|  |  |
| --- | --- |
|  | C:\Users\nuria.torroja\AppData\Local\Temp\Temp1_INEGI_a.zip\INEGI_a.png |

**INFORME DE RESULTADOS 2021**



Marzo 2022

|  |  |
| --- | --- |
|  |  |

ANEXO 3
****

# Anexo 3: Marco conceptual y resultados de indicadores para programas de información cuyo insumo son registros administrativos

En 2019, el Comité aprobó tres indicadores de precisión para los programas de información cuyo único insumo son registros administrativos. Para una mejor comprensión de estos indicadores en el cuadro A.3.1 se presentan algunos conceptos generales.

**A.3.1. Conceptos generales aplicados en los indicadores de precisión para
programas de información cuyo único insumo son registros administrativos**

|  |  |
| --- | --- |
| **Registro administrativo**: | Conjunto de Datos que fue generado con fines operacionales o como parte de las funciones de una institución pública o privada sobre un tipo de objeto, sujeto, acción, hecho o evento, y obtenidos sistemáticamente con base en un formato específico ya sea impreso, digital u otro y bajo un marco de funciones y facultades formalmente establecidas en instrumentos jurídicos o reglamentarios. |
| **Proveedor de información**: | Individuo o institución pública o privada que proporciona datos de un conjunto de unidades (por ejemplo, registros administrativos) a través de un Protocolo de Intercambio. Para algunos programas de información, el proveedor de información es una instancia concentradora de varios registros administrativos y, en otros casos los proveedores son varias instancias que recopilan directamente la información de las unidades. |
| **Población**: | Conjunto de unidades pertenecientes a un grupo de personas, empresas, establecimientos, viviendas, o cualquier otro tipo de objetos, acciones o eventos, con base en ciertas características bien definidas, incluyendo límites sobre tiempo y espacio. |
| **Población objetivo**: | Población para la cual se requiere realizar mediciones o representaciones de los conceptos; este conjunto contiene a todos los dominios de estudio. Es decir, es la de interés estadístico, según el fenómeno que se quiere medir. |
| **Población de los registros administrativos**: | Población de la cual existe información en forma nominativa o individualizada en los registros administrativos. Es decir, es la población posible de captar mediante el registro administrativo. |
| **Dominio de estudio**: | Subconjunto de la población objetivo para el cual se requiere realizar mediciones o representaciones de los conceptos de forma separada. Una vez definida la población objetivo, un dominio de estudio corresponde a un subconjunto que está determinado por cuestiones temáticas, temporales o geográficas. Por ejemplo, los dominios de estudio asociados a: entidades federativas, sexo, grupos de edad, etc. En ocasiones los dominios de estudio son referidos como desagregaciones, agregaciones, subpoblaciones, etc. |

Cuando a la población del registro le falta parte de la población objetivo, se dice que existe subcobertura. Para conocer esta diferencia se requieren análisis especiales, por lo que no es posible definir un indicador único; sin embargo, es un análisis fundamental para determinar la pertinencia del programa. Los indicadores que aprobó el Comité no incluyen la subcobertura, sino que se enfocan a medir la calidad a partir de la población observada en el programa en cuestión.

A continuación, se describen los tres indicadores aprobados por el Comité en 2019.

1. **Tasa de sobrecobertura**

Esta tasa permite conocer en qué medida el registro administrativo en cuestión se apega a la población objetivo del programa. Para el cálculo de este indicador resulta útil responder la pregunta del siguiente diagrama:

**A.3.2. Pregunta relevante para la tasa de sobrecobertura**

|  |  |
| --- | --- |
|  |  |

Cabe señalar que los casos duplicados o las unidades que se consideraron fuera de la población objetivo que se utilizan para el cálculo de este indicador, se excluyen de los siguientes indicadores.

1. **Tasa de no respuesta a nivel unidad**

El propósito de este indicador es hacer un análisis más fino, buscando conocer la información que le falta al registro administrativo a través de la estimación de los casos a los que no se tuvo acceso. Esto es, cuando el proveedor de la información no envió una parte de los datos o envió datos no admisibles, el registro recibido está incompleto. En el diagrama A.3.3. se explican las distintas opciones que deben tomarse en cuenta para identificar la no respuesta a nivel de unidad.

**A.3.3. Componentes de la no respuesta a nivel unidad**

|  |  |
| --- | --- |
|  | Donde$N\_{g1}$ es el número total de unidades contenidas en los registros administrativos que no fueron proporcionados por los proveedores de información,$N\_{g2}$ es el número total de unidades bien identificadas de las cuales no se captó información aun cuando se tuvo acceso a los registros administrativos,$N\_{g3}$ es el número total de unidades de las cuales aun cuando no se sabe su identidad se infiere que faltan dado que en el registro administrativo están ausentes unidades correspondientes a cierto aspecto temporal o espacial (e.g., faltó la información de uno de los doce meses)$N\_{g4}$ es el número total de unidades en los registros administrativos a los que se tuvo acceso con información no admisible en al menos una variable$N\_{g5}$ es el número total de unidades en los registros administrativos a los que se tuvo acceso con información admisible en todas las variablesNota. Para realizar el cálculo se eliminan los casos considerados como sobrecobertura, incluyendo duplicados. |

La no respuesta o existencia de valores perdidos está presente en todo programa de información y puede inducir un sesgo en los resultados. El objetivo principal de la tasa de no respuesta a nivel unidad de observación es identificar la existencia de patrones en la no respuesta que pudieran inducir ese sesgo. Para este indicador no se adopta un valor de referencia; sin embargo, valores altos o la existencia de patrones entre las agrupaciones obligan a una revisión de los insumos. Algunas acciones que se pueden generar para corregir este tipo de problemas son mejorar o concretar protocolos de intercambio con los proveedores de información, analizar la conveniencia de usar métodos de imputación o alguna acción para homologar conceptos.

1. **Tasa de no respuesta a nivel variable**

Este indicador busca determinar si a nivel variable falta información del fenómeno a medir en la base de datos final. La tasa de no respuesta a nivel variable se debe interpretar en conjunto y de forma adicional a la tasa de no respuesta a nivel unidad, pues esta última incluye aquellos casos en donde toda la unidad, incluyendo la información de la variable en cuestión, es considerada como no respuesta. La tasa de no respuesta a nivel variable tiene como principal objetivo identificar la existencia de patrones en la no respuesta en una variable en particular. Cuando se presentan variables con alto porcentaje de no respuesta convendrá revisar los objetivos del programa de información para especificar los indicadores relevantes a ser calculados y eliminar aquellas variables que no contribuyen a este objetivo, o si se consideran variables indispensables analizar la conveniencia de usar métodos de imputación o establecer estrategias con los proveedores del registro administrativo para mejorar la respuesta. La pregunta relevante para este caso debe realizarse para cada una de las variables:

**A.3.4. Pregunta relevante para la tasa de no respuesta a nivel variable**

|  |  |
| --- | --- |
|  | Para una variable 𝑍 se consideran sólo las unidades donde se requiere información de ésta, de manera que para un dominio de estudio 𝑔 se calcula como:Donde:𝑁\_𝑍𝑔1 es el número total de unidades en donde en la variable 𝑍 no hay información o ésta desde un inicio corresponde a un código de no respuesta𝑁\_𝑍𝑔2 es el número total de unidades en donde en la variable 𝑍 hay información diferente a un código de no respuesta, pero al realizar la codificación, validación o edición se define como no respuesta𝑁\_𝑍𝑔 es el número total de unidades donde se requiere información de la variable 𝑍 |

En la segunda sesión 2019 del Comité, se aprobaron tres indicadores de uso interno para ser reportados en los metadatos por los responsables de programas de información basados en registros administrativos a partir del segundo semestre de 2019, considerando los siguientes puntos:

* el reporte se debe realizar a más tardar a los 30 días naturales contados a partir de la publicación de la información;
* se debe usar el formato definido en el grupo de trabajo;
* en el caso de la tasa de no respuesta a nivel unidad, para cada programa se incluye en la sección de metadatos una nota técnica explicando el algoritmo específico usado para su cálculo.

Con el fin de dar un panorama general de los resultados de los indicadores reportados en 2021, en el cuadro A.3.5 se presentan los promedios de las tasas de no respuesta a nivel unidad de observación y de las tasas de sobrecobertura. Como puede observarse, para la tasa de no respuesta a nivel unidad de los programas Accidentes de Tránsito Terrestre (ATUS), Transporte Urbano de Pasajeros (ETUP), Sacrificio de Ganado en Rastros Municipales (ESGRM) y Estadística de Divorcios (ED) en 2021 se presenta una recuperación parcial. En contraparte, en 2021 se observa un aumento de la no respuesta en los programas Finanzas Públicas Estatales y Municipales (EFIPEM), y Estadística de Museos (EM). En cuanto a las tasas de sobrecobertura, se observa consistencia a lo largo del tiempo en la mayoría de los programas.

**A.3.5. Comparativo 2019 – 2020 – 2021 del promedio de las tasas de sobrecobertura y de no respuesta a nivel unidad**

| **UA** | **Programa** | **Acrónimo** | **Número de casos totales** | **Promedio de la tasa de sobrecobertura (%)** | **Promedio de la tasa de no respuesta a nivel unidad (%)** |
| --- | --- | --- | --- | --- | --- |
| **2019** | **2020** | **2021** | **2019\*** | **2020\*** | **2021\*\*** | **2019\*** | **2020\*** | **2021\*\*** |
| DGEE | Finanzas Públicas Estatales y Municipales | EFIPEM | 4 | 2 | 2 | 0 | 0 | 0 | 7 (ponderado) | 5 (ponderado) | 5(ponderado) |
| DGEE | Accidentes de Tránsito Terrestre | ATUS | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 4 | **3** |
| DGEE | Transporte Urbano de Pasajeros | ETUP | 6 | 10 | 12 | 0 | 0 | 0 | 0 | 2 | **0** |
| DGEE | Vehículos de Motor Registrados en Circulación  | VMRC | 6 | 12 | 12 | 0 | 0 | 0 | 0 | 0 | 0 |
| (0 anual) /a | (0 anual) /a |
| DGEE | Sacrificio de Ganado en Rastros Municipales | ESGRM | 10 | 12 | 12 | 0 | 0 | 0 | 3 | 5 | **4** |
| (0 anual) /a | (**0** anual) /a |
| DGES | Estadística de Nacimientos Registrados | ENR | - | - | 33 | - | - | 0 | - | - | 0 |
| DGES | Estadísticas de Defunciones Registradas | EDR | 33 | 33 | 33 | 0 | 0 | 0 | 0 | 0 | 0 |
| DGES | Estadísticas de Defunciones Fetales | EDF | 33 | - | 33 | 0 | - | **1** | 0 | - | 0 |
| DGES | Estadística de Matrimonios | EMAT | 33 | 33 | 33 | 0 | 0 | 0 | 0 | 0 | 0 |
| DGES | Estadística de Divorcios | ED | 33 | 33 | 33 | - | 0 | 0 | 0 | 1 | **0** |
| DGES | Estadísticas de Salud en Establecimientos Particulares | ESEP | 33 | 33 | 33 | 0 | 0 | 0 | 0 | 0 | 0 |
| DGES | Estadística de Museos | EM | - | 33 | 33 | - | 0 | 0 | - | 0 | **1** |
| DGES | Estadísticas sobre Relaciones Laborales de Jurisdicción Local | ERLAJUL | - | 33 | 33 | - | 0 | 0 | - | 0 | 0 |

/a Los programas VMRC y ESGRM reportaron, en 2021, una versión anualizada de los indicadores correspondientes al periodo de referencia de la información (2020).

\* Estos datos son los reportados en el Informe Anual de Resultados 2019 y 2020 respectivamente.

\*\* Estos datos corresponden a los indicadores registrados en metadatos durante 2021, independientemente del período de publicación de la información.

En el cuadro A.3.6., se presentan medidas de resumen para la tasa de no respuesta a nivel variable, con la información reportada en 2019, 2020 y 2021. Se observa una recuperación en la no respuesta media a nivel variable para prácticamente todos los programas exceptuando el de Estadísticas de Matrimonios (EMAT). Sin embargo, al revisar las tasas de no respuesta máximas reportadas, se observa que los programas Estadísticas de Defunciones Fetales (EDF), Estadísticas de Nacimientos Registrados (ENR), Estadística de Divorcios (ED) y Estadísticas sobre Relaciones Laborales de Jurisdicción Local (ERLAJUL) presentan casos con variables con no respuesta superiores a 80%, por lo que es necesario revisar si estas variables contribuyen al objetivo del programa y, en su caso, definir estrategias para mejorar la respuesta.

**A.3.6. Valores medios y máximos alcanzados en tasa de no respuesta a nivel variable**

| **Programa** | **Acrónimo** | **Número de variables** | **Tasa de no respuesta a nivel variable (%)** |
| --- | --- | --- | --- |
| **Media** | **Máxima** |
| **2019\*** | **2020\*** | **2021\*\*** | **2019\*** | **2020\*** | **2021\*\*** | **2019\*** | **2020\*** | **2021\*\*** |
| Finanzas Públicas Estatales y Municipales | EFIPEM | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Accidentes de Tránsito Terrestre | ATUS | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| Transporte Urbano de Pasajeros | ETUP | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 |
| Vehículos de Motor Registrados en Circulación  | VMRC | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| (2 anual) /a | (2 anual) /a |
| Sacrificio de Ganado en Rastros Municipales | ESGRM | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| (0 anual) /a | (0 anual) /a |
| Estadísticas de Defunciones Registrados | EDR | 39 | 39 | 39 | 4 | 6 | **3** | 13 | 100 | 38 |
| Estadísticas de Defunciones Fetales | EDF | 45 | - | 45 | 8 | - | 5 | 70 | - | 86 |
| Estadísticas de Nacimientos Registrados | ENR | 34 | - | 33 | 7 | - | 7 | 78 | - | 92 |
| Estadística de Matrimonios | EMAT | 18 | 18 | 18 | 7 | 7 | **8** | 56 | 100 | 76 |
| Estadística de Divorcios | ED | 54 | 54 | 54 | 7 | 8 | 8 | 44 | 90 | 91 |
| Estadísticas de Salud en Establecimientos Particulares | ESEP | - | - | 1 | - | 0 | 0 | - | 0 | 0 |
| Estadística de Museos | EM |  - | 23 | 23 | -  | 0 | 0 | -  | 9 | 14 |
| Estadísticas sobre Relaciones Laborales de Jurisdicción Local | ERLAJUL | - | 7 | 8 | - | 12 | **11** | - | 100 | 100 |

/a Los programas VMRC y ESGRM reportaron en 2021, una versión anualizada de los indicadores correspondientes al periodo de referencia de la información (2020).

\* Estos datos son los reportados en el Informe Anual de Resultados 2019 y 2020 respectivamente

\*\* Estos datos corresponden a los indicadores registrados en metadatos durante 2021, independientemente del período de publicación de la información.