



Agosto 2025

Encuesta Nacional de Participación Cultural y Comportamiento Lector

ENPCCL 2024

MANUAL DE USUARIO DE LA BASE DE DATOS

Instituto Nacional de Estadísticas
Subdepartamento de Estadísticas de Condiciones de Vida



Índice

CAPÍTULO I: INTRODUCCIÓN	4
CAPÍTULO II: ENCUESTA NACIONAL DE PARTICIPACIÓN CULTURAL Y COMPORTAMIENTO LECTOR	5
2.1. OBJETIVOS	6
2.2. ACCESO Y SOLICITUD DE LOS DATOS	6
2.3. FICHA TÉCNICA.....	6
CAPÍTULO III: ESTRUCTURA DEL FORMULARIO	8
CAPÍTULO IV: INDICACIONES PARA EL USO DE LA BASE DE DATOS	11
4.1. DISEÑO MUESTRAL	11
4.2. UNIDAD DE ANÁLISIS	14
4.3. USO DE FACTORES DE EXPANSIÓN	15
4.4. TEST DE HIPÓTESIS	16
4.5. MANEJO DE DATOS CON CATEGORÍAS NO APLICA	19
CAPÍTULO V: INDICADORES AGREGADOS.....	21
CAPÍTULO VI: CALIDAD DE LAS ESTIMACIONES	31

Índice de Tablas

Tabla 1. Ficha Técnica ENPCCL 2024.....	7
Tabla 2. Estructura del formulario ENPCCL 2024.....	10
Tabla 3. Unidades de análisis en ENPCCL 2024.....	14
Tabla 4. Factores de expansión asociados a las diferentes unidades de análisis.....	15
Tabla 6. Indicadores agregados	21

Índice de Figuras

Figura 1. Flujo de Evaluación de Calidad de las Estimaciones	32
--	----

INSTITUTO NACIONAL DE ESTADÍSTICAS

Subdirección Técnica

Departamento de Estadísticas Demográficas Y Sociales

Subdepartamento de Estadísticas de Condiciones de Vida

ENCUESTA NACIONAL DE PARTICIPACIÓN CULTURAL Y COMPORTAMIENTO LECTOR 2024

Manual de usuario de la base de datos

Agosto 2025

Dirección: Morandé 801, piso 22, Santiago de Chile

Teléfono: (56) 22892 4000

Sitio web: www.ine.gob.cl

Correo electrónico: : ine@ine.gob.cl

X: @INE_Chile

Instagram: Chile.INE

LinkedIn: Instituto Nacional de Estadísticas de Chile

Facebook: @ChileINE

CAPÍTULO I: INTRODUCCIÓN

El Ministerio de las Culturas, las Artes y el Patrimonio (MINCAP) en conjunto con el Instituto Nacional de Estadísticas (INE), realizaron la Encuesta Nacional de Participación Cultural y Comportamiento Lector (ENPCCL) durante el año 2024. La encuesta permite contar con información útil tanto para la toma de decisiones como para generar conocimiento sistemático sobre la participación cultural y el comportamiento lector que presenta el país.

El presente documento corresponde al “Manual de Usuario de la Base de Datos” de la ENPCCL, cuyo objetivo es servir de apoyo para la correcta utilización de los microdatos, así como las programaciones que se requieran para hacer uso de los datos y replicar los indicadores provistos.

El manual se estructura en seis capítulos. El primero corresponde a esta introducción; el segundo detalla los objetivos estratégicos y aspectos técnicos de la ENPCCL 2024, mientras que el tercero expone la arquitectura del cuestionario. La cuarta sección proporciona directrices para el manejo de las bases de datos, y la quinta desarrolla la metodología de construcción de indicadores agregados mediante el software estadístico R, incluyendo la sintaxis respectiva. En el sexto capítulo se especifica el protocolo de evaluación de la calidad estadística de los indicadores, con énfasis en el uso del paquete de control de calidad para encuestas de hogares.

Adicionalmente, se pone a disposición el detalle de la base de datos en un documento anexo publicado bajo el título Diccionario de variables ENPCCL 2024, que complementa el presente manual.

CAPÍTULO II: ENCUESTA NACIONAL DE PARTICIPACIÓN CULTURAL Y COMPORTAMIENTO LECTOR

La ENPCCL surge a partir de un convenio de colaboración entre la SCA y el INE, marcando la primera edición en la que se integran dos estudios estadísticos previos: la Encuesta Nacional de Participación Cultural y la Encuesta de Comportamiento Lector.

Dado el carácter innovador de este proceso estadístico, se incorporaron mejoras basadas en experiencias tanto nacionales como internacionales, así como en pruebas de los cuestionarios¹. Esta versión incluye actualizaciones relevantes en el diseño y en el sistema de indicadores, alineándose con los cambios recientes en el acceso a la cultura.

¹ Para mayor detalle consultar el Informe Metodológico ENPCCL 2024.

2.1. OBJETIVOS

Objetivo general

Caracterizar las prácticas de participación cultural de la población de 15 años y más, a partir de una encuesta con muestra representativa de la población urbana nacional y regional.

Objetivos específicos

- Medir las diferentes prácticas de participación cultural según dominios de estimación (país y región).
- Caracterizar las prácticas de participación cultural a nivel de práctica y características de las personas.
- Contar con información actualizada y específica respecto del comportamiento y hábitos de lectura de la población residente en zonas urbanas en Chile.

2.2. ACCESO Y SOLICITUD DE LOS DATOS

Las bases de datos de la ENPCCL 2024 se encuentran publicadas en el portal web del Instituto Nacional de Estadísticas (INE), específicamente dentro de la sección de estadísticas sociales de “Condiciones de vida y cultura”. Para acceder a ellas, se debe ingresar a la página institucional, y seguir la siguiente ruta: Estadísticas > Condiciones de vida y cultura > Participación Cultural y Comportamiento Lector > BASES DE DATOS > 2024.

Las bases de datos disponibles de la ENPCCL 2024 se encuentran en formato “.Rdata”, “.dta”, “.csv” (CommaSeparated Values) y “.sav” (SPSS).

2.3. FICHA TÉCNICA

La información presente en la base de datos corresponde a las respuestas dadas por los informantes Kish encuestados², que corresponden a 12.263 casos muestrales, y a la información sociodemográfica de los integrantes de sus hogares.

² El informante Kish corresponde al integrante del hogar que proporciona la información del cuestionario central, seleccionado aleatoriamente entre todas las personas residentes habituales de la vivienda.

Tabla 1. Ficha Técnica ENPCCL 2024

ENCUESTA NACIONAL DE PARTICIPACIÓN CULTURAL Y COMPORTAMIENTO LECTOR 2024	
Nombre de la base de datos	Base de datos – ENPCCL 2024
Nombre del archivo (Rdata)	Base de datos – ENPCCL 2024.Rdata
Nombre del archivo (dta)	Base de datos – ENPCCL 2024.dta
Nombre del archivo (SPSS)	Base de datos – ENPCCL 2024.sav
Nombre del archivo (CSV)	Base de datos – ENPCCL 2024.csv
Formato de archivo	.Rdata (R), dta, .sav (SPSS) y .csv (texto plano delimitado por comas)
N° de registros	34.787
N° de informantes Kish	12.263
N° de variables	553

CAPÍTULO III: ESTRUCTURA DEL FORMULARIO

El formulario de la ENPCCL 2024 se compone de tres secciones; 1) Hoja de ruta, 2) Registro de personas en el hogar y 3) Cuestionario central (Tabla 2). A continuación, se presenta una breve descripción del formulario ENPCCL 2024³. Para mayor detalle consultar el Informe Metodológico ENPCCL 2024.

HOJA DE RUTA

La hoja de ruta tiene como objetivo clasificar las unidades muestrales seleccionadas para la ENPCCL 2024. La clasificación establece la elegibilidad de las unidades, es decir, si corresponden a la población objetivo de la encuesta o no.

Se basa en la adaptación institucional de los códigos de disposición recomendados por la American Association for Public Opinion Research (AAPOR) para la aplicación de encuestas a hogares y personas

REGISTRO DE PERSONAS EN EL HOGAR (RPH)

Recolecta información sociodemográfica sobre las personas residentes habituales para poder analizar los resultados desagregados y así poder contextualizar el fenómeno en función de las características de la población. Esta sección contiene la caracterización demográfica, que incluye sexo, edad, género, orientación sexual, discapacidad, migración, pertenencia a pueblos originarios, nivel educativo, ocupación y nivel socioeconómico.

Para ENPCCL 2024 se incluye una sección sobre comportamiento lector en el hogar, con preguntas sobre la cantidad de libros disponibles en este y actividades lectoras realizadas con menores de 15 años en el hogar de la persona informante Kish, en caso de existan integrantes de dicho rango etario.

³ Para mayor detalle consultar el Informe Metodológico ENPCCL 2024.

CUESTIONARIO CENTRAL

Dentro del Cuestionario Central, la participación cultural se compone de siete módulos. Sin embargo, los indicadores principales se encuentran en los módulos de actividades artísticas y patrimoniales, infraestructuras culturales y comportamiento lector.

El Cuestionario Central inicia con el módulo de Cultura general. El siguiente módulo corresponde al de Actividades artísticas y patrimoniales, que utiliza una medición asincrónica, es decir, separa el módulo en una primera parte correspondiente a la identificación de 11 Actividades artísticas y patrimoniales, y su asistencia en la vida y en los últimos doce meses; seguido de una segunda parte de caracterización de la última asistencia a ocho Actividades artísticas y patrimoniales⁴. El módulo cierra con la sección de Motivos de no asistencia a actividades artísticas y patrimoniales.

El módulo de Infraestructuras culturales también utiliza la medición asincrónica, primero identificando seis Infraestructuras culturales, y su asistencia en la vida y en los últimos doce meses; una segunda parte de caracterización de la última asistencia a tres Infraestructuras culturales⁵. El módulo sigue con la sección de Motivos de no asistencia a infraestructuras culturales y finaliza con la sección de Acceso a infraestructuras culturales.

El módulo de Lectura indaga sobre el comportamiento lector de la persona informante, para luego caracterizar la lectura de libros en aquellos casos que hayan declarado leer libros por al menos 15 minutos en los últimos 12 meses y que leyeran por gusto y/u ocio.

El módulo de Entorno digital indaga sobre la participación cultural digital en internet y la realización de actividades culturales digitales simultaneas. Se finaliza con los módulos de Participación inventiva y Otros indicadores.

⁴ Para mayor detalle consultar el Informe Metodológico ENPCCL 2024.

⁵ Para mayor detalle consultar el Informe Metodológico ENPCCL 2024.

Tabla 2. Estructura del formulario ENPCCL 2024

SECCIÓN	MÓDULOS	SUBSECCIÓN
HOJA DE RUTA	HOJA DE RUTA	Identificación de la vivienda
		Identificación de visitas de la vivienda
REGISTRO DE PERSONAS EN EL HOGAR	Datos Sociodemográficos	Identificación de Hogares
		Registro de Personas en el Hogar
		Comportamiento y hábitos de lectura en el hogar
CUESTIONARIO CENTRAL	Cultura general	Sección única
	Actividades artísticas y patrimoniales	Asistencia a actividades artísticas y patrimoniales (<i>screening</i>)
		Caracterización asistencia actividades artísticas y patrimoniales
		Motivos de no asistencia a actividades artísticas y patrimoniales
	Infraestructuras culturales	Asistencia a infraestructuras culturales (<i>screening</i>)
		Caracterización asistencia a infraestructuras culturales
		Motivos de no asistencia a infraestructuras culturales
		Acceso a infraestructuras culturales
	Comportamiento lector	Lectura
		Libros
	Entorno digital	Sección única
	Participación inventiva	Sección única
	Otros indicadores	Sección única
REGISTRO DE PERSONAS EN EL HOGAR	Orientación sexual	Sección única

Fuente: Informe Metodológico ENPCCL 2024.

CAPÍTULO IV: INDICACIONES PARA EL USO DE LA BASE DE DATOS

En este apartado se incluyen aspectos necesarios para estimar los resultados de la ENPCCL 2024, a saber, las unidades de análisis correspondientes a cada indicador, los factores de expansión asociados a esa unidad de análisis y el plan de muestras complejas recomendado para cada indicador.

4.1. DISEÑO MUESTRAL

Dado que la ENPCCL 2024 utiliza un diseño muestral complejo, que considera tres etapas de selección, las estimaciones deben incluir medidas de precisión a través del uso de muestras complejas⁶ donde se definen las características asociadas al diseño muestral. Para el caso de la ENPCCL, las variables que definen el diseño muestral complejo son: estratos, conglomerados y factor de expansión.

Dada la configuración del diseño muestral y los resultados del levantamiento, se hace necesaria la construcción de nuevas variables que lo caractericen, de manera que se garanticen estimaciones insesgadas y consistentes de los errores muestrales, para ello se realizan agrupaciones de los estratos y conglomerados:

- **Pseudoestratos (Estrato_ENPC)**: Corresponde a la identificación de los pseudo estratos asociados al diseño muestral.
- **Pseudoconglomerados (Varunit)**: Corresponde a la identificación de los pseudo conglomerados asociados al diseño muestral.

Por último, es necesario incluir el factor de expansión asociado a la unidad de análisis a estimar:

- **Fexp_p**: Corresponde a la cantidad de personas en la población que representa cada persona entrevistada en base a sus características a nivel regional.

⁶ Para trabajar con diseños complejos se suelen emplear procedimientos relacionados con ajuste de modelos (tales como Surveyreg en SAS, CSGLM en SPSS, LINCOM en Stata, LM en R, entre otros).

- **Fexp_h**: Corresponde a la cantidad de hogares en la población que representa cada hogar entrevistado en base a sus características a nivel regional.

Para generar el plan de análisis de muestra complejas se deben emplear los siguientes elementos:

- Estrato: **Estrato_ENPC**
- Conglomerado: **Varunit**
- Factor de expansión:
 - **Fexp_p**
 - **Fexp_h**

A continuación, se incluye, a modo de ejemplo, la sintaxis utilizada en el software R para crear el plan de muestra compleja que permite estimar el porcentaje de personas que asistieron a actividades culturales y/o patrimoniales y/o infraestructuras culturales en los últimos 12 meses, el error estándar, el intervalo de confianza y el coeficiente de variación.

En R Studio, previamente, para generar el diseño muestral complejo (DC), es necesario cargar las librerías *tidyverse* y *survey*. Luego, se debe filtrar la base de datos ENPCCL por `kish_dummy = 1` y, utilizando el factor de expansión según el nivel de estimación requerido, construir el diseño con el siguiente código:

```
enpccl2024 <- haven::read_sav("Base de datos - ENPCCL 2024.sav")
enpccl2024_Kish <- enpccl2024 %>% filter(kish_dummy==1)
DC <- svydesign(ids = ~Varunit ,
               strata = ~Estrato_ENPC,
               weights = ~fexp_p,
               check.strata = TRUE,
               data = enpccl2024_Kish)
options(survey.lonely.psu = "certainty")
```

A modo de ejemplo, se estima el porcentaje de participación cultural a nivel nacional, el error estándar, el intervalo de confianza al 95% y el coeficiente de variación:

```

prop_filtro_pc <- data.frame("estimacion"=numeric(1),"se"=numeric(1))

prop_filtro_pc %>%
  mutate(estimacion = coef(svymean(~filtro_pc, design = DC)),
         se = SE(svymean(~filtro_pc, design = DC))) %>%
  mutate(enpccl2024 %>% summarise(gl = n_distinct(Varunit)-n_distinct(Est
rato_ENPC))) %>%
  mutate(t = qt(p = 0.95 + 0.05/2, df = gl),
         li = estimacion - se*t,
         ls = estimacion + se*t,
         cv = se / estimacion * 100)

```

4.2. UNIDAD DE ANÁLISIS

La ENPCCL 2024 tiene dos unidades de análisis (Tabla 3). Es importante hacer referencia exacta a las unidades que corresponden a cada pregunta para poder leer correctamente los resultados de la encuesta.

Tabla 3. Unidades de análisis en ENPCCL 2024

NIVEL	UNIDAD DE ANÁLISIS	DEFINICIÓN
PERSONA	Informante Kish	Personas de 15 años o más que habitan en las zonas urbanas de las 129 comunas que conforman la muestra de ENPCCL 2024
HOGAR	Hogar	Hogares residentes en viviendas particulares ocupadas de zonas urbanas de las 129 comunas que conforman la muestra de la ENPCCL 2024

Fuente: Informe Metodológico ENPCCL 2024.

Relación entre unidades de análisis y observación

La participación cultural, al igual que otros módulos de la encuesta, se mide a partir de la selección aleatoria de una persona por vivienda, por lo que la unidad de análisis es la persona y la unidad de observación corresponde al informante Kish. Para el caso del comportamiento y hábitos de lectura en el hogar (subsección de datos demográficos), esta corresponde a información a nivel de hogar del informante Kish. En el primer caso las estimaciones se construyen exclusivamente a partir de las declaraciones de la persona seleccionada aleatoriamente dentro del hogar, respecto a su participación en actividades culturales y/o patrimoniales y/o en infraestructuras culturales. En el segundo caso las estimaciones se construyen en base a lo declarado por la persona informante idónea y/o la persona informante Kish, en caso de que esta fuese la persona informante idónea de su hogar.

De este modo, si queremos describir una variable a nivel de personas, una manera correcta sería decir “Un XX% de los habitantes del área urbana del país ha asistido a actividades culturales y/o patrimoniales y/o infraestructuras culturales en los últimos doce meses” (filtro_pc == 1), mientras que en las preguntas acerca del comportamiento lector del hogar del módulo de Datos Sociodemográficos, una forma apropiada de describir el resultado sería diciendo: “El XX% de los hogares del área urbana del país cuentan con 10 libros o menos” (rph_comp_lector_libros == 1).

4.3. USO DE FACTORES DE EXPANSIÓN

Para una correcta lectura de los datos es necesario utilizar el factor de expansión adecuado a cada unidad de análisis. El factor de expansión de la ENPCCL 2024 está construido según el diseño muestral⁷, por lo que se interpreta como la cantidad de personas u hogares que representa sobre el total y es calculado como el inverso de la probabilidad de selección de la unidad muestral.

Los factores de expansión permiten obtener estimaciones para cada una de las variables consideradas en la ENPCCL 2024, lo que depende de la unidad de análisis asociada a cada variable:

- **Fexp_p**: Factor de expansión regional para variables a nivel de Persona.
- **Fexp_h**: Factor de expansión regional para variables a nivel de Hogar.

A continuación, en el cuadro se incluye una guía que detalla el factor de expansión asociado a las principales variables y su unidad de análisis correspondiente según sección del cuestionario.

Tabla 4. Factores de expansión asociados a las diferentes unidades de análisis

MÓDULOS	SUBSECCIÓN/VARIABLE	UNIDAD DE ANÁLISIS	FACTOR DE EXPANSIÓN
---------	---------------------	--------------------	---------------------

⁷ El diseño muestral de ENPCCL corresponde a un diseño complejo trietápico. Para mayor detalle consultar Informe Metodológico ENPCCL 2024.

Datos Sociodemográficos	Comportamiento y hábitos de lectura en el hogar ⁸	Hogar	fexp_h
Cultura general	Todas las preguntas	Persona	fexp_p
Actividades artísticas y patrimoniales	Asistencia a actividades artísticas y patrimoniales (<i>screening</i>)	Persona	fexp_p
	Caracterización asistencia actividades artísticas y patrimoniales	Persona	fexp_p
	Motivos de no asistencia a actividades artísticas y patrimoniales	Persona	fexp_p
Infraestructuras culturales	Asistencia a infraestructuras culturales (<i>screening</i>)	Persona	fexp_p
	Caracterización asistencia a infraestructuras culturales	Persona	fexp_p
	Motivos de no asistencia a infraestructuras culturales	Persona	fexp_p
	Acceso a infraestructuras culturales	Persona	fexp_p
Comportamiento lector	Lectura	Persona	fexp_p
	Libros	Persona	fexp_p
Entorno digital	Todas las preguntas	Persona	fexp_p
Participación inventiva	Todas las preguntas	Persona	fexp_p
Otros indicadores	Todas las preguntas	Persona	fexp_p
Orientación sexual	Todas las preguntas	Persona	fexp_p

4.4. TEST DE HIPÓTESIS

La base de datos permite realizar pruebas de hipótesis⁹ para determinar si las proporciones y/o medias observadas en un indicador varían con diferencias estadísticamente significativas entre distintos grupos de análisis. En la revisión de difusión de la ENPCCL 2024 se realizaron diferentes contrastes entre subpoblaciones, a continuación, se mencionan algunos de ellos:

⁸ Las preguntas que usan factor de expansión hogar corresponden a: rph_comp_lector_actividades_a, rph_comp_lector_actividades_b, rph_comp_lector_actividades_c, rph_comp_lector_actividades_d, rph_comp_lector_actividades_e, rph_comp_lector_actividades_f, rph_comp_lector_actividades_g, rph_comp_lector_actividades_h y rph_comp_lector_libros

⁹ La realización de contrastes de hipótesis debe estar sujeta a evaluación. Se recomienda considerar criterios de calidad de la desagregación, composición de la variable, potencia de la prueba, entre otros.

- Contraste de proporciones y/o medias según sexo:
 - Hombre - Mujer
- Contraste de proporciones y/o medias según tramo etario:
 - Jóvenes - Adultas
 - Jóvenes - Adultas mayores
 - Adultas – Adultas mayores
- Contraste de proporciones y/o medias según nivel socioeconómico
 - Bajo – Medio
 - Bajo – Alto
 - Medio - Alto
- Contraste de proporciones y/o medias según nivel educacional
 - Básico – Medio
 - Básico – Superior
 - Medio – Superior
- Contraste de proporciones y/o medias según presencia de persona en situación de discapacidad en el hogar de la persona entrevistada.
 - Con integrante - Sin integrante
- Contraste de proporciones y/o medias según situación de discapacidad de la persona entrevistada.
 - Con situación de discapacidad - Sin situación de discapacidad
- Contraste de proporciones y/o medias según presencia de niñas o niños de 6 años en el hogar de la persona entrevistada.
 - Con presencia de niña o niño de 6 años - Sin presencia de niña o niño de 6 años

En términos generales, se busca comparar la proporción y/o media P_i de personas u hogares pertenecientes al grupo i (por ejemplo, hombres), respecto a un indicador determinado, con la proporción y/o media P_k del mismo indicador en otro grupo k (por ejemplo, mujeres), mediante el siguiente contraste:

$H_0: P_i = P_k$ versus $H_1: P_i \neq P_k$

Donde:

Hipótesis nula (H_0): No existe diferencia significativa entre las proporciones y/o medias del indicador entre los grupos i , y k : $P_i = P_k$ (por ejemplo, no existe diferencia estadísticamente significativa entre hombres y mujeres).

Hipótesis alternativa (H_1): Existe una diferencia significativa entre las proporciones y/o medias del indicador entre los grupos i , y k : $P_i \neq P_k$ (por ejemplo, existe diferencia estadísticamente significativa entre hombres y mujeres).

Para evaluar estas diferencias, se utiliza el test t adaptado a diseños muestrales complejos (*svyttest* del paquete *survey*). Esta herramienta permite incorporar las características del diseño como estratificación, conglomerados y ponderadores, lo que asegura inferencias válidas. Se debe considerar que la variable de interés sea binaria.

A modo de ejemplo, se presenta la comparación entre hombres y mujeres para el indicador de Participación cultural (*filtro_pc*):

```
# Filtramos por Kish
enpcc12024_Kish <- enpcc12024 %>% filter(kish_dummy==1)

#Definimos diseño complejo
DC <- svydesign(ids = ~Varunit ,
               strata = ~Estrato_ENPC,
               weights = ~fexp_p,
               check.strata = TRUE,
               data = enpcc12024_Kish)
options(survey.lonely.psu = "certainty")
#Aplicamos t-test
tt <- svyttest(formula = filtro_pc ~ rph_sexo, design = DC)
#Imprimir resultado en consola
print(tt)
```

Para la aplicación del test t, la variable de agrupación debe ser dicotómica, es decir, tener únicamente dos categorías. En casos donde la variable de interés presenta más de dos categorías, como nivel socioeconómico (NSE) que incluye tres niveles, es necesario realizar

comparaciones por pares. Esto implica filtrar o excluir una de las categorías para contrastar solo las dos restantes, repitiendo el proceso para cada par de interés.

4.5. MANEJO DE DATOS CON CATEGORÍAS NO APLICA

Para la construcción de algunos indicadores y el cálculo de sus respectivas proporciones, es necesario realizar un filtro previo en el que se marginen del cálculo aquellas observaciones a los que no aplica el indicador. Esto responde a la necesidad de calcular el porcentaje de participación en actividades culturales y/o infraestructuras sobre la población efectiva que puede realizar esta acción. Por ejemplo, para el cálculo del promedio de libros leídos por gusto y/u ocio (libros_ocio) es necesario que las personas entrevistadas reportaran leer libros y que declaran leer, en general, por gusto y/u ocio, es decir, se debe aplicar un filtro de lectura de libros y lectura por gusto y/u ocio. En el agregado, esto se reflejará en que el promedio de libros leídos por gusto y/u ocio se calculará con aquellas personas que leyeron libros y leyeron por gusto y/u ocio.

A continuación, se agrega el procedimiento en R de filtro por lectura de libros y lectura por gusto y/u ocio para el indicador “Libros Ocio” a nivel nacional:

#Para crear el filtro de libros leídos por ocio y/u gusto es necesario considerar quienes leen libros (lectura_libros) y quienes leen por gusto y/u ocio.

```
# Filtramos por Kish, Lectura de Libros y Lectura por gusto y/o ocio
enpccl2024_libros<-enpccl2024 %>%
  filter(kish_dummy==1) %>%
  filter(lectura_libros %in% c(1:6) & lectura_tipo_3==1)
```

#Definimos diseño complejo

```
DC_libros <- svydesign(ids = ~Varunit ,
  strata = ~Estrato_ENPC,
  weights = ~fexp_p,
  check.strata = TRUE,
  data = enpccl2024_libros)
```

```
options(survey.lonely.psu = "certainty")
```

```
promLibrosOcio <- survey::svymean(~libros_ocio,design = DC_libros,na.rm =
T)
```

Un procedimiento similar puede ser aplicado para estimar otros indicadores cuyas variables de origen contienen la categoría de no aplica, como, por ejemplo, algunas variables de comportamiento lector en el hogar, que usan el factor de expansión de hogar, requieren la existencia de un integrante de 14 años o menos ($\text{niño} == 1 \mid \text{niño_2} == 1$)¹⁰ o de 6 a 14 años ($\text{niño_2} == 1$)¹¹ en el hogar kish, o las de caracterización de asistencia a actividades culturales y/o patrimoniales (filtro_teatro , filtro_cine , entre otras), para así asegurar que se están aplicando análisis sobre la población correcta y no sobre aquellos casos que deben excluirse dado que la situación consultada no es aplicable a su caso (Ej.: asistencia a una obra de teatro de personas que no hayan asistido a una obra de teatro en los últimos doce meses).

¹⁰ Las preguntas en las que aplica esta regla corresponden a $\text{rph_comp_lector_actividades_a}$, $\text{rph_comp_lector_actividades_b}$, $\text{rph_comp_lector_actividades_c}$, $\text{rph_comp_lector_actividades_d}$ y $\text{rph_comp_lector_actividades_e}$

¹¹ Las preguntas en las que aplica esta regla corresponden a $\text{rph_comp_lector_actividades_f}$, $\text{rph_comp_lector_actividades_g}$ y $\text{rph_comp_lector_actividades_h}$

CAPÍTULO V: INDICADORES AGREGADOS

Los indicadores agregados (Tabla 6), corresponden a los principales indicadores que son creados a partir de otras variables presentes en la base de datos. Son los indicadores más importantes para los objetivos de la ENPCCL 2024, dado que miden la participación cultural y el comportamiento lector de personas.

Tabla 5. Indicadores agregados

VARIABLE	NOMBRE DE VARIABLE	CÓDIGOS	FACTOR DE EXPANSIÓN	DOMINIO DE ESTIMACIÓN
pc1	Participación cultural - Asistencia	1 a 10	Persona	Nacional Regional
filtro_pc	Participación cultural	1 Sí 0 No	Persona	Nacional Regional
ind_var	Participación cultural – Variedad	0 Sin asistencia 1 Baja 2 Media 3 Alta	Persona	Nacional Regional
ind_int	Participación cultural – Intensidad	0 Sin frecuencia 1 Baja 2 Media 3 Alta	Persona	Nacional Regional
pc_2017	Participación cultural ENPC 2017	0 a 14	Persona	Nacional Regional
pc_unesco	Participación cultural Unesco/Protocolo San Salvador	0 a 17	Persona	Nacional Regional

VARIABLE	NOMBRE DE VARIABLE	CÓDIGOS	FACTOR DE EXPANSIÓN	DOMINIO DE ESTIMACIÓN
pc_b	Asistencia a actividades culturales y patrimoniales fuera del hogar	0 a 11	Persona	Nacional Regional
pc_c	Asistencia a infraestructuras culturales	0 a 6	Persona	Nacional Regional
gc_a	Gasto en actividades culturales	1 Sí 0 No	Persona	Nacional Regional
i_int	Interacción cultural en entornos digitales	1 Sí 0 No	Persona	Nacional Regional
i_sim	Interacción cultural digital: simultaneidad de actividades	1 Sí 0 No	Persona	Nacional Regional
i_acc	Acceso a espacios culturales	1 Sí 0 No	Persona	Nacional Regional
lt_f	Lectura	1 Sí 0 No	Persona	Nacional Regional
pi	Participación cultural inventiva	0 a 9	Persona	Nacional Regional
ind_inc_inf	Incentivo de prácticas culturales y lectoras en la infancia	1 Sí 0 No	Persona	Nacional Regional
ind_inc_pc	Incentivo de prácticas lectoras en la infancia	1 Sí 0 No	Persona	Nacional Regional
ind_inc_cl	Incentivo en la infancia: Comportamiento lector	1 Sí 0 No	Persona	Nacional Regional

VARIABLE	NOMBRE DE VARIABLE	CÓDIGOS	FACTOR DE EXPANSIÓN	DOMINIO DE ESTIMACIÓN
ind_lec_dif	Autoeficacia lectora	1 Sí 0 No	Persona	Nacional Regional
ind_asist_biblio	Asistencia a biblioteca	1 Sí 0 No	Persona	Nacional Regional

La forma correcta de interpretación de estos indicadores, por tanto, considera tener en cuenta el uso correcto del factor de expansión, por ejemplo:

Participación cultural: Corresponde a las personas que asistieron a una o más actividades culturales y patrimoniales o infraestructuras culturales de un listado de diez actividades/infraestructuras en los últimos doce meses¹².

A modo de ejemplo, se incluye la sintaxis de construcción de los indicadores utilizando R:

Participación cultural

```
condpc1<- c("filtro_teatro",
            "filtro_danza",
            "filtro_musica",
            "filtro_cine",
            "filtro_arte",
            "filtro_patrimonio",
            "filtro_biblio",
            "filtro_museo",
            "filtro_arqueo",
            "filtro_memoria")

enpcc12024$pc1 <- ifelse(test = enpcc12024$kish_dummy == 0,
                        yes = NA,
                        no = rowSums(enpcc12024[, condpc1] == 1, na.rm =
TRUE))
```

Participación cultural – Asistencia

¹² Para mayor detalle respecto de la construcción de indicadores consultar Informe Metodológico ENPCCL 2024

```
condfiltro_pc<- (enpcc12024_Kish$filtro_teatro==1|
  enpcc12024_Kish$filtro_danza==1|
  enpcc12024_Kish$filtro_musica==1|
  enpcc12024_Kish$filtro_cine==1|
  enpcc12024_Kish$filtro_arte==1|
  enpcc12024_Kish$filtro_patrimonio==1|
  enpcc12024_Kish$filtro_biblio==1|
  enpcc12024_Kish$filtro_museo==1|
  enpcc12024_Kish$filtro_arqueo==1|
  enpcc12024_Kish$filtro_memoria==1)

enpcc12024_Kish$filtro_pc<-ifelse(test = condfiltro_pc, yes = 1,
  no = 0)
```

Participación cultural – Variedad

```
enpcc12024$ind_var <- ifelse(test = enpcc12024$kish_dummy==0,
  yes = NA,
  no = ifelse(enpcc12024$pc1 == 0, 0,
    ifelse(enpcc12024$pc1 %in%
1:2, 1,
    ifelse(enpcc12024$pc1
%in% 3:4, 2,
    ifelse(enpcc12024$pc1 >= 5, 3, NA))))))

enpcc12024$ind_var <- labelled::labelled(enpcc12024$ind_var,
  labels = c("Sin asistencia" =
0,
  "Baja" = 1,
  "Media" = 2,
  "Alta" = 3))
```

Participación cultural – Intensidad

#Creación de variable auxiliar de cantidad de veces para exposición artística considerando valores negativos

```
enpcc12024$arte_veces_aux<-ifelse(test = enpcc12024$kish_dummy == 0,
  yes = NA,
```



```

no = ifelse(test = enpcc12024$arte_veces <
0,

yes = 1,
no = enpcc12024$arte_veces))

#Variables usadas para el cálculo numérico de intensidad

condIntensidad <- c(
  "teatro_veces",
  "danza_veces",
  "musica_veces",
  "cine_veces",
  "arte_veces_aux",
  "filtro_patrimonio",
  "biblio_veces",
  "filtro_memoria",
  "filtro_arqueo",
  "museo_veces")

enpcc12024$ind_int <- ifelse(test = enpcc12024$kish_dummy==0,
yes = NA,
no =
rowSums(enpcc12024[,condIntensidad],na.rm=T))

enpcc12024$ind_int1 <- ifelse(test = enpcc12024$kish_dummy==0,
yes = NA,
no = ifelse(enpcc12024$ind_int == 0,0,
ifelse(enpcc12024$ind_int
%in% 1:3,1,

ifelse(enpcc12024$ind_int %in% 4:9,2,

ifelse(enpcc12024$ind_int >= 10, 3, NA))))))

enpcc12024$ind_int <- labelled::labelled(enpcc12024$ind_int,
labels = c("Sin frecuencia" =
0,

"Baja" = 1,
"Media" = 2,
"Alta" = 3))

```

Participación cultural 2017

```

enpcc12024$filtro_videojuegos_frec <- enpcc12024$videojuegos_frec %in% c(
1,2,3,4,5,6)

enpcc12024$filtro_lectura_libros <- enpcc12024$lectura_libros %in% c(1,2,
3,4,5,6)

condpc_2017<-c("filtro_teatro",
               "filtro_circo",
               "filtro_danza",
               "filtro_musica",
               "filtro_cine",
               "filtro_arte",
               "filtro_ppoo",
               "filtro_rel",
               "filtro_biblio",
               "filtro_museo",
               "filtro_natural",
               "filtro_arqueo",
               "filtro_videojuegos_frec",
               "filtro_lectura_libros")

enpcc12024$pc_2017 <- ifelse(test = enpcc12024$kish_dummy == 0,
                             yes = NA,
                             no = rowSums(enpcc12024[, condpc_2017
] == 1, na.rm = TRUE))

```

Asistencia a espacios culturales

```

condpc_b<- c("filtro_biblio",
             "filtro_museo",
             "filtro_cc",
             "filtro_natural",
             "filtro_arqueo",
             "filtro_memoria")

enpcc12024$pc_b <- ifelse(test = enpcc12024$kish_dummy == 0,
                           yes = NA,
                           no = rowSums(enpcc12024[, condpc_b] == 1
, na.rm = TRUE))

```

Participación en actividades culturales fuera del hogar

```

condpc_c<- c( "filtro_teatro",
               "filtro_opera",
               "filtro_circo",
               "filtro_danza",
               "filtro_musica",
               "filtro_cine",
               "filtro_arte",

```

```

        "filtro_artesania",
        "filtro_patrimonio",
        "filtro_ppoo",
        "filtro_rel")

enpccl2024$pc_c <- ifelse(test = enpccl2024$kish_dummy == 0,
                        yes = NA,
                        no = rowSums(enpccl2024[, condpc_c] == 1
, na.rm = TRUE))

```

Participación cultural UNESCO

```
enpccl2024$pc_unesco<- (enpccl2024$pc_b + enpccl2024$pc_c)
```

Gasto en actividades culturales

```

condgc_a<- (enpccl2024$gc_teatro==1|
            enpccl2024$gc_danza==1|
            enpccl2024$gc_musica==1|
            enpccl2024$gc_cine==1|
            enpccl2024$gc_arte==1|
            enpccl2024$gc_artesania==1|
            enpccl2024$gc_libros==1|
            enpccl2024$gc_artesania_obj==1)

enpccl2024$gc_a <- ifelse(test = enpccl2024$kish_dummy == 0,
                        yes = NA,
                        no = ifelse(test = condgc_a, yes = 1,
                                    no = 0))

```

Interacción cultural en internet

```

condi_int<- (enpccl2024$internet_videojuegos==1|
            enpccl2024$internet_musica==1|
            enpccl2024$internet_conciertos==1|
            enpccl2024$internet_podcast==1|
            enpccl2024$internet_teatro==1|
            enpccl2024$internet_pelicula==1|
            enpccl2024$internet_infra==1)

enpccl2024$i_int <-ifelse(test = enpccl2024$kish_dummy == 0,
                        yes = NA,
                        no = ifelse(test = condi_int, yes = 1,
                                    no = 0))

```

Simultaneidad

```
sim_vars <- c("sim_a", "sim_b", "sim_c", "sim_d", "sim_e",
             "sim_f", "sim_g", "sim_h", "sim_i", "sim_j")
```

```
# Contar 1
sim_sum <- rowSums(enpccl2024[sim_vars] == 1, na.rm = TRUE)

# Dicotomizar
enpccl2024$i_sim <- ifelse(test = enpccl2024$kish_dummy == 0,
                           yes = NA,
                           no = ifelse(test = sim_sum >= 1, yes = 1
, no = 0))
```

Conocimiento de espacios culturales

```
condi_acc<- (enpccl2024$acceso_teatro %in% c(1,2,3,4,5)|
             enpccl2024$acceso_cine %in% c(1,2,3,4,5)|
             enpccl2024$acceso_biblio %in% c(1,2,3,4,5)|
             enpccl2024$acceso_museo %in% c(1,2,3,4,5)|
             enpccl2024$acceso_cc %in% c(1,2,3,4,5))

enpccl2024$i_acc<-ifelse(test = enpccl2024$kish_dummy == 0,
                        yes = NA,
                        no = ifelse(test = condi_acc, yes = 1,no
= 0))
```

Lectura

```
condlt_f<- (enpccl2024$lectura_diarios %in% c(1,2)|
            enpccl2024$lectura_laboral %in% c(1,2)|
            enpccl2024$lectura_comic %in% c(1,2)|
            enpccl2024$lectura_libros %in% c(1,2))

enpccl2024$lt_f<-ifelse(test = enpccl2024$kish_dummy == 0,
                        yes = NA,
                        no = ifelse(test = condlt_f, yes = 1,no =
0))
```

Incentivo infancia – Participación cultural

```
condind_inc_pc<- (enpccl2024$otros_infra == 1|
                  enpccl2024$otros_teatro ==1|
                  enpccl2024$otros_cine == 1|
                  enpccl2024$otros_musica == 1)

enpccl2024$ind_inc_pc<-ifelse(test = enpccl2024$kish_dummy == 0,
                              yes = NA,
                              no = ifelse(test = condind_inc_pc, y
es = 1, no = 0))
```

Incentivo infancia - Comportamiento lector

```
condind_inc_cl<- (enpccl2024$incent_cuentos == 1|
                  enpccl2024$incent_libros == 1)

enpccl2024$ind_inc_cl<-ifelse(test = enpccl2024$kish_dummy == 0,
                              yes = NA,
                              no = ifelse(test = condind_inc_cl, y
es = 1,no = 0))
```

Incentivo infancia

```
condind_inc_inf<- (enpccl2024$ind_inc_pc == 1|
                  enpccl2024$ind_inc_cl ==1)

enpccl2024$ind_inc_inf<-ifelse(test = enpccl2024$kish_dummy == 0,
                              yes = NA,
                              no = ifelse(test = condind_inc_inf,
yes = 1,no = 0))
```

Autoeficacia lectora

```
condind_lec_dif<- (enpccl2024$lectura_dificultad_1 %in% c(1,2)|
                  enpccl2024$lectura_dificultad_2 %in% c(1,2)|
                  enpccl2024$lectura_dificultad_3 %in% c(1,2)|
                  enpccl2024$lectura_dificultad_4 %in% c(1,2)|
                  enpccl2024$lectura_dificultad_5 %in% c(1,2))

enpccl2024$ind_lec_dif_manual<-ifelse(test = enpccl2024$kish_dummy == 0,
                                       yes = NA,
                                       no = ifelse(test = condind_lec_dif,
yes = 1, no = 0))
```

Asistencia a biblioteca

```
condind_asist_biblio<- (enpccl2024$filtro_biblio == 1)

enpccl2024$ind_asist_biblio<-ifelse(test = enpccl2024$kish_dummy == 0,
                                    yes = NA,
                                    no = ifelse(test = condind_asist_biblio, yes = 1,no = 0))
```

Participación cultural inventiva

```
condpi<- c("inv_malabarismo",
            "inv_teatro",
            "inv_danza",
            "inv_musica",
            "inv_cine",
            "inv_plasticas",
            "inv_artesania",
```

```
      "inv_foto",  
      "inv_cuentos")  
enpccl2024$pi <- ifelse(test = enpccl2024$kish_dummy == 0,  
                        yes = NA,  
                        no = rowSums(enpccl2024[, condpi] == 1, na.rm =  
TRUE))
```

CAPÍTULO VI: CALIDAD DE LAS ESTIMACIONES

La evaluación de la calidad estadística de las estimaciones se basa en el “Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares”, publicado por el INE en febrero de 2020. Este estándar surge como parte de la producción de estándares metodológicos para proveer al Sistema Estadístico Nacional (SEN) y busca establecer los criterios que permitan caracterizar qué tan precisa y confiable es la información que se publica respecto a los atributos de la población que se pretende caracterizar.

De acuerdo con este estándar, cada estimación se evalúa siguiendo criterios específicos. En términos de diseño muestral, se considera el tamaño de la muestra y los grados de libertad de la estimación, mientras que en medidas de dispersión se utiliza el error estándar cuando el estimador corresponde a una proporción o razón entre 0 y 1. Para el resto de las estimaciones (medias, quintiles, etc.), se utiliza el coeficiente de variación.

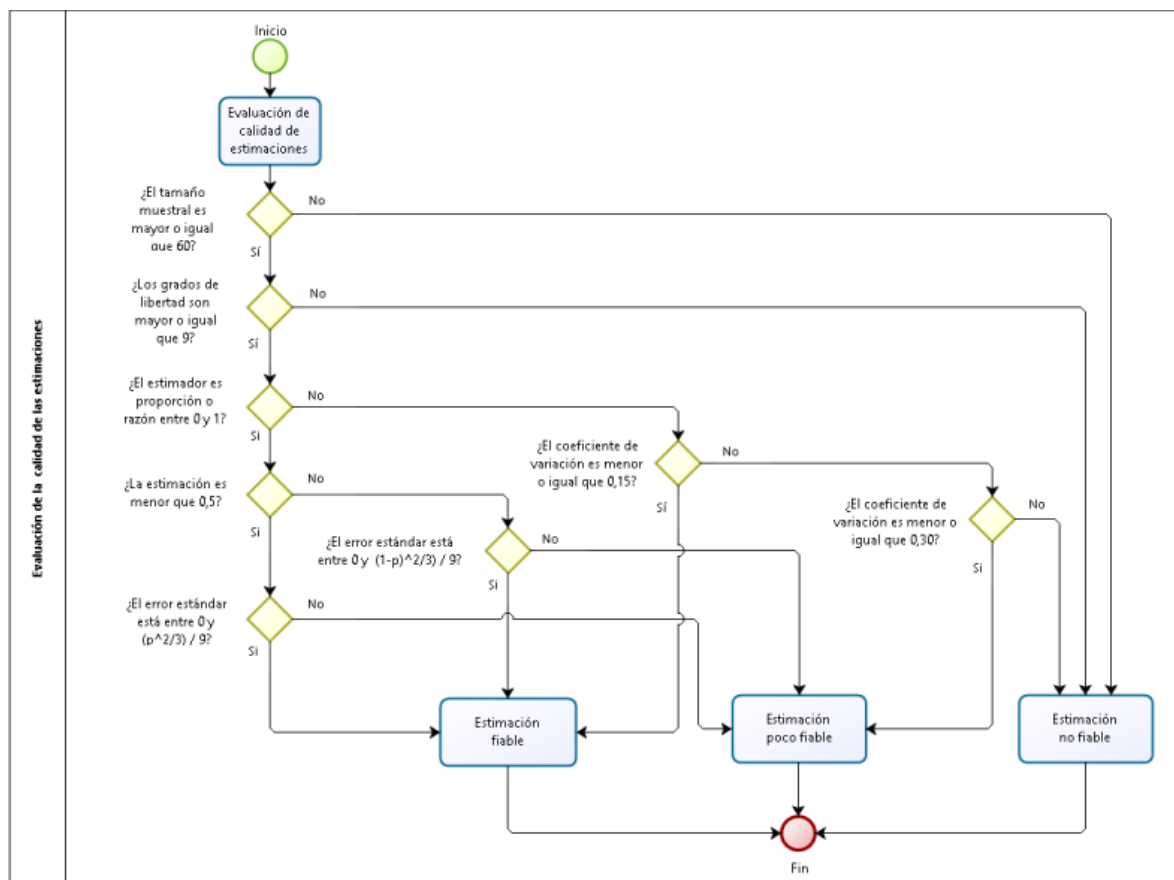
De acuerdo con el flujograma de decisiones definido en el estándar, las estimaciones pueden ser clasificadas en tres tipos: estimaciones fiables, estimaciones poco fiables y estimaciones no fiables. Esta clasificación de las estimaciones es reportada tanto en la síntesis de resultados como en los tabulados publicados.

- Las *estimaciones fiables* corresponden a aquellas que cumplen con todos los criterios de calidad propuestos, por tanto, pueden ser reconocidas como estimaciones confiables y precisas de un parámetro poblacional.
- Las *estimaciones poco fiables* son aquellas que cumplen los criterios del tamaño muestral, al menos 60 observaciones, y grados de libertad, al menos 9, pero que no se ajustan a los requerimientos establecidos para el error estándar máximo o el coeficiente de variación menor o igual al umbral del 15%. Es por esta razón que se recomienda utilizar esta estimación con precaución para interpretar características poblacionales, ya que podrían llevar a conclusiones poco acertadas.
- Las *estimaciones no fiables* corresponden a aquellas que no cumplen con alguno de los dos primeros criterios establecidos para el tamaño muestral y/o los grados de libertad. También se incluyen dentro de esta categoría a aquellas estimaciones que

no son proporciones ni razones entre 0 y 1 cuyo coeficiente de variación supere 30%. Para estos casos se debe consignar que no se recomienda el uso de estas estimaciones.

Para aplicar este flujo de calidad a las estimaciones se sugiere instalar y utilizar la librería “calidad” desarrollada por el INE para este propósito, y que se encuentra alojada en el repositorio CRAN del lenguaje R. Esta librería permite evaluar la calidad de las estimaciones, por ejemplo, de las medias de las variables de la base de datos, de forma ágil porque el output que genera incluye los estadígrafos (error estándar, número de observaciones, error cuadrático medio, entre otros) usados para evaluar la calidad y una columna con el resultado cualitativo (“fiable”, “poco fiable”, “no fiable”), lo que posibilita una inspección visual de los criterios y un resultado de los mismos.

Figura 1. Flujo de Evaluación de Calidad de las Estimaciones



Fuente: Instituto Nacional de Estadísticas de Chile