede afectar la calidad de los análisis estadísticos y las inferencias derivadas de los registros administrativos. Para abordar estos problemas, es esencial realizar una validación y limpieza exhaustivas de los datos, identificar y revisar los criterios de inclusión de las unidades en los registros para garantizar una representación más precisa de la población de interés. Además, es importante considerar y documentar cualquier sesgo potencial asociado con la sobrecobertura al interpretar los resultados.</p> <p>Mientras más cercanos al 100% sean los resultados, indicarán una mayor sobrecobertura, es decir, hay más unidades en los registros administrativos de lo que debería haber según la estimación (stock) de la población de referencia mediante una encuesta o una fuente auxiliar.</p>
Fórmula de cálculo	$TSob = \frac{\text{Total de unidades del registro no contenidas en el registro de referencia}}{\text{Total de unidades del registro administrativo}} \times 100$ <p>El numerador representa el total de unidades del registro administrativo que no se encuentran contenidas en el registro de referencia, asumiendo que este registro contiene a la totalidad de la población objetivo que se está estudiando.</p> <p>Por su parte el denominador considera al total de unidades del registro administrativo analizado para fines estadísticos.</p>
Referencias	DANE 2018 Metodología de Diagnóstico de los Registros Administrativos para su Aprovechamiento Estadístico.

## 2.10 Tasa de edición

Nombre del Indicador	Tasa de Edición
Definición	Razón entre el total de valores editados de una variable y el total de valores de la variable con errores (rangos, longitud, dominios) e inconsistencias del conjunto de datos analizado.

Etapa de evaluación	Calidad del proceso
Principio de calidad asociado	Precisión
Uso e interpretación	<p>La edición corresponde a la actividad destinada a detectar y corregir errores (inconsistencias) identificados en la etapa de entrada del registro administrativo. Esta actividad, considera una variedad de procedimientos y procesos utilizados para detectar y manejar errores en los datos.</p> <p>En particular, este indicador busca conocer la proporción de valores editados respecto a aquellos valores que presentan errores (rangos, longitud, dominios) de cada variable. Este indicador debe ser utilizado cuando en el tratamiento/procesamiento de la variable del registro, ya se hayan tomado las decisiones en cuanto a la edición de datos y, por tanto, se dé cuenta cuantitativamente de la proporción de valores editados. Junto a lo anterior, este indicador puede ser calculado para el total de valores de la variable, así como para el total de valores con errores (ponderado).</p> <p>Si este indicador presenta valores menores al 100%, deben tomarse acciones de justificación y debida documentación de las decisiones por la cuales no fueron editados todos los casos identificados.</p> <p>Es importante considerar que se deben documentar las decisiones metodológicas en el proceso de edición debido a que esto permite tener una mayor comprensión sobre los datos y los análisis que se realicen. Lo anterior permite realizar una retroalimentación a la oficina administrativa.</p>
Fórmula de cálculo	$TEDI = \frac{\text{Número de valores editados}}{\text{Número total de valores de la variable con errores o inconsistencias}} \times 100$ <p>El numerador considera al total de valores editados, mientras que el denominador abarca al total de valores con errores o inconsistencias detectados en el proceso de revisión e inspección de los datos.</p> <p>El denominador incluye el total de valores de la variable que presentan errores que ya han sido identificados dentro del proceso de evaluación de la calidad de entrada.</p>
Referencias	“Informe técnico de definición de indicadores base para la medición de la calidad del Registro Estadístico de Empresa – REE” (Aguilera, 2021).

## 2.11 Tasa de imputación

Nombre del Indicador	Tasa de imputación
Definición	Razón entre el número de valores imputados y el número total de valores para la variable analizada.
Etapa de evaluación	Calidad del proceso
Principio de calidad asociado	Precisión
Uso e interpretación	La imputación es un procedimiento que se emplea para asignar valores de reemplazo para datos faltantes, inválidos o inconsistentes dentro del análisis de

	<p>determinadas variables, y que, por su importancia relativa en el análisis y generación de resultados, deben contar con dicha información.</p> <p>El objetivo principal de la tasa de imputación es tener un parámetro o medida que permita identificar el porcentaje de observaciones que se imputan para una variable determinada o las más importantes del registro. La tasa de imputación se debe calcular y comparar sobre los diferentes dominios del registro administrativo y en caso de que sea necesario, profundizar en ciertos subgrupos poblacionales de interés.</p> <p>La imputación es un procedimiento utilizado ante la falta de valores para una variable de un registro administrativo (en las encuestas, de la no respuesta y no la no respuesta al ítem, por ejemplo), pero en algunos casos puede usarse cuando los datos de una variable analizada presenten inconsistencias o mediante análisis estadístico se verifique una determinada observación no debe ser parte del conjunto de datos, en cuya situación se puede optar por la decisión metodológica de imputar (reemplazar) información a estos datos. Resulta muy importante que la decisión de imputar sea ponderada en función del comportamiento aleatorio o no de la ausencia de los datos y de la existencia de un registro que conceptualmente aborde la misma población objetivo de estudio estadístico, pues, en caso de no garantizar estos elementos, la decisión de imputación puede carecer de fundamento metodológico o temático.</p> <p>En la literatura, en general, a esta tasa se la reconoce o puede asociarse a la tasa de imputación “no ponderada” (previa a la aplicación de factores de expansión a los datos), diferenciándola de la tasa de imputación “ponderada”, la cual permite conocer la contribución relativa de los valores imputados a la estimación de este ítem/variable.</p>
Fórmula de cálculo	$TI = \frac{\text{Núm. de valores imputados para la variable}}{\text{Núm. total de valores de la variable}} \times 100$ <p>El numerador incluye todos los valores que fueron imputados para la variable de interés. Por su parte, el denominador incluye el total de valores pertinentes de la variable en un periodo de tiempo especificado.</p>
Referencias	<p>Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 34).</p> <p>ESS Quality and Performance Indicators (UNECE, 2014, pág. 17).</p>

## 2.12 Tasa de codificación

<b>Nombre del Indicador</b>	Tasa de codificación
<b>Definición</b>	Razón entre el número de valores codificados y el total de valores a codificar en la variable.
<b>Etapas de evaluación</b>	Calidad del proceso
<b>Principio de calidad asociado</b>	Precisión
<b>Uso e interpretación</b>	La codificación corresponde al proceso que permite convertir información verbal o textual en códigos que representan categorías dentro de un esquema de clasificación estadística, para facilitar el procesamiento de datos (INE, 2022).

	<p>En el contexto de registros administrativos, el indicador busca conocer la magnitud de valores que lograron codificarse a la clasificación deseada, considerando que puede corresponder a una clasificación estandarizada, es decir, asociable a esquema de clasificación uniforme, exhaustivo y jerárquico que provee el marco conceptual de aplicación, ya sea nacional o internacional, o a un esquema ad-hoc o útil para el análisis del registro.</p> <p>En el caso en que el proceso de codificación esté compuesto por codificación manual y automatizada, deberán calcularse por separado y referenciando a cada subconjunto de datos, para reflejar la intensidad con la cual se utiliza cada método de codificación.</p> <p>Mientras la tasa más se acerque a uno (1), la variable se encontrará apta para su uso a los fines estadísticos planificados.</p>
Fórmula de cálculo	$TC_{cod} = \frac{\text{Núm. de datos codificados}}{\text{Núm. total de datos de entrada de la variable a codificar}} \times 100$ <p>El denominador incluye a todos los datos de la variable que lograron ser codificados a un esquema de clasificación. Por su parte, el numerador considera al total de los datos de la variable que se está analizando.</p>
Referencias	Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 33).

### 2.13 Tasa de unidades no vinculables

Nombre del Indicador	Tasa de unidades no vinculables
Definición	Razón entre el número de unidades no vinculables y el total de unidades que contiene el registro.
Etapa de evaluación	Calidad de proceso
Principio de calidad asociado	Precisión
Uso e interpretación	<p>La definición de la unidad observable para la construcción y cálculo de este indicador está sujeta a la construcción de una unidad estadística que será estudiada y que permitirá disponer de una unidad de análisis para los datos tratados. Esta decisión debe tomarse en el diseño y planificación de la estadística basada en registros administrativos o en el diseño del uso estadístico del registro dentro del proceso productivo.</p> <p>Se enfoca en medir la calidad de la(s) variable(s) de vinculación, buscando conocer la cantidad de objetos u observaciones de esta variable que no se encuentran en condiciones de ser vinculados, previamente a la acción de vincular. Este indicador es importante puesto que permite identificar aquellas unidades (viviendas, hogares, personas, empresas, establecimientos, etc.) que de acuerdo con el estado de la variable del registro (el RUN o RUT, por ejemplo) no se encuentran aptas de ser</p>

	<p>vinculables con otro conjunto de datos en caso de ser requerido o no entregan garantías de vinculaciones uno a uno.</p> <p>Este indicador permite conocer la magnitud de aquellas unidades que no pueden vincularse con otro conjunto de datos y quedarán fuera de los análisis. De acuerdo con esto, hay que conocer si los problemas de calidad de esta variable son puntuales o se focalizan en ciertos grupos poblacionales y las posibles causas de estos problemas.</p>
Fórmula de cálculo	$T \text{ no vinculables} = \frac{\text{Número de unidades no vinculables}}{\text{Número total de unidades}} \times 100$ <p>El numerador incluye a las unidades u observaciones que son parte de la variable de interés que no están en condiciones de ser vinculadas. El denominador considera el total de unidades del registro que deben ser vinculadas.</p>
Referencias	DANE 2018 Metodología de Diagnóstico de los Registros Administrativos para su aprovechamiento estadístico.

## 2.14 Tasa de vinculación

Nombre del Indicador	Tasa de vinculación de base A a B ( $PE_{a \rightarrow b}$ )
Definición	Razón de unidades vinculadas desde un conjunto de datos “a” a otro conjunto de datos base “b”, expresado en porcentaje.
Etapas de evaluación	Calidad del proceso
Principio de calidad asociado	Precisión
Uso e interpretación	<p>Este indicador mide efectivamente el resultado de la vinculación entre dos conjuntos de datos distintos, siendo una medida muy importante de la calidad en el uso estadístico de los registros administrativos. Una baja tasa de vinculación puede indicar que los conjuntos de datos cubren diferentes partes de la población o responden a distintas poblaciones objetivo; que el proceso de vinculación no identifica todas las conexiones que existen entre las unidades en cada conjunto de datos o reflejar problemas de calidad referente a las características de la variable de vinculación, las cuales podrían requerir edición previa.</p> <p>Este indicador está orientado para su uso cuando sea necesario vincular uno o más conjuntos de datos con un registro considerado como base, idealmente, la calidad del proceso de vinculación dependerá del estado y características de la variable “llave” o de identificación, en cuanto a sus características. Cabe señalar que puede ser importante analizar este indicador en subpoblaciones específicas de interés.</p> <p>En caso de que en el procesamiento no se realicen procesos de integración entre uno a más registros, este indicador no debe ser utilizado.</p>
Fórmula de cálculo	$PE_{a \rightarrow b} = \frac{\text{Núm. unidades vinculadas de la base A a la base B}}{\text{Núm. total de unidades de la base A que se espera se vinculen a la base B}} \times 100$



	<p>Base A: conjunto de datos que se requiere vincular con B.</p> <p>Base B: conjunto de datos al que se le vincula la información de A.</p>
Referencias	Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 29).

## 2.15 Tasa de robustez de valores atípicos

Nombre del Indicador	Tasa de robustez de valores atípicos (TRVA)
Definición	Razón entre los valores atípicos corregidos, en relación con el total de valores atípicos detectados.
Etapa de evaluación	Calidad del proceso
Principio de calidad asociado	Precisión
Uso e interpretación	<p>El propósito del indicador es medir el porcentaje de valores atípicos que fueron corregidos o verificados correctos. Este indicador medirá la calidad de un proceso de detección de valores atípicos en las variables claves de análisis del registro.</p> <p>Se debe distinguir dentro del cálculo del indicador, aquellos valores atípicos que efectivamente pueden ser considerados como correctos o válidos de aquellos valores atípicos que deben ser corregidos pues son considerados como un error. Sobre este último grupo corresponde aplicar el indicador.</p> <p>Es importante considerar junto al cálculo del indicador, la documentación que dé cuenta de los criterios y metodología adoptada para la detección y corrección de un valor atípico.</p>
Fórmula de cálculo	$TRVA = \frac{\text{Núm. de valores atípicos corregidos}}{\text{Núm. total de valores atípicos de la variable}}$ <p>El numerador considera a los valores atípicos que fueron corregidos durante el procesamiento.</p> <p>El denominador considera el total de valores atípicos incluyendo a aquellos que tras los análisis se consideran válidos.</p>
Referencias	Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 35).

## 2.16 Utilización de métodos de Anonimización

Nombre del Indicador	Utilización de métodos de anonimización u otros
Definición	Mide la aplicación de método de anonimización al registro administrativo con forma a los lineamientos y orientaciones INE.
Etapa de evaluación	Calidad de proceso

Principio de calidad asociado	Accesibilidad y credibilidad
Uso	<p>La anonimización corresponde a un 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.</p> <p>En se sentido, tal como se realiza en otros procesos estadísticos, como encuestas y censos, es una buena práctica que se debe utilizar en el uso estadístico de los registros administrativos, considerando que, en algunos casos, las bases de datos o conjunto de datos resultante en el tratamiento del registro puede ser publicado y difundido.</p> <p>De acuerdo con lo anterior, este indicador mide efectivamente que esta actividad se haya realizado con el fin de garantizar la privacidad de individuos o instituciones (personas naturales y empresas, por ejemplo).</p>
Fórmula de cálculo	<p><b>Cumplimiento ideal:</b> se realiza el proceso de anonimización de las unidades de observación.</p> <p><b>No cumplimiento:</b> no se realiza el proceso de anonimización de las unidades de observación.</p>
Referencias	Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 45).

### 2.17 Grado de coherencia con otras Fuentes

Nombre del Indicador	Grado de coherencia con otras fuentes (GCF)
Definición	Comparación entre los resultados obtenidos del análisis del registro administrativo con otras fuentes de información.
Etapas de evaluación	Calidad de proceso
Principio de calidad asociado	Coherencia y comparabilidad
Uso e interpretación	<p>El nivel de coherencia busca conocer en qué términos, los resultados provenientes del uso del registro son válidos si los comparamos con fuentes externas al registro utilizado, las que pueden corresponder a otros registros y/o a resultados de encuestas o censos. Este indicador no tiene como propósito que los resultados entre el registro utilizado y la fuente externa sean los mismos, se busca garantizar que la interpretación de estos sea coherente.</p> <p>Aunque el grado de coherencia es una medida cualitativa y no genera un resultado o valor cuantitativo, es importante resguardar y realizar este tipo de análisis para garantizar y asegurar la validez del uso del registro, sobre todo cuando el uso reviste de alta complejidad. Si existen alcances o limitantes en el uso e interpretación de fuentes externas, hay que señalarlas y documentarlas.</p> <p>Este indicador debe ser calculado en la medida de que sea pertinente a la complejidad del uso estadístico del registro administrativo.</p>

Fórmula de cálculo	<p><b>Cumplimiento ideal:</b> se realiza el análisis con otras fuentes externas y la comparación es coherente con el registro administrativo en términos de tendencias, magnitudes u otros análisis pertinentes.</p> <p><b>Cumplimiento intermedio:</b> existe coherencia al comparar ciertos niveles de tendencias y/o magnitudes (u otros análisis), sin embargo, en análisis pormenorizados por subpoblaciones o variables de interés, se observan diferencias entre ambas fuentes.</p> <p><b>No cumplimiento:</b> no se observa coherencia entre las fuentes ni a nivel de agregados ni en análisis particulares de subpoblaciones ni variables de interés.</p>
Referencias	Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources (UNECE, 2017, pág. 36).

## 2.18 Tiempo o lapso de desfase en la entrega de resultados finales

Nombre del Indicador	Tiempo (lapso) de desfase en la entrega de resultados finales
Definición	Diferencia (en días) entre la fecha de publicación de los resultados finales para las personas usuarias y el periodo de referencia del registro administrativo.
Etapa de evaluación	Calidad de proceso
Principio de calidad asociado	Oportunidad
Uso e interpretación	<p>La oportunidad de los resultados estadísticos es definida como el tiempo transcurrido entre el final del evento o fenómeno descrito y su disponibilidad, en ese sentido, mientras menor sea el tiempo transcurrido entre la publicación de resultados finales y el periodo de referencia, en este caso, del registro administrativo, más oportuno serán los resultados en la toma de decisiones. Por contrapartida, mientras mayor es el tiempo entre los resultados y el periodo de referencia, la toma de decisiones por parte de los usuarios podría basarse en información desactualizada.</p> <p>En cuanto a las unidades de medida para el cálculo de la diferencia en el tiempo de publicación y el periodo de referencia del registro, el objetivo es que la unidad seleccionada refleje y exprese con la mayor precisión posible el posible desfase entre ambos periodos, puede ser el número de días, semanas o meses, desde el último día del periodo de referencia hasta el día de publicación de los resultados completos y finales.</p> <p>Las unidades de medida pueden expresarse en días, meses y años dependiendo de la naturaleza del tema analizado.</p>
Fórmula de cálculo	<p>Para una variable clave específica:</p> $T = F1 - F2$

	<p>Donde:</p> <p><math>d_{final}</math> = Fecha de publicación de los resultados finales para las personas usuarias.</p> <p><math>d_{refp}</math> = Último día (fecha) del periodo de referencia de las estadísticas.</p>
Referencias	ESS Quality and Performance Indicators (UNECE, 2014, pág. 20).

## 2.19 Puntualidad en la publicación de resultados

Nombre del Indicador	Puntualidad de publicación de resultados
<b>Definición</b>	Diferencia entre la fecha de publicación de los datos y la fecha establecida en el calendario de publicación.
Etapa de evaluación	Calidad de proceso
<b>Principio de calidad asociado</b>	Puntualidad
<b>Uso e interpretación</b>	<p>La puntualidad o la relación de cumplimiento de la puntualidad de los resultados, se mide de acuerdo con lo que establece el calendario de publicación para el INE; en el caso de ser otra institución la que publique los resultados, la fecha comprometida de publicación puede estar reflejada en otro instrumento o documentación.</p> <p>El hecho que exista una diferencia entre la fecha comprometida y la fecha real de publicación es indicador de que existen potenciales problemas en algún componente en la generación de los resultados. En ese sentido, es importante comunicar oportunamente a los usuarios el retraso en la publicación, así como las razones de dichos retrasos y una nueva fecha de publicación.</p> <p>Para los equipos responsables de la producción, las razones de la entrega tardía o no puntual deben indicarse junto con su efecto en el producto estadístico, lo que significa que, debido a la entrega tardía de datos, los procedimientos de garantía de calidad para toda la serie de productos, es posible que no estén completos.</p> <p>Para las personas usuarias: se deben dar algunas explicaciones a los usuarios sobre la publicación no puntual.</p> <p>El no cálculo de este indicador por razones de no disponibilidad de una fecha definida de publicación oficial transgrede una buena práctica estadística relacionada con el principio de Puntualidad.</p>
<b>Fórmula de cálculo</b>	<p>Para los productores:</p> <p>Puntualidad en la entrega de datos <math>P_t</math></p> $P_t = F1 - F2$ <p>Donde:</p> <p>F1 = Fecha real de publicación de los resultados.</p> <p>F2 = Fecha prevista de la publicación de los resultados (por ejemplo, en la página web institucional).</p> <p>Unidades de medida - formato de referencia: días calendario.</p>
<b>Referencias</b>	ESS Quality and Performance Indicators (UNECE, 2014, pág. 21).

### 3. GLOSARIO

**Anonimización:** 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 (INE Chile, 2022).

**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 los mismos y que son almacenados sistemáticamente para su posterior uso (INE Chile, 2022).

**Calidad estadística:** Es el grado o nivel de cumplimiento de las propiedades que deben tener los procesos y productos estadísticos, orientado a satisfacer las necesidades de información de instituciones, organizaciones y personas usuarias (INE Chile, 2022).

**Censo:** Operación estadística realizada sobre el conjunto completo de unidades pertenecientes a una población objetivo en un momento determinado del tiempo (INE Chile, 2022).

**Clave primaria:** Uno o más campos en una tabla de atributos que identifica de forma única una instancia, fila o registro específico (INE Chile, 2022).

**Codificación:** Proceso que permite convertir información verbal o textual en códigos que representan categorías dentro de un esquema de clasificación estadística, para facilitar el procesamiento de datos (INE Chile, 2022).

**Dato administrativo:** Conjunto de datos recopilados a través de procesos administrativos y que provienen de una fuente administrativa. Se entiende que el dato administrativo, se encuentra en un estado previo a cualquier procesamiento o validación que pueda ser efectuada por parte de la oficina estadística (INE Chile, 2022).

**Diccionario de Variables:** El diccionario de variable o también conocido como diccionario de datos, contiene la información que describe las variables utilizadas en un conjunto de datos. Este diccionario proporciona detalles sobre cada variable, incluyendo su nombre, tipo, definición, categorías, unidades de medida y cualquier otra información relevante para facilitar la comprensión y el uso adecuado de los datos (DANE, 2022).

**Edición de datos:** Modificar y corregir errores e inconsistencias en la información recopilada, siguiendo un criterio metodológico establecido posterior a un proceso de verificación y detección de posibles fallos (INE Chile, 2022).

**Estándar estadístico:** Los estándares definen y establecen especificaciones y características uniformes para los productos o servicios. Refieren a un conjunto integral de conceptos estadísticos, definiciones, clasificaciones y modelos, métodos y procedimientos utilizados para lograr el tratamiento uniforme de los problemas estadísticos en los procesos, o entre ellos, y a través del tiempo y del espacio (CEPAL, 2022).

**Evaluación de la Calidad:** Evaluar el grado en que se cumplen los requisitos de calidad según las necesidades o expectativas establecidas (CEPAL, 2022).

**Fecha de referencia del dato:** La fecha a la que hacen alusión los datos recopilados en el registro, indicando o permitiendo reconocer el momento específico al que corresponden las observaciones (Wallgren & Wallgren, 2022).

**Imputación de datos:** Proceso de asignar valores de reemplazo para datos faltantes, inválidos o inconsistentes, garantizando estimaciones de alta calidad y registros coherentes. Se realiza mediante ajustes basados en datos de encuestados similares, considerando estratos y otros factores (INE Chile, 2022).

**Libro de códigos:** Refiere a la información dentro del diccionario de variables o de datos que describe los códigos utilizados en las variables. Estos códigos representan categorías o etiquetas que facilitan la gestión e interpretación de datos (IBM, 2023).

**Metadato Estadístico:** Datos que proporcionan la información esencial para la correcta utilización e interpretación del conjunto de datos y las estadísticas por parte de los usuarios. Estos describen los datos generados, incluyendo detalles sobre la metodología, el trabajo de campo, el procesamiento, el análisis y la calidad, entre otros aspectos, de una operación estadística específica (INE Chile, 2022)

**Oficina Nacional de Estadísticas (ONE):** Organismo de estadística líder en un sistema estadístico nacional. “Oficina Nacional de Estadística” e “Instituto Nacional de Estadística” significan lo mismo. En general, la oficina nacional de estadística asume un rol de coordinación dentro del sistema estadístico nacional y es responsable del desarrollo, producción y difusión de estadísticas oficiales en múltiples dominios estadísticos (CEPAL, 2022, pág. 59).

**Población objetivo:** Conjunto completo de unidades estadísticas en un área geográfica definida y en una fecha determinada, sobre los que se desea información y se requieren estimaciones, para ser medidas mediante una encuesta para informar sobre uno o más temas específicos (INE Chile, 2022).

**Producción Estadística:** Conjunto de subprocesos de transformación de insumos obtenidos a partir de la recolección, almacenamiento, procesamiento y análisis necesarios para la obtención de productos estadísticos, enmarcados dentro de un sistema institucional y cuyo propósito es la generación de información estadística (INE Chile, 2022).

**Producto Estadístico:** Resultados físicos o digitales de una operación estadística, en general corresponden a publicaciones de información mediante la presentación de datos y metadatos, que buscan satisfacer las necesidades de la comunidad usuaria (INE Chile, 2022).

**Registro Administrativo:** Conjunto de datos recopilados y utilizados para fines administrativos por una entidad pública o privada sobre un tipo de hecho, evento, acción, objeto, sujeto, obtenidos sistemáticamente con base en un formato específico ya sea impreso, digital u otro y dentro del marco de sus atribuciones (INE Chile, 2022).

**Registro Estadístico:** Registro administrativo procesado para propósitos estadísticos. Se crean mediante el procesamiento de objetos de registros administrativos de modo que los conjuntos de objetos y las variables satisfagan necesidades estadísticas. El resultado es una base de datos derivada de la transformación o integración de uno o varios registros administrativos, por ejemplo, los registros estadísticos de personas, empresas, hechos vitales, entre otros (INE Chile, 2022).

**Sistema Estadístico Nacional (SEN):** El conjunto de organizaciones y unidades estadísticas (organismos de estadística) de un país que desarrolla, produce y difunde estadísticas oficiales en nombre del Gobierno nacional (y de otros niveles de gobierno). Es responsabilidad de cada país definir el alcance de su SEN (CEPAL, 2022, pág. 60).

**Sobrecobertura:** Inclusión de registros de más unidades de observación como parte de la población objetivo, resultando en un recuento total superior al esperado (Wallgren & Wallgren, 2022).

**Subcobertura:** Ausencia de registros de algunas unidades de observación que corresponden a la población objetivo, resultando en un recuento total inferior al esperado (Wallgren & Wallgren, 2022)

**Validación de datos:** Métodos y procesos para evaluar y verificar la integridad y conformidad de los datos recopilados con respecto a un conjunto predefinido de reglas, estándares o criterios OCDE, 2007).

**Variable:** Una variable es un atributo medible de un objeto. Cuando se captan datos, recibimos valores que son mediciones de esos atributos. Lo que se compila en una matriz de datos son esas mediciones. Una variable puede ser cuantitativa o cualitativa. Las características que se pueden describir con un número (como la edad de una persona), se pueden medir con variables tanto cuantitativas (como edad en años) como cualitativas (como joven, de edad mediana o adulto mayor) (INE Chile, 2022).

**Variable de vinculación:** Permite conectar y asociar datos entre registros, facilitando la identificación y consolidación de información relacionada en diferentes bases de datos dentro de un conjunto de datos (INE Chile, 2022).

**Vinculación de datos:** Proceso de fusionar información proveniente de dos o más fuentes, con el objetivo de consolidar en una única base de datos hechos o individuos relacionados (OCDE, 2007).

## 4. REFERENCIAS

- Aguilera, M. (2021). *Informe técnico de definición de indicadores base para la medición de la calidad del Registro Estadístico de Empresas REE*.
- CEPAL. (2022). *Guía para la implementación del marco de aseguramiento de la calidad para procesos y productos estadísticos*. .
- Cepal. (2023). *CÓDIGO REGIONAL DE BUENAS PRÁCTICAS EN ESTADÍSTICAS*.
- DANE. (2022). *DANE*. Recuperado el Enero de 2024, de <https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/registros-administrativos/programa-de-fortalecimiento#linea-1>
- DANE. (2023). *Fortalecimiento de los registros administrativos para su uso estadístico*.
- IBM. (2023). *IBM*. Recuperado el Enero de 2024, de [https://www.ibm.com/docs/es/spss-statistics/saas?topic=features-codebook#idh\\_codebook\\_variables](https://www.ibm.com/docs/es/spss-statistics/saas?topic=features-codebook#idh_codebook_variables)
- INE. (2022). *Glosario de conceptos estadísticos* .
- INE. (2022). *Orientaciones para la aplicación del Instrumento de Medición de Calidad de Registros Administrativos*. Santiago.
- INE Chile. (2022). *Glosario de conceptos estadísticos*.
- INE Chile. (2022). *Orientaciones para la aplicación del Instrumento de Medición de la Calidad de Registros Administrativos (IMCRA)*. Santiago.
- Naciones Unidas. (2019). *Manual del marco nacional de aseguramiento de*. New York.
- OCDE. (2007). *Glosario de Términos Estadísticos*.
- UNECE. (2014). *ESS Quality and Performance Indicators* .
- UNECE. (2017). *Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources* .
- UNECE. (2020). *United Nations Guidelines on Statistical Business Registers*.
- Wallgren, A., & Wallgren, B. (2022). *Register-based Statistics: Registers and the National Statistical System*.