

# Guía de uso de Base de Datos

Encuesta Nacional  
sobre Desnutrición  
Infantil - ENDI

Septiembre, 2024





# Guía de uso de Base de Datos de la Encuesta Nacional sobre Desnutrición Infantil - ENDI

**Dirección/Departamento:**

Dirección de Estadísticas Sociodemográficas

**Unidad:**

Gestión de estadísticas permanentes a hogares

**Elaborado por:**


Margoth Herrera  
José Castellanos  
Gabriela Obregón  
Mayra Pilaguano

**Revisado por:**

Vladimir Tipán

**Aprobado por:**

Estefanía Encalada



## Índice de contenido

|   |    |
|---|----|
| 1. Introducción .....   | 4  |
| 2. Recomendaciones para el usuario .....  | 4  |
| 3. Datos generales de la operación estadística .....                              | 5  |
| 4. Formularios de captura de información .....                                    | 5  |
| 5. Descripción de la Base de Datos.....   | 9  |
| 5.1 Tablas relacionales .....   | 9  |
| 5.2 Variables .....   | 12 |
| 5.3 Número de registros por sección.....  | 15 |
| 5.4 Proceso para unir tablas relacionadas.....                                    | 15 |
| 5.5 Clasificaciones y Nomenclaturas .....   | 17 |
| 6. Tratamiento previo a la publicación.....                                       | 18 |
| 6.1 Anonimización.....  | 18 |
| 6.2 Tratamiento de valores perdidos, imputaciones y validaciones.....             | 20 |
| 7. Réplica e interpretación de indicadores principales.....                       | 21 |
| 7.1 Fichas metodológicas para indicadores.....                                    | 21 |
| 7.2 Factores de expansión para encuestas .....                                    | 22 |
| 7.3 Coeficientes de variación y error asociado al estimador (para encuestas) .... | 22 |
| Referencias bibliográficas .....  | 23 |

## Índice de ilustraciones

|   |    |
|---|----|
| Figura 1. Formulario del Hogar – f1 .....                                       | 6  |
| Figura 2. Formulario MEF – f2.....  | 7  |
| Figura 3. Formulario de Desarrollo Infantil – f3 .....                          | 8  |
| Figura 4. Formato de ficha metodológica de indicadores.....                     | 21 |
| Figura 5. Sección de interpretación del indicador en la ficha metodológica..... | 21 |

## Índice de cuadros

|  |    |
|--|----|
| Cuadro 1. Información de la operación estadística.....                               | 5  |
| Cuadro 2. Descripción de la Base de Datos .....                                      | 9  |
| Cuadro 3. Estructura de codificación para variables que provienen de preguntas ..... | 13 |
| Cuadro 4. Tipo de relación entre tablas de la Base de Datos .....                    | 16 |
| Cuadro 5. Clasificaciones y nomenclaturas utilizadas .....                           | 17 |
| Cuadro 6. Variables anonimizadas .....   | 18 |
| Cuadro 7. Ejemplo de creación de variables codificadas para identificación única ... | 19 |
| Cuadro 8. Tratamiento de valores perdidos en la Base de Datos .....                  | 20 |

# 1. Introducción

La *Encuesta Nacional sobre Desnutrición Infantil* (ENDI) es una operación estadística por muestreo probabilístico que consta en el *Programa Nacional Estadístico 2021-2025*, cuya población objetivo son las niñas y niños menores de 5 años. Para el segundo año consecutivo de levantamiento se investigaron 22.250 viviendas, obteniendo 20.110 viviendas efectivas, con cobertura geográfica en todo el territorio nacional.

La ENDI surge como una operación estadística que busca evaluar la *Estrategia Nacional Ecuador Crece Sin Desnutrición* (ENECSDI) y contar con información sobre la situación de desnutrición infantil de las niñas/os menores de 5 años, con la finalidad de servir de base para el desarrollo de políticas y programas nacionales que contribuyan a minimizar esta problemática tanto como sea posible.

Es necesidad del Estado ecuatoriano contar con información acerca de la desnutrición infantil, para lo cual se han generado indicadores con la finalidad de evaluar políticas relacionadas con la erradicación de la desnutrición crónica en niñas/os menores de 5 años, en los próximos años. Los resultados muestran que el problema de la desnutrición requiere un gran esfuerzo de las instituciones (programas y medidas sociales) que necesita ser evaluado.

La ENDI contribuirá al cumplimiento del *Plan Nacional de Desarrollo 2021 – 2025*, en el objetivo No. 6 el cual establece: “Garantizar el derecho a la salud integral, gratuita y de calidad”, meta: “6.4.1. Reducir 6 puntos porcentuales la desnutrición crónica infantil en menores de 2 años” (Secretaría Técnica Ecuador Crece Sin Desnutrición Infantil).

La ENDI está compuesta por 3 formularios. El Formulario 1 (*Hogar*) recaba información de miembros del hogar, características generales, actividades económicas, datos de la vivienda, seguridad alimentaria, antropometría, prueba de anemia, y pruebas de E. coli y cloro en el agua para beber. El Formulario 2 (*Madres en Edad Fértil con niñas/os menores de 5 años*) recaba información referente a la historia de embarazos y nacimientos, lactancia materna, salud en la niñez y programas de primera infancia. El Formulario 3 (*Desarrollo Infantil*), obtiene información relacionada a oportunidades de juego y aprendizaje, disciplina infantil, desarrollo de lenguaje del niño/a y estado emocional de su cuidador.

El presente documento contiene guías y definiciones a tener en consideración para el uso adecuado de las tablas que conforman la Base de Datos generada por la ENDI.

**Material relacionado:** Toda la documentación relacionada con la encuesta, así como la Base de Datos para descarga se puede encontrar en la siguiente dirección: [https://www.ecuadorencifras.gob.ec/encuesta\\_nacional\\_desnutricion\\_infantil/](https://www.ecuadorencifras.gob.ec/encuesta_nacional_desnutricion_infantil/)

## 2. Recomendaciones para el usuario

Para un adecuado manejo de la Base de Datos de la ENDI, se recomienda:

- Revisar la presente guía.
- Revisar los formularios y la estructura y relaciones de las tablas que conforman la Base de Datos.
- Identificar preguntas o indicadores que se desea analizar a partir de las variables presentes en las tablas.

- Utilizar el factor de expansión normalizado apropiado, según la tabla y variables que se estén utilizando.
- Los archivos de la Base de Datos se publican en las siguientes extensiones: .csv (Excel), .dta (STATA) y .rds (R). Se encuentran disponibles para descarga en el sitio web de la encuesta. Puede trabajar con el software estadístico de su preferencia.

### 3. Datos generales de la operación estadística

En el Cuadro 1, se presentan los datos generales de la ENDI<sup>1</sup>:

**Cuadro 1.** Información de la operación estadística

| N° | Dimensión                                   | Descripción  |
|----|---|--|
| 1  | <b>Nombre de la operación</b>               | <i>Encuesta Nacional sobre Desnutrición Infantil – ENDI 2023-2024</i>  |
| 2  | <b>Población Objetivo</b>                   | Todos los niños/as menores de 5 años   |
| 3  | <b>Unidad de observación</b>                | Niñas/os menores de 5 años   |
| 4  | <b>Unidad de análisis</b>                   | Niñas/os menores de 5 años   |
| 5  | <b>Cobertura geográfica</b>                 | Nacional, regional, provincial, urbano – rural   |
| 6  | <b>Tamaño de la muestra</b>                 | 22.250 viviendas visitadas   |
| 7  | <b>Viviendas efectivas</b>                  | 20.110 viviendas efectivas con menores de 5 años   |
| 8  | <b>Desagregación de la información</b>      | Nacional, Nacional Urbano, Nacional Rural y provincial   |
| 9  | <b>Comparabilidad de la serie histórica</b> | <i>Encuesta de Condiciones de Vida – ECV: 2006 y 2014</i><br><i>Encuesta Nacional de Salud y Nutrición – ENSANUT: 2012 y 2018</i><br><i>Encuesta Nacional de Desnutrición Infantil – ENDI: 2022-23</i> |

Fuente: Instituto Nacional de Estadística y Censos.

### 4. Formularios de captura de información

Debido a la variedad de temas presentes en cada formulario, se deberá identificar los individuos a los que se destinan las preguntas en las diferentes secciones y partes. Por ejemplo:

- Todos los miembros del hogar
- Mujeres en edad fértil
- Madres entre 10 a 49 años con hijos menores de 5 años
- Niñas/os menores de 3 años
- Niñas/os menores de 5 años

<sup>1</sup> Para profundizar en los puntos mencionados en el Cuadro 1 remítase al Documento Metodológico de la ENDI 2023-24



A continuación, se detalla el contenido y las temáticas abordadas en cada uno de los 3 formularios que componen la ENDI 2023-24.

### Formulario 1: Hogar

**Objetivo:** Recabar información sociodemográfica y económica de todos los miembros de hogar, estado nutricional de los menores, seguridad alimentaria y calidad del agua para beber.

#### Secciones que lo conforman:

**Carátula:** Ubicación geográfica y muestral, ubicación de la vivienda, datos del informante, resultado de la entrevista, personal responsable, fecha de la entrevista.

**Sección 1:** Registro de los miembros del hogar.

**Sección 2:** Actividades económicas para personas de 10 años y más.

**Sección 3:** Datos de la vivienda y el hogar.

**Sección 4:** Seguridad alimentaria.

**Sección 5:** Antropometría.

**Sección 6:** Prueba de anemia.

**Sección 7:** Prueba de E. coli y cloro.

Figura 1. Formulario del Hogar – f1

**INEC** INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS

GESTIÓN DE ESTADÍSTICAS PERMANENTES A HOGARES  
ENCUESTA NACIONAL DE DESNUTRICIÓN INFANTIL - ENDI 2023-2024

**1**

FORMULARIO  DE

**CONFIDENCIALIDAD:**  
LOS DATOS PROPORCIONADOS POR EL INFORMANTE SON Estrictamente confidenciales y serán utilizados únicamente con fines estadísticos de acuerdo al artículo 21 de la Ley de Estadística

**A. UBICACIÓN GEOGRÁFICA Y MUESTRAL**

1. Área Urbana  2. Área Rural

3. Provincia

4. Cantón

5. Cabecera cantonal o parroquial

6. Conglomerado

7. Zona

8. Sector

9. No. Secuencial de la vivienda original

10. Número de hogares en la vivienda

11. Hogar

12. Período

**B. RESULTADO DE LA ENTREVISTA**

**RESULTADO DE LA ENCUESTA**

1. Completa (efectiva)

2. Rechazo

3. Nadie en casa

4. Vivienda temporal

5. Vivienda desocupada

6. Vivienda en construcción

7. Vivienda inhabitable o destruida

8. Vivienda convertida en negocio

9. Vivienda sin menores de 5 años

10. Otra razón, cuál?

OBSERVACIONES:

**C. IDENTIFICACIÓN Y UBICACIÓN DE LA VIVIENDA EFECTIVA**

Manzana:  Edificio:

Calle:

Localidad:

Lote No.  Bloque No.  Patio No.

Piso No.  Casa No.  Depart. No.

**D. DATOS DEL HOGAR**

Número de miembros del hogar:

Nombre del informante calificado:  N° Telf.:

N° Cel.:

Escriba la altitud /metros sobre el nivel del mar (msnm):

**E. PERSONAL RESPONSABLE**

RESPONSABLE ZONAL: Ci:

SUPERVISOR: Ci:

ENCUESTADOR: Ci:

ANTROPOMETRISTA: Ci:

REVISOR DE CALIDAD: Ci:

**F. FECHA DE LA ENTREVISTA**

Fecha:  DÍA  MES  AÑO  2  0  2

Fuente: Instituto Nacional de Estadística y Censos.

### Formulario 2: Mujeres en Edad Fértil, salud de la niñez y lactancia materna

**Objetivo:** Recolectar información sobre historial de embarazos y nacimientos de las madres en edad fértil (10 a 49 años) que tienen hijos menores de 5 años al momento de la entrevista, así como sobre lactancia materna y salud en la niñez.

**Secciones que lo conforman:**

**Carátula:** Ubicación geográfica y muestral, ubicación de la vivienda, datos de la informante, resultado de la entrevista, personal responsable, fecha de la entrevista.

**Sección 1:** Características generales de la entrevistada.

**Sección 2:** Historia de embarazos y nacimientos.

**Sección 3:** Lactancia materna (niñas/os menores de 3 años).

**Sección 4:** Salud en la niñez (niñas/os menores de 5 años).

**Sección 5:** Programas de primera infancia (niñas/os menores de 5 años).

**Figura 2.** Formulario MEF – f2

**INEC** GESTIÓN DE ESTADÍSTICAS PERMANENTES A HOGARES  
ENCUESTA NACIONAL DE DESNUTRICIÓN INFANTIL - ENDI 2023- 2024

CONFIDENCIALIDAD:  
LOS DATOS PROPORCIONADOS POR EL INFORMANTE SON ESTRICTAMENTE CONFIDENCIALES Y SERÁN UTILIZADOS ÚNICAMENTE CON FINES ESTADÍSTICOS DE ACUERDO AL ARTÍCULO 21 DE LA LEY DE ESTADÍSTICA

**MUJERES EN EDAD FÉRTIL (MEF DE 10 A 49 AÑOS), SALUD DE LA NIÑEZ Y LACTANCIA MATERNA** 2

**A. UBICACIÓN GEOGRÁFICA Y MUESTRAL**

|   |  |  |
|---|--|--|
| 1. Área Urbana <input type="checkbox"/>                 | 2. Área Rural <input type="checkbox"/> | 9. No. Secuencial de la vivienda original <input type="text"/> |
| 3. Provincia <input type="text"/>                       | 4. Cantón <input type="text"/>         | 10. Número de hogares en la vivienda <input type="text"/>      |
| 5. Cabeceira cantonal o parroquial <input type="text"/> | 6. Conglomerado <input type="text"/>   | 11. Hogar <input type="checkbox"/>                             |
| 7. Zona <input type="text"/>                            | 8. Sector <input type="text"/>         | 12. Período <input type="text"/>                               |
|   |  | 13. Manzana <input type="text"/>                               |
|   |  | 14. Edificio <input type="text"/>                              |

INFORMANTE: DIRECTO

**B. DATOS DE LA INFORMANTE**

| COD. PER. FORM. DEL HOGAR | PRIMER NOMBRE        | SEGUNDO NOMBRE       | PRIMER APELLIDO      | SEGUNDO APELLIDO     |
|---------------------------|----------------------|----------------------|----------------------|----------------------|
| <input type="text"/>      | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |

**C. FECHA DE LA ENTREVISTA**

DÍA   MES   AÑO  2  0  2

**D. PERSONAL RESPONSABLE**

|                     |     |                      |
|---------------------|-----|----------------------|
| SUPERVISOR:         | CL. | <input type="text"/> |
| ENCUESTADOR:        | CL. | <input type="text"/> |
| REVISOR DE CALIDAD: | CL. | <input type="text"/> |

**E. RESULTADO DE LA ENTREVISTA**

|                                     |                      |
|-------------------------------------|----------------------|
| 1. COMPLETA (efectiva)              | <input type="text"/> |
| 2. RECHAZO                          | <input type="text"/> |
| 3. AUSENTE                          | <input type="text"/> |
| 4. NO TIENE HIJOS MENORES DE 5 AÑOS | <input type="text"/> |
| 5. OTRA CUÁL?                       | <input type="text"/> |

FIN DE LA ENTREVISTA

**RECUERDE: " Antes de iniciar la entrevista con la MEF, procure lograr el máximo de privacidad posible"**

**SECCIÓN I. CARACTERÍSTICAS GENERALES DE LA ENTREVISTADA**

ENCUESTADOR/A: En el caso que corresponda, seleccione o registre la respuesta dada por la informante

|     |  |   |
|-----|--|---|
| 100 | ¿Cuál es su fecha de nacimiento?<br><small>(Solicite la cédula de identidad o partida de nacimiento)</small> | DÍA <input type="text"/> <input type="text"/> MES <input type="text"/> <input type="text"/> AÑO <input type="text"/> <input type="text"/> |
| 101 | Entonces, ¿Cuántos años cumplidos tiene?   | Años cumplidos <input type="text"/> <input type="text"/>  |

**Fuente:** Instituto Nacional de Estadística y Censos.

**Formulario 3: Desarrollo infantil para niños/as menores de 5 años**

**Objetivo:** Obtener información de los menores de 5 años sobre oportunidades de juego y aprendizaje en el hogar, formas de disciplina infantil, desarrollo de lenguaje y estado emocional de la madre o cuidador principal.

**Secciones que lo conforman:**

**Carátula:** Ubicación geográfica y muestral, ubicación de la vivienda, datos del informante y del niño seleccionado, resultado de la entrevista, personal responsable, fecha de la entrevista.

**Sección 0:** Selección del niño/a menor de 5 años.

**Sección 1:** Oportunidades de juego en el hogar para menores de 5 años.

**Sección 2:** Disciplina infantil.

**Sección 3:** Lenguaje de niños/as de 12 a 18 meses (1 año a 1 año y 5 meses).

**Sección 4:** Lenguaje de niños/as de 19 a 30 meses (1 año 7 meses a 2 años y 6 meses).

**Sección 5:** Lenguaje de niños/as de 31 a 42 meses (2 años 7 meses a 3 años y 6 meses).

**Sección 6:** Lenguaje – Test de vocabulario en imágenes PEABODY – TVIP para niños/as de 43 a 59 meses (3 años 7 meses a 4 años y 11 meses).

**Sección 7:** Estado emocional de la madre o cuidador principal (en ausencia permanente de la madre) de niñas/os menores de 5 años.

**Figura 3.** Formulario de Desarrollo Infantil – f3

**INEC** **GESTIÓN DE ESTADÍSTICAS PERMANENTES A HOGARES**  
**ENCUESTA NACIONAL DE DESNUTRICIÓN INFANTIL**  
**ENDI**

**DESARROLLO INFANTIL** 3  
**PARA NIÑOS/AS MENORES DE 5 AÑOS**

CONFIDENCIALIDAD:  
LOS DATOS PROPORCIONADOS POR EL INFORMANTE SON ESTRICTAMENTE CONFIDENCIALES Y SERÁN UTILIZADOS ÚNICAMENTE CON FINES ESTADÍSTICOS DE ACUERDO AL ARTÍCULO 21 DE LA LEY DE ESTADÍSTICA

**A. UBICACIÓN GEOGRÁFICA Y MUESTRAL**

|   |  |                                      |  |
|---|--|--------------------------------------|--|
| 1. Área Urbana <input type="checkbox"/> | 2. Área Rural <input type="checkbox"/> | 7. Zona                              |  |
| 3. Provincia                            |  | 8. Sector                            |  |
| 4. Cantón                               |  | 9. No. Secuencial de la vivienda     |  |
| 5. Cabeecera cantonal o parroquial      |  | 10. Número de hogares en la vivienda |  |
| 6. Conglomerado                         |  | 11. Hogar                            |  |
|   |  | 12. Período                          |  |

**B. UBICACIÓN DE LA VIVIENDA**

Manzana:  Edificio:

**C.1 DATOS DEL NIÑO/A SELECCIONADO**

| COD. PER. | EDAD |       | PRIMER NOMBRE | SEGUNDO NOMBRE | PRIMER APELLIDO | SEGUNDO APELLIDO |
|-----------|------|-------|---------------|----------------|-----------------|------------------|
|           | Años | Meses |               |                |                 |                  |
|           |      |       |               |                |                 |                  |

**C.2 DATOS DE LA PERSONA QUE INFORMA (MADRE O CUIDADOR PRINCIPAL)**

| COD. PER. INF. | Relación con el niño/a | PRIMER NOMBRE | SEGUNDO NOMBRE | PRIMER APELLIDO | SEGUNDO APELLIDO |
|----------------|------------------------|---------------|----------------|-----------------|------------------|
|                |                        |               |                |                 |                  |

RELACION CON EL NIÑO/A: (Padre=1, madre=2, hermana=3, abuelo/a=4, tío/a=5, padrastro o madrastra=6, otros familiares=7 y otros no familiares=8)

**D. RESULTADO DE LA ENTREVISTA**

|                        |  |
|------------------------|--|
| 1. COMPLETA (efectiva) |  |
| 2. RECHAZO             |  |
| 3. AUSENTE             |  |
| 4. OTRA ¿CUÁL?         |  |

**E. PERSONAL RESPONSABLE**

SUPERVISOR:

ENCUESTADOR:

REVISOR DE CALIDAD:

**F. FECHA DE LA ENTREVISTA**

DÍA  MES  AÑO

**SECCIÓN 0. SELECCIÓN DEL NIÑO/A MENOR DE 5 AÑOS**

ENCUESTADOR/A: Registre con el código 1 en el casillero del niño/a seleccionado de acuerdo a una selección aleatoria. Añote el código 2 en cada uno de los casilleros de los niños no seleccionados.

| No. | 2. COD. PER. | 3. NOMBRE DE NIÑOS/AS MENORES DE 5 AÑOS | 4. FECHA DE NACIMIENTO (Solicite la cédula de identidad o partida de nacimiento) | 5. ¿CUÁNTOS AÑOS Y MESES CUMPLIDOS TIENE (...)       | 6. SEL. SI.1 No. 2 |
|-----|--------------|---|--|--|--------------------|
|     |              |   |  |  |                    |
| 1   |              |   | DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/>       | AÑOS <input type="text"/> MESES <input type="text"/> |                    |
| 2   |              |   | DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/>       | AÑOS <input type="text"/> MESES <input type="text"/> |                    |
| 3   |              |   | DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/>       | AÑOS <input type="text"/> MESES <input type="text"/> |                    |
| 4   |              |   | DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/>       | AÑOS <input type="text"/> MESES <input type="text"/> |                    |
| 5   |              |   | DÍA <input type="text"/> MES <input type="text"/> AÑO <input type="text"/>       | AÑOS <input type="text"/> MESES <input type="text"/> |                    |

**1. ¿CUÁNTOS NIÑOS/AS MENORES DE 5 AÑOS HAY EN ESTE HOGAR?**

Fuente: Instituto Nacional de Estadística y Censos.



## 5. Descripción de la Base de Datos

### 5.1 Tablas relacionales

La Base de Datos ENDI 2023-24 está conformada por 6 sets de datos a los que denominaremos tablas. Las tablas están relacionadas entre sí, y sus nombres son:

1. *Personas*
2. *Hogar*
3. *MEF*
4. *Lactancia*
5. *Salud en la niñez*
6. *Desarrollo infantil*

Cada tabla tiene el objetivo de proporcionar al usuario los datos con un nivel de granularidad que le permita realizar análisis específicos relacionados con las temáticas investigadas considerando el alcance en las desagregaciones y cobertura habilitadas por la operación estadística.

En el Cuadro 2, se detalla la información<sup>2</sup> relevante sobre cada una de las tablas que conforman la Base de Datos:

**Cuadro 2.** Descripción de la Base de Datos

| Criterio                                 | Descripción  |
|--|--|
| <b>Tabla 1 – Personas</b>                |  |
| <b>Nombre del archivo (extensión):</b>   | <i>BDD_ENDI_R2_f1_personas</i><br>(.csv / .dta / .rds)   |
| <b>Unidad de análisis (granularidad)</b> | Persona miembro de hogar   |
| <b>Formulario asociado:</b>              | Formulario del Hogar – f1  |
| <b>Secciones del formulario:</b>         | Sección 1: <i>Registro e información de los miembros</i><br>Sección 2: <i>Actividades económicas</i><br>Sección 5: <i>Antropometría</i><br>Sección 6: <i>Prueba de Anemia</i>  |
| <b>Número de filas (registros):</b>      | <b>93.242</b> miembros de hogar  |
| <b>Número de columnas (variables):</b>   | <b>124</b> columnas desglosadas en:<br>Llave primaria: <b>1</b> ( <i>id_per</i> )<br>*Llave primaria secundaria: <b>1</b> ( <i>id_mef</i> )<br>Llaves foráneas: <b>1</b> ( <i>id_hogar</i> )<br>Inform. geográfica y muestral: <b>11</b> (incluye factor de expansión <i>fexp</i> )<br>Sección 1: <b>30</b> (provienen de 21 preguntas más <i>persona</i> o número de miembro dentro del hogar)<br>Sección 2: <b>33</b> (provienen de 26 preguntas)<br>Sección 5: <b>17</b> (provienen de 7 preguntas)<br>Sección 6: <b>10</b> (provienen de 6 preguntas más <i>altitud</i> para ajustar hemoglobina)<br>Calculadas: <b>20</b> (en función de variables preexistentes) |

<sup>2</sup> Las descripciones aquí presentadas corresponden a los datos de territorio continental. Los correspondientes a Galápagos se publican por separado con sus propios factores de expansión y sin variables de desagregación, pero guardan la misma estructura.

|  |   |
|--|---|
| <b>Variables para match o join</b>       | <i>id_per, id_hogar, id_mef</i>   |
| <b>Tabla 2 – Hogar</b>                   |   |
| <b>Nombre del archivo (extensión):</b>   | <i>BDD_ENDI_R2_f1_hogar (.csv / .dta / .rds)</i>  |
| <b>Unidad de análisis (granularidad)</b> | Hogar   |
| <b>Formulario asociado:</b>              | Formulario del Hogar – f1   |
| <b>Secciones del formulario:</b>         | Sección 3: <i>Datos de vivienda y hogar</i><br>Sección 4: <i>Seguridad alimentaria</i><br>Sección 7: <i>Prueba de E. coli y cloro residual</i>  |
| <b>Número de filas (registros):</b>      | <b>20.186</b> hogares efectivos   |
| <b>Número de columnas (variables):</b>   | <b>133</b> columnas desglosadas en:<br>Llave primaria: <b>1</b> ( <i>id_hogar</i> )<br>Llaves foráneas: <b>0</b><br>Inform. geográfica y muestral: <b>10</b> (incluye factor de expansión <i>fexp</i> )<br>Sección 3: <b>41</b> (proviene de 29 preguntas)<br>Sección 4: <b>20</b> (proviene de 4 preguntas)<br>Sección 7: <b>53</b> (proviene de 28 preguntas)<br>Calculadas: <b>8</b> (en función de variables preexistentes) |
| <b>Variables para match o join</b>       | <i>id_hogar</i>   |
| <b>Tabla 3: MEF</b>                      |   |
| <b>Nombre del archivo (extensión):</b>   | <i>BDD_ENDI_R2_f2_mef (.csv / .dta / .rds)</i>  |
| <b>Unidad de análisis (granularidad)</b> | Madre entre 10 a 49 años con hijo/s menores de 5 años   |
| <b>Formulario asociado:</b>              | Formulario MEF (Mujer en Edad Fértil) – f2  |
| <b>Secciones del formulario:</b>         | Sección 1: <i>Características generales de la madre</i><br>Sección 2: <i>Historia de embarazos y nacimientos</i>  |
| <b>Número de filas (registros):</b>      | <b>19.846</b> mujeres en edad fértil miembros de hogar que son madres de al menos un menor de 5 años  |
| <b>Número de columnas (variables):</b>   | <b>289</b> columnas desglosadas en:<br>Llave primaria: <b>1</b> ( <i>id_mef</i> )<br>Llaves foráneas: <b>1</b> ( <i>id_hogar</i> )<br>Inform. geográfica y muestral: <b>10</b> (incluye factor de expansión <i>fexp</i> )<br>Sección 1: <b>19</b> (proviene de 8 preguntas)<br>Sección 2: <b>258</b> (proviene de 36 preguntas)<br>Calculadas: <b>0</b>   |
| <b>Variables para match o join</b>       | <i>id_mef, id_hogar</i>   |

| <b>Tabla 4: Lactancia</b>                |  |
|--|--|
| <b>Nombre del archivo (extensión):</b>   | <i>BDD_ENDI_R2_f2_lactancia (.csv / .dta / .rds)</i>   |
| <b>Unidad de análisis (granularidad)</b> | Menor de 3 años miembro de hogar   |
| <b>Formulario asociado:</b>              | Formulario MEF (Mujer en Edad Fértil) – f2   |
| <b>Secciones del formulario:</b>         | Sección 3: <i>Lactancia materna</i>  |
| <b>Número de filas (registros):</b>      | <b>12.141</b> menores de 3 años miembros de hogar que conviven con su madre  |
| <b>Número de columnas (variables):</b>   | <b>71</b> columnas desglosadas en:<br>Llave primaria: <b>1</b> ( <i>id_per</i> )<br>*Llave primaria secundaria: <b>1</b> ( <i>id_mef_per</i> )<br>Llaves foráneas: <b>2</b> ( <i>id_hogar, id_mef</i> )<br>Inform. geográfica y muestral: <b>10</b> (incluye factor de expansión <i>fexp_lm</i> )<br>Sección 3: <b>57</b> (proviene de 25 preguntas)<br>Calculadas: <b>0</b>   |
| <b>Variables para match o join</b>       | <i>id_per, id_mef_per, id_hogar, id_mef</i>  |
| <b>Tabla 5: Salud en la niñez</b>        |  |
| <b>Nombre del archivo (extensión):</b>   | <i>BDD_ENDI_R2_f2_salud_ninez (.csv / .dta / .rds)</i>   |
| <b>Unidad de análisis (granularidad)</b> | Menor de 5 años cuya madre es miembro de hogar   |
| <b>Formulario asociado:</b>              | Formulario MEF (Mujer en Edad Fértil) – f2   |
| <b>Secciones del formulario:</b>         | Sección 4: <i>Salud en la niñez</i><br>Sección 5: <i>Programas de primera infancia</i>   |
| <b>Número de filas (registros):</b>      | <b>22.188</b> menores de 5 años cuya madre es miembro de hogar   |
| <b>Número de columnas (variables):</b>   | <b>413</b> columnas desglosadas en:<br>Llave primaria: <b>1</b> ( <i>id_per</i> <sup>3</sup> )<br>*Llaves primarias secundarias: <b>2</b> ( <i>id_mef_per, id_hijo_ord</i> )<br>Llaves foráneas: <b>2</b> ( <i>id_hogar, id_mef</i> )<br>Inform. geográfica y muestral: <b>10</b> (incluye factor de expansión <i>fexp</i> )<br>Sección 4: <b>379</b> (proviene de 102 preguntas, más el número de orden del hijo <i>ord_hijo</i> con respecto a su madre - último o penúltimo)<br>Sección 5: <b>19</b> (proviene de 15 preguntas)<br>Calculadas: <b>0</b> |
| <b>Variables para match o join</b>       | <i>id_per, id_mef_per, id_hogar, id_mef</i>  |

<sup>3</sup> Explicación adicional sobre las llaves primarias de esta tabla en el apartado **5.4 Proceso para unir tablas relacionadas**

| <b>Tabla 6: Desarrollo infantil</b>      |  |
|--|--|
| <b>Nombre del archivo (extensión):</b>   | <i>BDD_ENDI_R2_f3_desarrollo_inf (.csv / .dta / .rds)</i>  |
| <b>Unidad de análisis (granularidad)</b> | Menor de 5 años miembro de hogar   |
| <b>Formulario asociado:</b>              | Formulario de Desarrollo Infantil – f3   |
| <b>Secciones del formulario:</b>         | Sección 0: <i>Selección del menor de 5 años</i><br>Sección 1: <i>Oportunidades de juego</i><br>Sección 2: <i>Disciplina infantil</i><br>Sección 3: <i>Lenguaje - 1 año a 1 año 6 meses</i><br>Sección 4: <i>Lenguaje - 1 año 7 meses a 2 años 6 meses</i><br>Sección 5: <i>Lenguaje - 2 años 7 meses a 3 años 6 meses</i><br>Sección 6: <i>Lenguaje - 3 años 7 meses a 4 años 11 meses</i><br>Sección 7: <i>Estado emocional del cuidador</i>  |
| <b>Número de filas (registros):</b>      | <b>9.836</b> menores de 5 años miembros de hogar <sup>4</sup>  |
| <b>Número de columnas (variables):</b>   | <b>393</b> columnas desglosadas en:<br>Llave primaria: <b>1</b> ( <i>id_per</i> )<br>Llaves foráneas: <b>2</b> ( <i>id_hogar, id_cuid</i> )<br>Inform. geográfica y muestral: <b>10</b> (incluye factor de expansión <i>fexp_di</i> )<br>Sección 0: <b>7</b> (provienen de 6 preguntas más <i>relacion_cuidadora</i> que indica la relación del cuidador/a con el menor)<br>Sección 1: <b>53</b> (provienen de 6 preguntas)<br>Sección 2: <b>23</b> (provienen de 3 preguntas)<br>Sección 3: <b>51</b> (provienen de 51 preguntas)<br>Sección 4: <b>51</b> (provienen de 51 preguntas)<br>Sección 5: <b>51</b> (provienen de 51 preguntas)<br>Sección 6: <b>136</b> (provienen de 11 preguntas, una de las cuales corresponde a las respuestas del Test de Vocabulario en Imágenes Peabody)<br>Sección 7: <b>8</b> (de 1 pregunta)<br>Calculadas: <b>0</b> |
| <b>Variables para match o join</b>       | <i>id_per, id_hogar, id_cuid</i>   |

**Fuente:** Instituto Nacional de Estadística y Censos.

## 5.2 Variables

Cada una de las variables existentes se encuentra definida en los diccionarios de variables adjuntos a la Base de Datos publicada.

En el Cuadro 2 se mencionaron los siguientes tipos de variables: llave primaria, llaves foráneas, variables correspondientes a cada pregunta de los formularios, variables calculadas, y variables de información geográfica y muestral. A continuación, se describe cada tipo.

<sup>4</sup> La muestra es menor debido a que este formulario se levanta solo durante la segunda mitad de la ronda de encuesta, y además solo se selecciona 1 niño por hogar.

### Llave primaria

Es aquella que cumple dos funciones: 1. Identifica de manera única cada registro (fila) en esa tabla y 2. Sirve para hacer match o unificación de tablas cuando aparece en otras tablas. Se caracteriza porque no puede contener valores duplicados ni valores faltantes. Cada tabla posee su llave primaria. En la Base de Datos estas variables se codifican como: **id\_per, id\_mef, id\_hogar, id\_mef\_per, id\_hijo\_ord**

### Llaves foráneas

Son aquellas variables presentes en una tabla que, al ser llave primaria de otra tabla, permiten vincular ambas tablas mediante *join*. Permiten mantener la integridad referencial entre tablas. Se caracterizan porque en su rol de llave foránea pueden contener valores duplicados y, dependiendo del caso, pueden aceptar valores faltantes. Estas variables son: **id\_hogar, id\_mef, id\_cuid**

### Variables correspondientes a las preguntas de los formularios

Se corresponden directamente con cada una de las preguntas de los formularios, por lo que se las puede ubicar dentro de cada sección siguiendo la estructura de los formularios. Sus nombres codificados fueron construidos y estructurados considerando los criterios que se describen en el Cuadro 3.

**Cuadro 3.** Estructura de codificación para variables que provienen de preguntas

| Formulario | Sección | Parte                              | Núm. pregunta desde - hasta | Desglose de pregunta *(en caso de existir) | Codificación completa (concatenación) desde - hasta |
|------------|---------|------------------------------------|-----------------------------|--|---|
| f1         | s1      | -                                  | 1 - 18                      | 1<br>2<br>3<br>.<br>.<br>.<br>.            | f1_s1_1 ... f1_s1_18                                |
|            | s2      |                                    | 1 - 26                      |  | f1_s2_1 ... f1_s2_26                                |
|            | s3      |                                    | 1 - 29                      |  | f1_s3_1 ... f1_s3_29_d                              |
|            | s4      |                                    | 1 - 4                       |  | f1_s4_1_1 ... f1_s4_4_f                             |
|            | s5      |                                    | 1 - 7                       |  | f1_s5_1 ... f1_s5_7                                 |
|            | s6      |                                    | 1 - 6                       |  | f1_s6_1 ... f1_s6_6                                 |
|            | s7      |                                    | 1 - 23                      |  | f1_s7_1 ... f1_s7_23                                |
| f2         | s1      | -                                  | 0 - 107                     | a<br>b<br>c<br>.<br>.                      | f2_s1_0 ... f2_s1_107                               |
|            | s2      |                                    | 200 - 235                   |  | f2_s2_200 ...<br>f2_s2_235_f_anios_15               |
|            | s3      |                                    | 301 - 324                   |  | f2_s3_301_a_cod ... f2_s3_324                       |
|            | s4      | a, b, c,<br>d, e, f,<br>g, h, i, j | 400a - 501                  | v  | f2_s4_400a ... f2_s4j_501_b_v                       |
|            | s5      | -                                  | 500 - 514                   | dia<br>mes/es<br>anio/s                    | f2_s5_500 ... f2_s5_514                             |
| f3         | s0      | -                                  | 1                           | cod  | f3_s0_1_cod ... f3_s0_1c_meses                      |
|            | s1      | -                                  | 100 - 105                   |  | f3_s1_100_a ... f3_s1_105_g1_5                      |
|            | s2      | a, b                               | 200 - 202                   | f3_s2a_200_a ... f3_s2b_202                |   |
|            | s3      | -                                  | 300 - 350                   | .  | f3_s3_300 ... f3_s3_350                             |
|            | s4      | -                                  | 400 - 450                   | .  | f3_s4_400 ... f3_s4_450                             |
|            | s5      | -                                  | 500 - 550                   | .  | f3_s5_500 ... f3_s5_550                             |
|            | s6      | -                                  | 600 - 610                   | etc.                                       | f3_s6_600 ... f3_s6_610                             |
|            | s7      | -                                  | 700                         |  | f3_s7_700_a ... f3_s7_700_h                         |

Fuente: Instituto Nacional de Estadística y Censos.

## Variables calculadas

Son variables generadas a partir de procesar variables primarias levantadas en la encuesta, ya que no fueron recopiladas o captadas directamente con el instrumento de recolección. Son apropiadas para ser utilizadas en reportes y desagregaciones para el cálculo de indicadores, y son:

- **etnia:** agrega las categorías de identificación de las personas según su cultura y sus costumbres agrupando las etnias minoritarias, calculada a partir de *f1\_s1\_10* (pregunta 10 del f1 - Sección 1).
- **edaddias:** edad de las personas expresada en días, calculada en función de la fecha de nacimiento de la persona y la fecha de entrevista.
- **grupo\_edad\_nin:** grupo de edad al que pertenece el menor de 5 años, expresado en meses.
- **nivins\_mef:** nivel de instrucción de la mujer en edad fértil, agrupado, calculado a partir de nivel de instrucción y años aprobados.
- **quintil:** quintil de ingreso per cápita calculado en función de los ingresos reportados por el informante en f1 – Sección II (actividades económicas).
- **pobreza:** indica si en ese momento la persona está bajo el umbral de pobreza por ingreso per cápita, calculado en función de los ingresos reportados por el informante en f1 – Sección II (actividades económicas).
- **nbi\_1:** indica la condición de *Necesidades Básicas Insatisfechas*, calculado en función de varias dimensiones que indican privaciones en capacidad económica, acceso a educación, acceso a vivienda, acceso a servicios básicos, así como hacinamiento.
- **dcronica\_2, dcronica, dcronica2\_5:** indicadores de la condición de desnutrición crónica infantil en cada grupo de edad, calculadas a partir de las variables del f1 – Sección 5 (antropometría).
- **dglobal\_2, dglobal, dglobal2\_5:** indicadores de la condición de desnutrición global infantil en cada grupo de edad, calculadas a partir de las variables del f1 – Sección 5 (antropometría).
- **daguda\_2, daguda, daguda2\_5:** indicadores de la condición de desnutrición aguda infantil en cada grupo de edad, calculadas a partir de las variables del f1 – Sección 5 (antropometría).
- **ane6\_59, ane6\_59\_new, ane6\_23, ane6\_23\_new":** indicadores de la condición de anemia infantil en cada grupo de edad, calculadas a partir de las variables del f1 – Sección 6 (prueba de anemia) y de la *altitud*.
- **cloro\_v, cloro\_v\_new, cloro\_f, cloro\_f\_new, ecoli\_v, ecoli\_vnew, ecoli\_f, ecoli\_f\_new:** indicadores de la condición de calidad de agua que beben los hogares, calculadas a partir de las variables del f1 – Sección 7 (prueba de E. coli y cloro).

## Variables de información geográfica y muestral

Contienen información relacionada con la ubicación geográfica de la vivienda, así como información relativa al muestreo. Están presentes en todas las tablas. Algunas se ubican en las carátulas de los formularios. Se detallan a continuación:

- **id\_upm:** variable identificadora de cada Unidad Primaria de Muestreo (UPM) codificada.
- **id\_viv:** variable identificadora de cada vivienda codificada.
- **fecha\_anio, fecha\_mes, fecha\_dia:** variables que indican el momento del tiempo en el que fue realizada la entrevista de acuerdo al plan de muestreo.
- **area:** variable que permite desagregar por sector urbano o rural.
- **prov:** variable que permite desagregar por provincia (excluye la provincia de Galápagos en la Base de Datos de territorio continental).



- **region:** variable que permite desagregar por región natural del Ecuador (Costa, Sierra y Amazonía, excluye región *Insular* en la Base de Datos de territorio continental). Se construye a partir de la variable *prov*.
- **parr\_pri:** variable que permite identificar si el hogar pertenece o no a una parroquia priorizada según desnutrición crónica infantil.
- **estrato:** variable indicadora del estrato socioeconómico del hogar, construida en diseño muestral y que para los fines de ENDI es una desagregación socioeconómica y territorial que define estratos dentro de cada área de residencia y provincia o dominio autorepresentado.
- **fexp, fexp\_lm, fexp\_di:** variables que contienen el factor de expansión normalizado que se debe aplicar, dependiendo de las tablas que se estén utilizando: *fexp\_lm* para la tabla de lactancia materna, *fexp\_di* para la tabla de desarrollo infantil, y *fexp* para el resto de tablas. El factor de expansión normalizado se construye multiplicando el factor de expansión que le corresponde a cada caso, por un factor de normalización. Para explicación adicional revisar el apartado: [7.2 Factores de expansión para encuestas](#).

### 5.3 Número de registros por sección

Es importante recordar y aclarar que no todas las secciones de los formularios registran la misma cantidad de casos, ya que cada una corresponde a un tema y/o grupo de edad específico, por lo que en las tablas existen casos que no son completos o que no presentan valores en absolutamente todas las variables, ni en todas las tablas. Ejemplos:

- En la tabla *BDD\_ENDI\_R2\_f2\_lactancia* (formulario 2, sección 3) que contiene información sobre lactancia materna, solo se encontrarán los registros que cumplan la condición del grupo de edad: niñas/os menores de 3 años, y cuyas madres accedieron a proporcionar la información (no rechazo).
- En la tabla *BDD\_ENDI\_R2\_f3\_desarrollo\_inf* (formulario 3) que contiene información sobre desarrollo infantil, solo se encontrarán los registros de los últimos 6 meses de encuesta y correspondientes a un menor de 5 años que fue seleccionado por cada hogar levantado en ese periodo. Las secciones 3, 4, 5 y 6 de desarrollo de lenguaje, segmentadas por grupo de edad, contienen información únicamente para los niños que cumplen el criterio de su grupo de edad.

Cabe mencionar además que, particularmente en la tabla *BDD\_ENDI\_R2\_f2\_salud\_ninez*, que contiene información sobre salud de las niñas/os menores de 5 años, en lo que corresponde al formulario 2, sección 4, partes A - F se encontrará una pequeña proporción de registros de niños/as que no son miembros de hogar debido a que en esas partes se incluye información de nacidos vivos menores de 5 años que han fallecido, o que no conviven con su madre miembro de hogar por otro motivo. Se los puede identificar con las variables que indican si el menor está vivo y si es miembro; además, al no ser miembros de hogar y no tener asignado un número de miembro dentro del hogar, su valor en la variable *id\_per* termina en 77.

### 5.4 Proceso para unir tablas relacionadas

Al trabajar con una tabla, dependiendo de la necesidad del investigador puede surgir la necesidad de traer información que se encuentra en alguna de las otras tablas de la Base de Datos, para este fin se deben utilizar las variables para *match* que se encuentran en cada una de las tablas. Por ejemplo:

**Ejemplo 1.** Si se encuentra trabajando con la tabla *BDD\_ENDI\_R2\_f1\_personas* y requiere variables de la tabla *BDD\_ENDI\_R2\_f1\_hogar*, en este caso debe utilizar el

identificador *id\_hogar* para realizar la unión de estas tablas. El factor de ponderación a utilizar es *fexp* que se encuentra en la tabla de personas.

**Ejemplo 2.** Si se encuentra trabajando con la tabla *BDD\_ENDI\_R2\_f2\_lactancia* y requiere utilizar información que se encuentra en la tabla de personas el identificador a utilizar para realizar esta unión es *id\_per*. El factor que se utiliza es *fexp\_lm* ya que el análisis se está realizando con el enfoque en la tabla de Lactancia Materna.

**Recuerde:** Las tablas *BDD\_ENDI\_R2\_f1\_personas*, *BDD\_ENDI\_R2\_f1\_hogar*, *BDD\_ENDI\_R2\_f2\_mef* y *BDD\_ENDI\_R2\_f2\_salud\_ninez* comparten la utilización de mismo factor de expansión: *fexp*. Mientras que *BDD\_ENDI\_R2\_f2\_lactancia* tiene su propio factor de expansión llamado *fexp\_lm*. De igual manera *BDD\_ENDI\_R2\_f2\_desarrollo\_inf* tienen su propio factor de expansión llamado *fexp\_di*.

En el Cuadro 4 se proporciona una explicación acerca de la forma en que las tablas se relacionan entre sí mediante las variables para *match* tales como llaves primarias y foráneas.

**Cuadro 4.** Tipo de relación entre tablas de la Base de Datos

| Relaciones      | Hogar  | MEF  | Lactancia   | Salud en la niñez   | Desarrollo infantil  |
|-----------------|--|--|---|---|--|
| <b>Personas</b> | <b>Varios-a-uno</b> (varios miembros a 1 hogar) mediante <b>id_hogar</b> | <b>Uno-a-uno</b> (1 mujer en edad fértil a 1 madre en edad fértil) mediante <b>id_mef</b> (también se podría utilizar su equivalente <b>id_per</b> ) | <b>Uno-a-uno</b> (1 <3años a 1 <3años) mediante <b>id_per</b><br><br><b>Uno-a-uno</b> (1 madre en edad fértil a 1 <3años) mediante <b>id_mef</b> (para lactancia solo se selecciona un <3años por cada madre) | <b>Uno-a-uno</b> (1 <5años a 1 <5años) mediante <b>id_per</b><br><br><b>Uno-a-varios</b> (1 madre en edad fértil a varios hijos <5años) mediante <b>id_mef</b> (para salud niñez se pueden seleccionar uno o dos hijos <5años por cada madre) | <b>Uno-a-uno</b> (1 <5años a 1 <5años) mediante <b>id_per</b><br><br><b>Uno-a-uno</b> (1 miembro de hogar cuidador a 1 <5años) mediante <b>id_per</b> equivalente a <b>id_cuid</b> (para desarrollo infantil se selecciona un <5años por cada hogar) |
| <b>Hogar</b>    |  | <b>Uno-a-varios</b> (1 hogar a varias madres en edad fértil) mediante <b>id_hogar</b>  | <b>Uno-a-varios</b> (1 hogar a varios <3años) mediante <b>id_hogar</b>  | <b>Uno-a-varios</b> (1 hogar a varios <5años) mediante <b>id_hogar</b>  | <b>Uno-a-uno</b> (1 hogar a 1 <5años) mediante <b>id_hogar</b> (para desarrollo infantil se selecciona solo un <5años por cada hogar)  |
| <b>MEF</b>      |  |  | <b>Uno-a-uno</b> (1 madre en edad fértil a 1 <3años) mediante <b>id_mef</b> (para lactancia solo se selecciona un <3años por cada madre)  | <b>Uno-a-varios</b> (1 madre en edad fértil a varios hijos <5años) mediante <b>id_mef</b> (para salud niñez se pueden seleccionar uno o dos hijos <5años por cada madre)  | <b>Uno-a-uno</b> (1 madre en edad fértil cuidadora a 1 <5años) mediante <b>id_mef</b> equivalente a <b>id_cuid</b> (para desarrollo infantil se selecciona un  |

|                   | Hogar | MEF | Lactancia | Salud en la niñez  | Desarrollo infantil  |
|-------------------|-------|-----|-----------|--|--|
| Lactancia         |       |     |           | Uno-a-uno (1 <3años a 1 <5años) mediante <b>id_per</b> o <b>id_mef_per</b> (el <3años en lactancia equivale al <5años en salud niñez que es menor de 3 años) | <5años por cada hogar)<br>Uno-a-uno (1 <3años a 1 <5años) mediante <b>id_per</b> (el <3años en lactancia equivale al <5años en desarrollo infantil que es menor de 3 años) |
| Salud en la niñez |       |     |           |  | Uno-a-uno (1 <5años a 1 <5años) mediante <b>id_per</b> (el <5años en salud niñez equivale al <5años cuando fue seleccionado en desarrollo infantil)                        |

Fuente: Instituto Nacional de Estadística y Censos.

El tipo de unión (*join* o *merge*) a utilizar depende del objetivo final de análisis. Según el caso será apropiado utilizar *left*, *right*, *inner* o *full/outer join* (especialmente si se está trabajando con lenguajes de análisis y manipulación de datos como R (*dplyr*), Python (*Pandas*) o SQL). En algunos casos se obtiene el mismo resultado utilizando varios tipos de unión, pero en otros el resultado es distinto. Esto debe ser analizado en cada caso particular, conjuntamente con el tipo de relación que hay entre las tablas (uno-a-uno, uno-a-varios, o varios-a-uno).

**Nota con respecto a la tabla *BDD\_ENDI\_R2\_f2\_salud\_ninez*:** *id\_per* funciona como identificador único y también para relaciones entre tablas en niños/as que son miembros de hogar (conviven con su madre). Pero si se considera también a niños/as que no son miembros (presentes únicamente en la tabla de Salud en la niñez) *id\_per* podría contener duplicados; en su lugar como identificador único de cada registro se recomienda usar *id\_hijo\_ord*, pero *id\_hijo\_ord* no se utiliza para relaciones al no estar presente en otras tablas.

## 5.5 Clasificaciones y Nomenclaturas

En el Cuadro 5 se enlistan las clasificaciones y nomenclaturas empleadas en la Base de Datos y se señalan las variables que las aplican.

**Cuadro 5.** Clasificaciones y nomenclaturas utilizadas

| N° | Nomenclaturas y clasificaciones usadas    | Versión / revisión | Variables que aplican clasificador | Aparecen en la tabla:  |
|----|---|--------------------|------------------------------------|--|
| 1  | Clasificador Geográfico Estadístico – DPA | 2021               | Provincia ( <i>prov</i> )          | <i>BDD_ENDI_R2_f1_personas</i><br><i>BDD_ENDI_R2_f1_hogar</i><br><i>BDD_ENDI_R2_f2_mef</i><br><i>BDD_ENDI_R2_f2_lactancia</i><br><i>BDD_ENDI_R2_f2_salud_ninez</i> |

Fuente: Instituto Nacional de Estadística y Censos.

**Nota:** Para más información revise el *Documento Metodológico* de la operación estadística y los documentos de clasificaciones internacionalmente homologadas.

## 6. Tratamiento previo a la publicación

### 6.1 Anonimización

Con el fin de garantizar la confidencialidad estadística y resguardar los datos de carácter personal de las unidades de observación y análisis, las Bases de Datos de la ENDI han atravesado por procedimientos de anonimización, mediante los cuales se omite información sensible y/o confidencial, permitiendo que la difusión de las estadísticas no vulnere los derechos a la protección de los datos personales garantizados en la Constitución, leyes y demás normativa vigente en el país. Dando cumplimiento a este mandato, tenga presente lo siguiente:

1. Los registros de personas constituyen una muestra del universo, lo cual reduce la probabilidad de identificar individuos específicos.
2. En las tablas se excluyen las variables de identificación directa de las personas (nombres y apellidos), así como otras variables que permitan su identificación indirecta (Ver Cuadro 6).
3. Para poder identificar a cada registro en la Base de Datos se cuenta con un identificador único a nivel de persona y de hogar; también se asigna un identificador único a cada vivienda y a cada UPM. Estos identificadores se forman mediante la concatenación de dígitos de las variables codificadas: *provincia(2dígitos) - cantón(2dígitos) - parroquia(2dígitos) - conglomerado(4dígitos) - vivienda(2dígitos)- hogar(2dígitos) - persona(2dígitos)*.

En este sentido en el Cuadro 6 se describen las variables que han sido anonimizadas para resguardar a la fuente directa de la información.

**Cuadro 6.** Variables anonimizadas

| Nº | Nombre de la tabla      | Formulario / Sección                           | Variables confidenciales  |
|----|-------------------------|--|---|
| 1  | BDD_ENDI_R2_f1_personas | f1 – Formulario 1 del hogar / <b>Carátula</b>  | Nombre de calle o camino.<br>Número domiciliario.<br>Número o letra de lote, bloque, torre u otro, piso, y número o letra de casa.<br>Pieza o departamento.<br>Nombre del informante calificado.<br>Número telefónico.<br>Personal responsable. |
|    |                         | f1 – Formulario 1 del hogar / <b>Sección 1</b> | Nombres y apellidos de las personas (pregunta 0).<br>Número de cédula de identidad (pregunta 6).  |

|   |                               |   |  |
|---|-------------------------------|---|--|
| 2 | BDD_ENDI_R2_f1_hogar          | f1 – Formulario 1 del hogar                                 | Ninguna.   |
| 3 | BDD_ENDI_R2_f2_mef            | f2 – Formulario 2 de la MEF / <b>Carátula</b>               | Ubicación de la vivienda, nombres y apellidos y otros datos de la informante (MEF), personal responsable.                    |
|   |                               | f2 – Formulario 2 de la MEF / <b>Sección 2</b>              | Nombres de los hijos nacidos vivos de la MEF.  |
| 4 | BDD_ENDI_R2_f2_lactancia      | f2 – Formulario 2 de la MEF / <b>Sección 3</b>              | Nombre del menor de 3 años de edad.  |
| 5 | BDD_ENDI_R2_f2_salud_ninez    | f2 – Formulario 2 de la MEF / <b>Sección 4</b>              | Nombre del menor de 5 años de edad.  |
| 6 | BDD_ENDI_R2_f3_desarrollo_inf | f3 – Formulario 3 de Desarrollo infantil / <b>Carátula</b>  | Ubicación de la vivienda y nombres y apellidos de la/el informante (cuidador) y del niño seleccionado, personal responsable. |
|   |                               | f3 – Formulario 3 de Desarrollo infantil / <b>Sección 0</b> | Nombres y apellidos de los menores de 5 años que hay en el hogar   |

Fuente: Instituto Nacional de Estadística y Censos.

El proceso de anonimización de la Base de Datos asegura que en los registros no sea posible identificar características de manera personal.

Para poder diferenciar un registro de otro y para habilitar la relación (*match*) entre tablas se crearon identificadores únicos codificados.

En el Cuadro 7 se detallan las variables que cumplen la función de identificadores únicos en las tablas, así como su estructura de codificación:

**Cuadro 7.** Ejemplo de creación de variables codificadas para identificación única

| Nombre codificado en Base de Datos | Ejemplo - Concatenación de códigos que la componen |        |           |              |          |       |   |         |                              |
|------------------------------------|--|--------|-----------|--------------|----------|-------|---|---------|------------------------------|
|                                    | Provincia  | Cantón | Parroquia | Conglomerado | Vivienda | Hogar | Persona madre (*en salud niñez y lact.) | Persona | Orden hijo (*en salud niñez) |
| id_upm                             | 04   | 02     | 50        | 0090         | x        | x     | x                                       | x       | x                            |
| id_viv                             | 04   | 02     | 50        | 0090         | 08       | x     | x                                       | x       | x                            |
| id_hogar                           | 04   | 02     | 50        | 0090         | 08       | 01    | x                                       | x       | x                            |
| id_per,<br>id_mef,<br>id_cuid      | 04   | 02     | 50        | 0090         | 08       | 01    | x                                       | 11      | x                            |
| id_mef_per                         | 04   | 02     | 50        | 0090         | 08       | 01    | 08                                      | 11      | x                            |
| id_hijo_ord                        | 04   | 02     | 50        | 0090         | 08       | 01    | 08                                      | 11      | 01                           |

Fuente: Instituto Nacional de Estadística y Censos.

## 6.2 Tratamiento de valores perdidos, imputaciones y validaciones

En el proceso de la generación de la información para la ENDI 2023-24, se observaron los valores registrados que se detallan en el Cuadro 8:

**Cuadro 8.** Tratamiento de valores perdidos en la Base de Datos

| Motivo de ausencia      | Descripción  | Tratamiento   |
|-------------------------|--|---|
| Ausentes por flujo      | Significa que la respuesta a una pregunta puede dirigir al informante a otras preguntas dependiendo de la opción seleccionada. Como resultado, es posible que las preguntas anteriores a este flujo no sean respondidas, quedando así vacías en casos específicos. | <p><b>NA</b> en formato <i>.rds</i></p> <ul style="list-style-type: none"> <li>• (carácter de punto) en formatos <i>.csv / .dta</i></li> </ul>  |
| Omisión de información  | Debido a que el informante omite proporcionar información a pesar de que la pregunta es aplicable, los valores se codifican de la siguiente forma.   | <ul style="list-style-type: none"> <li>- Año / mes / día de nacimiento de los miembros del hogar: <b>9999 / 99 / 99</b>.</li> <li>- Ingresos no reportados en Actividades Económicas: <b>999999</b></li> <li>- Día / mes / año en historial de embarazos y nacimientos: <b>99 / 99 / 9999</b></li> <li>- Meses / días que le dio de lactar solamente el pecho sin ningún otro líquido o complemento alimenticio: <b>99 / 99</b></li> <li>- Preguntas en las que No sabe/No responde: <b>88</b></li> </ul> |
| Limitaciones logísticas | Debido a temas como: condiciones climáticas, desastres naturales, áreas de difícil acceso, etc., los valores se codifican de la siguiente forma.   | <ul style="list-style-type: none"> <li>- Resultados de las pruebas de cloro y E. coli, tanto en fuente como en vaso: <b>999</b></li> </ul>  |

**Fuente:** Instituto Nacional de Estadística y Censos.

No se realizan imputaciones a los datos perdidos. Respecto a validaciones, se procede de la siguiente manera:

Se verifica en toda la Base de Datos que se encuentre la población a la cual se hace referencia en cada tabla. Es decir:

- Se verifica que las variables tengan valor no faltante
- Se verifica que las variables tengan el formato apropiado
- Se verifica que no existan casos duplicados
- Se analiza consistencia de la información, verificando que los valores sean coherentes entre variables
- Se observa que la información esté relacionada entre las diferentes variables
- Se verifica que en todas las tablas se encuentre la población a la cual se hace referencia. Es decir:
  - f1\_personas, debe contener todas las personas miembros de hogar.
  - f1\_hogar, se verifica que se encuentren todos los hogares.
  - f2\_mef, se verifica que contenga mujeres de 10 a 49 años.



- f2\_lactancia, se verifica que contenga niñas/os menores de 3 años.
- f2\_salud\_ninez, se verifica que contenga niñas/os menores de 5 años.
- f3\_desarrollo\_inf, se verifica que contenga niñas/os menores de 5 años.

## 7. Réplica e interpretación de indicadores principales

### 7.1 Fichas metodológicas para indicadores

Los principales indicadores de esta operación estadística han sido homologados, revisados y aprobados en Comisiones Especiales de Estadística, tomando en cuenta estándares internacionales y criterios de expertos. Estas consideraciones se han incluido en las fichas metodológicas de indicadores.

Las fichas metodológicas permiten generar y replicar indicadores con una misma metodología y sirven de insumo al usuario para su análisis.

**Figura 4.** Formato de ficha metodológica de indicadores

| FICHA METODOLÓGICA                         |               |
|--|---------------|
| NOMBRE DEL INDICADOR                       |               |
| DEFINICIÓN                                 |               |
| FÓRMULA DE CÁLCULO                         |               |
| DEFINICIÓN DE LAS VARIABLES RELACIONADAS   |               |
| METODOLOGÍA DE CÁLCULO                     |               |
| LIMITACIONES TÉCNICAS                      |               |
| UNIDAD DE MEDIDA O EXPRESIÓN DEL INDICADOR |               |
| INTERPRETACIÓN DEL INDICADOR               |               |
| FUENTE DE DATOS                            |               |
| PERIODICIDAD DEL INDICADOR                 |               |
| DISPONIBILIDAD DE LOS DATOS                |               |
| NIVEL DE DESAGREGACIÓN                     | GEOGRÁFICO    |
|  | GENERAL       |
|  | OTROS ÁMBITOS |
| INFORMACIÓN GEO - REFERENCIADA             |               |

|   |  |
|---|--|
| RELACION CON INSTRUMENTOS DE PLANIFICACIÓN NACIONAL E INTERNACIONAL |  |
| REFERENCIAS BIBLIOGRÁFICAS DE LA CONSTRUCCIÓN DEL INDICADOR         |  |
| FECHA DE ELABORACIÓN DE LA FICHA METODOLÓGICA                       |  |
| FECHA DE LA ÚLTIMA ACTUALIZACIÓN DE LA FICHA METODOLÓGICA           |  |
| CLASIFICADOR TEMÁTICO ESTADÍSTICO                                   |  |
| ELABORADO POR   |  |

| ANEXOS   |                    |
|--|--------------------|
| ANEXO 1: Algoritmo de cálculo del indicador SINTAXIS               |                    |
| 1. ENCUESTADO DE SINTAXIS-EJEMPLO                                  |                    |
| * TÍTULO DE LA SINTAXIS  |                    |
| * Cálculo de XXXX  |                    |
| * OPERACIÓN ESTADÍSTICA:   |                    |
| * Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMSE)      |                    |
| * ENTIDAD ENCUESTORA:  |                    |
| * Instituto Nacional de Estadística y Censos (INEC)                |                    |
| * UNIDAD TÉCNICA REFORMABLE:                                       |                    |
| * Dirección de Estudios de Población y Condiciones de Vida (DEPCV) |                    |
| * Fecha de elaboración:  | 15 de mayo de 2020 |
| * Fecha última modificación:                                       | 27 de mayo de 2020 |
| * Software estadístico:  | SPSS 23            |
| * Elaborado por:   |                    |
| * JANA 8888  |                    |
| * Analista de Estudios de Población y Condiciones de Vida          |                    |
| * Instituto Nacional de Estadística y Censos                       |                    |
| * JANA_8888@inec.gob.ec  |                    |
| * Revisado por:  |                    |
| * CCCC 0000  |                    |
| * Jefe de Estudios de Población y Condiciones de Vida              |                    |
| * Instituto Nacional de Estadística y Censos                       |                    |
| * CCCC_0000@inec.gob.ec  |                    |
| * Aprobado por:  |                    |
| * EEEE FFFF  |                    |
| * Director de Estudios de Población y Condiciones de Vida          |                    |
| * Instituto Nacional de Estadística y Censos                       |                    |
| * EEEE_FFFF@inec.gob.ec  |                    |
| 2. código de la sintaxis   |                    |
| AÑO DE REFERENCIA DE LA SINTAXIS                                   |                    |

**Fuente:** Instituto Nacional de Estadística y Censos.

Estas fichas contienen información detallada sobre el indicador y un espacio que ayuda con su interpretación, por lo tanto, se recomienda revisar el punto ilustrado en la Figura 5, que se encuentra dentro de la ficha del indicador:

**Figura 5.** Sección de interpretación del indicador en la ficha metodológica

| INTERPRETACIÓN DEL INDICADOR  |
|---|
| Por cada 100 niños/as menores de 5 años de edad se estiman XX niños con desnutrición crónica en el periodo determinado. |

**Fuente:** Instituto Nacional de Estadística y Censos.

**Nota:** Si planifica hacer una réplica exacta de los indicadores remítase a la *Ficha del indicador* y al archivo de *Sintaxis* de la ENDI 2023-24.

## 7.2 Factores de expansión para encuestas

Los factores de expansión para encuestas se deben tomar en cuenta para el desarrollo de productos analíticos y réplica de los resultados presentados en la operación estadística.

Debido a la naturaleza y flujos de los que dependen las preguntas presentes en una encuesta, ésta puede presentar entrevistas parciales entre los informantes efectivos. Esto dificulta comparar las estimaciones de totales de la misma población objetivo obtenidas entre las secciones/variables, puesto que estas varían dependiendo de la magnitud de la no respuesta parcial. Esta es una particularidad presente en encuestas como la ENDI, debido en gran medida a sus poblaciones objetivo.

Existen diferentes métodos para solventar este inconveniente. Uno de ellos es la imputación de los valores perdidos de las variables necesarias para los cálculos de los indicadores. Otro método consiste en calcular un factor de expansión normalizado para cada uno de los indicadores considerando la cobertura de las preguntas que construyen el indicador. En el *ICF International* se recomienda utilizar **factores normalizados** (ICF International, 2012).

### **Factor de expansión normalizado:**

Al trabajar con factores normalizados, el número de casos no ponderados coincide con el número de casos ponderados y se calcula multiplicando el factor de expansión de cada caso por un factor de normalización. El factor de expansión normalizado puede ser usado para estimar medias, proporciones, tasas y razones.

**Recuerde:** para un correcto análisis de la información o réplica de indicadores, se deben aplicar los factores de expansión normalizados apropiados para cada una de las tablas de la ENDI.

**Nota:** Si desea conocer más sobre la muestra y los factores de expansión, remítase al documento *Metodología del Diseño Muestral* de la ENDI 2023-24.

## 7.3 Coeficientes de variación y error asociado al estimador (para encuestas)

Se aconseja que una vez realizado el análisis de la información se revise los coeficientes de variación e intervalos de confianza para robustecer la interpretación de los resultados.

**Coeficiente de variación:** Medida relativa de la dispersión, que puede compararse para diferentes distribuciones y que expresa la desviación estándar como porcentaje de la media (Levin & Rubin, 2004, pág. 118).

**Intervalo de confianza:** Un rango de valores que tiene designada una probabilidad de que incluya el valor verdadero del parámetro de la población. (Levin & Rubin, 2004, pág. 312).

El intervalo de confianza y el coeficiente de variación complementa y facilita la lectura e interpretación de los indicadores.



## Referencias bibliográficas

ICF International. (2012). *Demographic and Health Survey Sampling and Household Listing Manual*. Calverton, Maryland.

Levin, R., & Rubin, D. (2004). *Estadística para Administración y Economía* (7ma edición ed.). Retrieved from <https://profefily.com/wp-content/uploads/2017/12/Estad%C3%ADstica-para-administraci%C3%B3n-y-economia-Richard-I.-Levin.pdf>

Secretaría Técnica Ecuador Crece Sin Desnutrición Infantil. (n.d.). [www.infancia.gob.ec](http://www.infancia.gob.ec). Retrieved septiembre 02, 2024, from <https://www.infancia.gob.ec/wp-content/uploads/downloads/2023/11/BasePoliticaNov2023V16FINAL.pdf>

**ENCUESTA  
NACIONAL SOBRE  
DESNUTRICIÓN  
*Infantil***

**INEC**

Buenas cifras,  
mejores vidas

