



Observatorio
Social



Uso de la Base de Datos

Encuesta Longitudinal de Primera Infancia Cuarta Ronda 2024

Noviembre 2025



Uso de la Base de Datos de la Encuesta Longitudinal de Primera Infancia 2024 – Cuarta Ronda

ELPI 2024

1. Introducción

El Ministerio de Desarrollo Social y Familia pone a disposición de investigadores, académicos, diseñadores de políticas públicas y público general las bases de datos de la cuarta ronda de la Encuesta Longitudinal de Primera Infancia (ELPI 2024).

Este documento tiene como objetivo entregar antecedentes y lineamientos técnicos para el correcto uso y análisis de la información contenida en las bases de datos. Se presenta una descripción general de la encuesta, la estructura de los archivos de datos que se publican, y las consideraciones metodológicas fundamentales que se deben tener en cuenta para el análisis, en particular las relacionadas con el diseño muestral complejo de la encuesta.

Para esta ronda, se publican tres bases de datos principales:

1. **Base de datos del Responsable Principal:** Contiene la información del cuestionario aplicado al responsable principal, incluyendo las características de todos los integrantes del hogar.
2. **Base de datos del Adolescente:** Incluye las respuestas del cuestionario autoaplicado por el o la adolescente.
3. **Base de datos de Evaluaciones:** Reúne los resultados de las escalas y tests aplicados, tanto al responsable principal como al adolescente.

Junto con las bases, se publica la documentación asociada, como los cuestionarios, el libro de códigos y este documento.

2. Antecedentes de la Encuesta ELPI 2024

La Encuesta Longitudinal de Primera Infancia (ELPI) es un estudio que sigue a una misma cohorte de individuos a lo largo del tiempo. Su objetivo principal ha sido caracterizar y analizar el desarrollo de niños y niñas a lo largo de su ciclo vital, considerando el efecto que pueden tener las características del hogar, de la principal persona responsable de su cuidado y del entorno sobre sus trayectorias de vida, con el fin de informar la política pública orientada a la infancia y adolescencia.

La cuarta ronda de ELPI, levantada en 2024, se centró en el seguimiento de la cohorte original de la encuesta, seleccionada en 2010. Por lo tanto, la población objetivo de esta medición son los adolescentes y jóvenes de entre 14 y 18 años que participaron en la primera ronda, junto a sus responsables principales.

El levantamiento de la encuesta fue adjudicado por el Ministerio de Desarrollo Social y Familia, a través de licitación pública, al Centro de Microdatos (CMD) de la Facultad de Economía y Negocios de la Universidad de Chile. El diseño del estudio contó, además, con la asesoría de un Panel de

Expertos y la colaboración del Instituto Nacional de Estadísticas (INE) para el desarrollo de los factores de expansión.

La recolección de datos se realizó de manera presencial en todo el país entre el 25 de mayo y el 15 de octubre de 2024, utilizando dispositivos móviles (modalidad CAPI). La muestra lograda fue de 10.003 niños, niñas y adolescentes. La base de datos del responsable principal, que incluye a todos los miembros de los hogares de estos adolescentes, contiene un total de 41.913 observaciones.

Es importante señalar que, debido a que la muestra incluye a adolescentes de hasta 18 años, se identificaron casos de jóvenes que vivían de manera independiente al momento de la entrevista. Para estos casos, se adaptó el flujo de la encuesta permitiéndoles responder también por las características de su hogar, por lo que no cuentan con un responsable principal informante. El diseño muestral de ELPI permite obtener resultados representativos a nivel nacional y por tramos de edad.

3. Instrumentos de Recolección

La cuarta ronda de la Encuesta Longitudinal de Primera Infancia (ELPI 2024) utilizó un conjunto de cuatro instrumentos de recolección de datos, aplicados a distintos informantes dentro del hogar del o la adolescente seleccionado(a). Estos instrumentos fueron administrados en entrevistas presenciales asistidas por dispositivos móviles (CAPI).

A continuación, se describen los cuatro instrumentos:

1. Cuestionario al Responsable Principal: Aplicado por un(a) encuestador(a) a la persona adulta identificada como la principal responsable del cuidado del adolescente.
2. Cuestionario al Adolescente: Cuestionario autoaplicado, respondido directamente por el o la adolescente en una tablet.
3. Evaluaciones al Responsable Principal: Conjunto de escalas autoaplicadas, respondidas por el responsable principal en una segunda tablet, para asegurar mayor privacidad.
4. Evaluación al Adolescente (TVIP): Test aplicado directamente al adolescente por un(a) encuestador(a) capacitado(a).

La **Tabla 1** presenta las especificaciones de cada uno de estos instrumentos, detallando la información que recogen y quién es la persona informante.

Tabla 1. Especificaciones de los instrumentos de recolección ELPI 2024

Instrumentos	¿Qué información recoge?	¿Quién responde?
Cuestionario Responsable Principal	al Recoge información sobre el hogar, el adolescente y el propio responsable principal, a través de los siguientes módulos: • Composición del hogar • Padre biológico no integrante • Madre conocida no integrante • Situación educacional • Educación • Salud del adolescente • Prácticas y condiciones de crianza • Roles familiares y expectativas • Ambiente familiar (Escala HOME-21) • Acontecimientos y eventos adversos • Salud del responsable principal • Situación laboral • Ingresos del hogar • Acceso a programas y subsidios • Vivienda y barrio	Persona adulta responsable principal del cuidado del o la adolescente.
Cuestionario Adolescente	al Indaga sobre las percepciones, hábitos y experiencias del adolescente en distintas áreas de su vida, a través de los siguientes módulos: • Actividades y participación • Educación y expectativas • Relaciones familiares • Salud y bienestar • Relaciones sociales y afectividad • Entorno y seguridad • Género y orientación sexual	El o la adolescente seleccionado(a).
Evaluaciones Responsable Principal	al Mide aspectos específicos del desarrollo socioemocional del adolescente y del bienestar del responsable principal, mediante las siguientes escalas estandarizadas: • Child Behavior Checklist (CBCL) • Center for Epidemiologic Studies Depression Scale (CESD-10) • Parent Stress Index (PSI) • Parenting Sense of Competence Scale (PSCS) • Adverse Childhood Experiences (ACEs)	Persona responsable principal del cuidado del o la adolescente.

Evaluación al Adolescente (TVIP)	Mide el vocabulario receptivo del adolescente como un indicador de su desarrollo cognitivo y verbal, a través de la aplicación del Test de Vocabulario en Imágenes Peabody (TVIP)	El o la adolescente seleccionado(a).
---	---	--------------------------------------

Fuente: Elaboración propia en base a los instrumentos de ELPI 2024.

4. Estructura de las Bases de Datos ELPI 2024

4.1. Descripción y Acceso a las Bases de Datos

Las bases de datos de la cuarta ronda de la Encuesta Longitudinal de Primera Infancia (ELPI 2024) se encuentran disponibles para uso público a través de archivos descargables en diversos formatos estadísticos (Stata, SPSS y R). Junto con los datos, se publica toda la documentación de respaldo, incluyendo los cuestionarios, el libro de códigos y el presente informe metodológico. Toda la información puede ser descargada desde el sitio web del Observatorio Social del Ministerio de Desarrollo Social y Familia.

Con el objetivo de resguardar la confidencialidad de la información y la privacidad de los participantes, las bases de datos han sido sometidas a rigurosos procesos de anonimización. Esto incluye la eliminación de identificadores directos (como nombres o direcciones) y la recodificación de variables geográficas sensibles.

4.2. Estructura y Vinculación de las Bases

ELPI 2024 distribuye sus datos en tres archivos separados, cada uno con una estructura y unidad de análisis distinta. Esta organización responde a la complejidad de los instrumentos aplicados y a la diversidad de informantes.

La vinculación entre estas tres bases de datos, así como con las bases de rondas anteriores (2010, 2012 y 2017), se realiza a través de la variable llave *folio*. Este identificador único representa al adolescente seleccionado(a) en la muestra original de 2010 y se mantiene constante a lo largo del tiempo.

En la **Tabla 2** se detalla la estructura de cada una de las bases de datos:

Tabla 2. Descripción de las bases de datos ELPI 2024

Base de Datos	Unidad de Observación	N° de Casos	Descripción del Contenido
Responsable Principal (RP)	Cada persona integrante del hogar.	41.913	Contiene la información del "Cuestionario al Responsable Principal". Incluye variables de caracterización de todos los miembros del hogar (demográficas, educacionales, laborales, de salud) y módulos específicos respondidos por el responsable principal sobre el adolescente y el hogar en general. Dentro de esta base, la variable tipo_persona permite identificar al adolescente seleccionado/a (1), al responsable principal (2), al responsable secundario si existe (3) y a otros integrantes del hogar (4). Al realizar análisis, es fundamental filtrar la base según la unidad de observación de interés.
Adolescente	Cada adolescente seleccionado(a) y entrevistado(a).	10.003	Contiene las respuestas del "Cuestionario al Adolescente". Incluye variables auto-reportadas sobre las actividades, expectativas, relaciones sociales y familiares, salud, bienestar y percepciones del o la adolescente.
Evaluaciones	Cada adolescente seleccionado(a) y entrevistado(a).	10.003	Contiene los resultados de las "Evaluaciones al Responsable Principal" y la "Evaluación al Adolescente". Incluye los puntajes y respuestas de las escalas estandarizadas (CBCL, CESD-10, PSI, PSCS, ACEs) y el test de vocabulario (TVIP).

Fuente: Elaboración propia en base a las bases de datos ELPI 2024.

4.3. Variables Construidas y Editadas

Las bases de datos publicadas no solo contienen las respuestas directas a los cuestionarios, sino que también incluyen un conjunto de variables procesadas por el equipo del Ministerio de Desarrollo Social y Familia para facilitar su análisis. Entre ellas destacan:

Variables de Ingreso Corregido: Se incluyen las variables de ingreso del trabajo y de pensiones con los valores imputados para los casos de no respuesta, permitiendo análisis económicos más robustos.

Escalas Estandarizadas: Para las distintas evaluaciones aplicadas (CBCL, CESD-10, etc.), se entregan los puntajes calculados según las normas de cada instrumento. Para más detalles respecto de estas consultar el

Variables de Diseño Muestral: Se incluyen las variables necesarias para el análisis de encuestas complejas, como el factor de expansión, los estratos y los conglomerados (ver sección 5).

Variables Editadas: Corresponden a variables que fueron corregidas para resolver inconsistencias detectadas durante el proceso de validación de datos.

5. Consideraciones para el Análisis de Datos

5.1. Diseño Muestral Complejo

La obtención de estimaciones a partir de encuestas con un diseño muestral complejo, como ELPI, requiere la especificación adecuada de las características de dicho diseño al momento del análisis. La consecuencia práctica de no considerar el diseño muestral es la obtención de varianzas incorrectas (generalmente subestimadas), lo que puede llevar a conclusiones erróneas sobre la significancia estadística de los resultados.

Para estimar varianzas correctamente con los datos de ELPI 2024, es necesario identificar y utilizar las variables que dan cuenta de su estructura de muestreo.

Variables del diseño muestral:

- **Factor de Expansión (Ponderador):** Variable f_exp . Este es el factor de expansión transversal para la ronda 2024. Representa el número de adolescentes de la población objetivo que cada caso entrevistado representa. Su uso es indispensable para obtener estimaciones representativas de la cohorte para el año 2024 a nivel poblacional (nacional, regional o por tramos de edad).
- **Estrato:** Variable estrato. La estratificación en ELPI corresponde a las comunas seleccionadas en la primera etapa del muestreo. Esta variable identifica los 116 estratos geográficos de los cuales se seleccionó la muestra y es crucial para el cálculo correcto de las varianzas.
- **Unidad Primaria de Muestreo (UPM) / Conglomerado:** Dado que la selección de adolescentes se realizó directamente dentro de cada estrato (comuna), no existe una variable de conglomerado de nivel superior. En los programas estadísticos, esto se especifica indicando que no hay un clúster (~1) para las bases a nivel de adolescente. Para la base de datos a nivel de persona del hogar, se debe usar *folio* como identificador de conglomerado para asegurar que los miembros de un mismo hogar no sean tratados como observaciones independientes.

5.2. Sintaxis para Programas Estadísticos

A continuación, se presenta la sintaxis básica para especificar el diseño complejo de ELPI 2024 en los programas estadísticos STATA, R y SPSS. Los ejemplos muestran cómo calcular la edad promedio de los adolescentes.

a. STATA

```
* Para bases a nivel de adolescente (Adolescente, Evaluaciones)
svyset [pw=f_exp], strata(estrato)

* Ejemplo de cálculo
svy: mean edad

* Para la base del Responsable Principal (a nivel de persona del hogar)
svyset folio [pw=f_exp], strata(estrato)

* Ejemplo de cálculo
svy: mean edad if tipo_persona == 1 // Filtrando por el adolescente
```

b. R (con el paquete survey)

```
# Cargar el paquete
library(survey)

# Especificar el diseño para bases a nivel de adolescente (ej. data_ado)
# Usamos ids = ~1 para indicar que no hay conglomerados dentro de los estratos
diseno_elpi_ado <- svydesign(ids = ~1, strata = ~estrato, weights = ~f_exp, data = data_ado)

# Ejemplo de cálculo
svymean(~edad, design = diseno_elpi_ado, na.rm = TRUE)

# Especificar el diseño para la base del Responsable Principal (ej. data_rp)
# Usamos ids = ~folio para agrupar por hogar
diseno_elpi_rp <- svydesign(ids = ~folio, strata = ~estrato, weights = ~f_exp, data = data_rp)

# Ejemplo de cálculo
svymean(~edad, design = subset(diseno_elpi_rp, tipo_persona == 1), na.rm = TRUE)
```

c. SPSS

* Para bases a nivel de adolescente (Adolescente, Evaluaciones)

CSPLAN ANALYSIS

/PLAN FILE='ruta_archivo\Plan_analisis_elpi2024_ado.csaplan'

/PLANVARS ANALYSISWEIGHT=f_exp

/SRSESTIMATOR TYPE=WOR

/PRINT PLAN

/DESIGN STRATA=estrato

/ESTIMATOR TYPE=WR.

* Ejemplo de cálculo

CSDESCRIPTIVES

/PLAN FILE='ruta_archivo\Plan_analisis_elpi2024_ado.csaplan'

/VARIABLES = edad

/STATISTICS = MEAN SE COUNT

/OPTIONS = TOTAL.

* Para la base del Responsable Principal (a nivel de persona del hogar)

CSPLAN ANALYSIS

/PLAN FILE='ruta_archivo\Plan_analisis_elpi2024_rp.csaplan'

/PLANVARS ANALYSISWEIGHT=f_exp

/SRSESTIMATOR TYPE=WOR

/PRINT PLAN

/DESIGN STRATA=estrato CLUSTER=folio

/ESTIMATOR TYPE=WR.

* Ejemplo de cálculo

CSDESCRIPTIVES

/PLAN FILE='ruta_archivo\Plan_analisis_elpi2024_rp.csaplan'

/VARIABLES = edad

/STATISTICS = MEAN SE COUNT

/OPTIONS = TOTAL.

5.3. Unidades de Análisis

Debido a la existencia de múltiples bases de datos con diferentes estructuras, es crucial que la persona usuaria identifique correctamente la unidad de análisis para cada pregunta de investigación:

- Análisis a nivel de adolescente: Para la mayoría de los análisis, la unidad de observación será el adolescente (10.003 casos). En este escenario, las variables contenidas en la base del Responsable Principal (ej. ingreso del hogar, nivel educativo del cuidador) y en la base de Evaluaciones (ej. puntaje CBCL) deben ser tratadas como atributos o características del adolescente. Para ello, se deben fusionar las bases de datos utilizando la variable llave *folio*. Todas las estimaciones expandidas con el factor f_{exp} serán representativas de la población de adolescentes de la cohorte.
- Análisis de otros miembros del hogar: La base de datos del Responsable Principal contiene información de todos los integrantes de los hogares. Si bien es posible realizar análisis descriptivos sobre este grupo (ej. caracterizar a los hermanos o a los responsables principales), es importante destacar que las estimaciones no son representativas de la población total de hermanos o de responsables principales de Chile. Los resultados describen únicamente las características de las personas que viven en los hogares de los adolescentes de la cohorte estudiada.