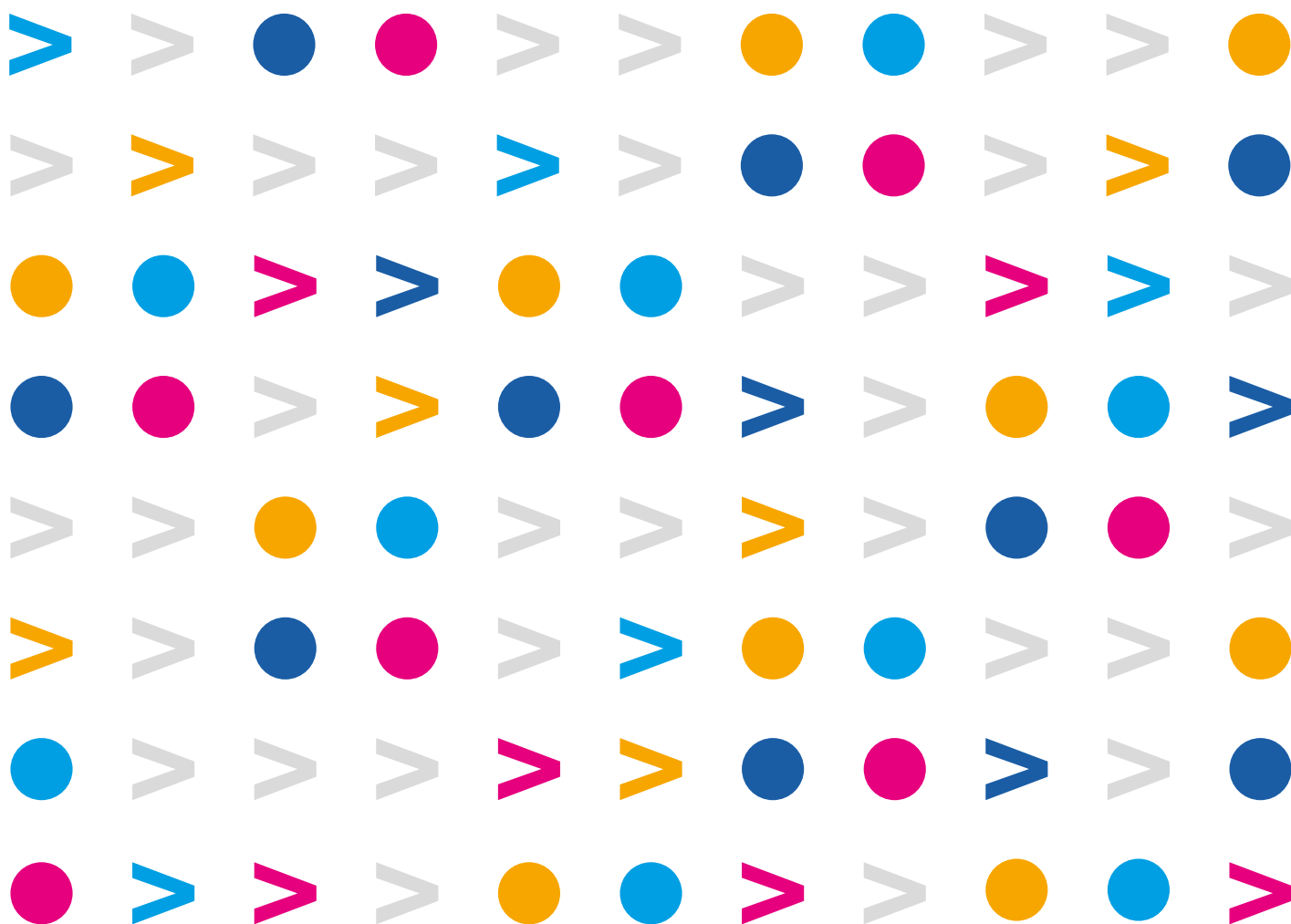




Manual de Uso de **Microdatos Censales**



Índice general

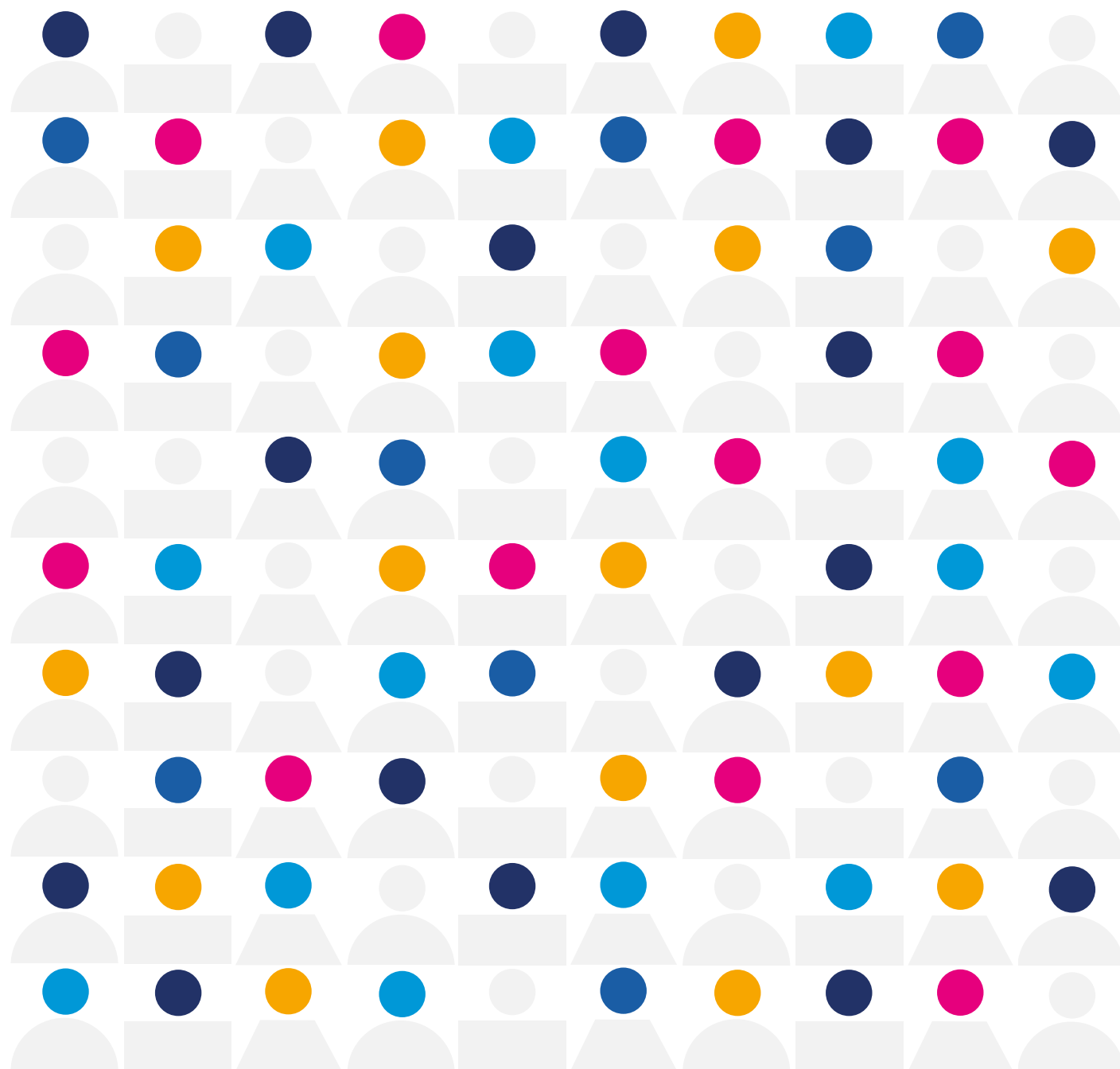
1	Presentación	5
2	Antecedentes y características generales del Censo 2024	7
	2.1. Objetivos del Censo 2024	8
	2.2. Metodología y unidades de análisis	8
	2.3. Modalidades de recolección	9
	2.4. Tipos de operativo de recolección	10
	2.5. Características del cuestionario censal	12
3	Acceso y solicitud de los datos	13
	3.1. Descarga y lectura de las bases de datos	14
4	Estructura de las bases de datos	15
	4.1. Organización de la base de datos	16
	4.2. Correspondencia entre preguntas del cuestionario y variables en la base de datos	21
	4.3. Variables calculadas	25
5	Control de divulgación estadística	45
6	Diccionario de variables	51
	6.1 Valores especiales	52
	6.2 Variables de viviendas	53
	6.3 Variables de hogar	56
	6.4 Variables de personas	59
	6.5 Diccionario de códigos territoriales	71
	6.6 Diccionario de códigos territoriales específicos	71

Definición y construcción de indicadores	72
7.1. Aspectos técnicos generales	73
7.1.1. Redondeo de decimales	73
7.1.2. Manejo de la no respuesta al ítem	73
7.1.3. Cargar paquetes	73
7.1.4. Tipos de variables y carga de datos	74
7.1.5. Desagregaciones temáticas y geográficas	74
7.2. Ficha de indicadores	75
7.2.1. dem01: Estructura sexo-edad (para pirámide poblacional)	75
7.2.2. dem02: Razón hombre-mujer (índice de masculinidad)	75
7.2.3. dem03: Índice de envejecimiento	76
7.2.4. dem04: Relación de dependencia total	76
7.2.5. dem05: Relación de dependencia de niños/as y jóvenes	77
7.2.6. dem06: Relación de dependencia de personas mayores	77
7.2.7. dem07: Porcentaje de hogares de personas mayores	78
7.2.8. dem08: Porcentaje de hogares con niños/as	79
7.2.9. dem09: Porcentaje de hogares unipersonales	79
7.2.10. dem10: Promedio de personas por hogar	80
7.2.11. dem11: Mediana de edad	80
7.2.12. dem12: Porcentaje de hogares con jefatura de hogar mujer	81
7.2.13. mig01: Porcentaje de población inmigrante internacional	81
7.2.14. fec01: Razón niño-mujer	82
7.2.15. fec02: Promedio de hijos nacidos vivos de mujeres entre 15-49 años (paridez media)	82
7.2.16. fec03: Porcentaje de mujeres entre 15-49 años con hijos	83
7.2.17. fec04: Porcentaje de mujeres entre 15-49 años sin hijos	83
7.2.18. mig_int01: Matriz de migración interna (nivel regional)	84
7.2.19. mig_int02: Indicadores derivados de la matriz de migración interna	85
7.2.20. mig_int03: Condición migratoria (nivel regional)	89
7.2.21. educ01: Promedio de años de escolaridad	90
7.2.22. educ02: Tasa de alfabetización	90
7.2.23. educ03: Tasa de asistencia neta a la educación parvularia	91
7.2.24. educ04: Tasa de asistencia neta a la educación básica	92

7.2.25. educ05: Tasa de asistencia neta a la educación media.....	92
7.2.26. educ06: Tasa de asistencia neta a la educación superior.....	93
7.2.27. disc01: Discapacidad.....	94
7.2.28. pueblos01: Porcentaje de la población que es o se considera perteneciente a un pueblo indígena u originario.....	95
7.2.29. pueblos02: Porcentaje de la población afrodescendiente.....	95
7.2.30. conm01: Matriz de conmutación laboral.....	96
7.2.31. conm02: Condición de conmutación laboral (nivel regional).....	97
7.2.32. hog01: Porcentaje de hogares con acceso a internet.....	98
7.2.33. hog02: Porcentaje de hogares con tenencia insegura de la vivienda.....	99
7.2.34. hog03: Tipología de hogar.....	99
7.2.35. viv01: Promedio de hogares en la vivienda.....	101
7.2.36. viv02: Promedio de personas por vivienda.....	101
7.2.37. viv03: Índice de hacinamiento.....	102
7.2.38. viv04: Índice de materialidad.....	103
7.2.39.viv05: Índice de tipo de vivienda.....	104
7.2.40. viv06: Viviendas irrecuperables.....	105
7.2.41. viv07: Hogares allegados.....	106
7.2.42. viv08: Núcleos allegados hacinados.....	106
7.2.44. viv10: Déficit habitacional cuantitativo.....	111
7.2.45. nr01: Tasa de no respuesta al ítem.....	112
7.2.46. na01: Frecuencia de casos en los que no aplica respuesta.....	113
7.2.47. tip_op01: Frecuencia de casos por tipo de operativo.....	113

1.

Presentación



El Censo de Población y Vivienda 2024 constituye el principal operativo estadístico nacional para la medición y caracterización de la población, los hogares y las viviendas del país. A diferencia de operativos censales anteriores, el Censo 2024, realizado entre los meses de marzo y julio de 2024, corresponde a un censo de derecho, en el que las personas fueron registradas según su lugar de residencia habitual.

El presente manual tiene como objetivo servir como guía para el uso de los microdatos del Censo 2024, proporcionando orientaciones técnicas sobre la estructura de las bases de datos, contenido y procesamiento. Está dirigido a personas e instituciones que realicen análisis estadísticos, investigación académica o generación de información para la toma de decisiones tanto en el sector público como en el sector privado. Entre los principales destinatarios se encuentran analistas de datos, investigadores, organismos gubernamentales, organizaciones internacionales e instituciones privadas.

Las bases de microdatos se encuentran publicadas en formato *csv* y *parquet*, y están disponibles para su descarga en el portal oficial del Instituto Nacional de Estadísticas (INE) www.ine.cl.

Las programaciones proporcionadas en este documento se desarrollaron para su ejecución en R¹, usando para esto los paquetes *arrow*² y la colección de paquetes *tidyverse*³.

Los contenidos de este manual se dividen en siete capítulos. Después de la presentación, el segundo capítulo entrega los antecedentes y características generales del censo, incluyendo sus objetivos, metodología, unidades de análisis, cobertura geográfica, contenido del cuestionario y modalidades de recolección. El tercer capítulo describe el proceso para acceder a la base de datos, los formatos disponibles y las instrucciones para su apertura.

El cuarto capítulo presenta la estructura de los microdatos, describe las variables para la identificación geográfica, las llaves que permiten la vinculación entre las bases de vivienda, hogar y persona, así como la correspondencia entre el cuestionario publicado en la página web⁴ y las variables contenidas en los microdatos. El quinto capítulo resume los procedimientos de anonimización aplicados para resguardar el secreto estadístico y la confidencialidad de las personas.

Finalmente, los capítulos seis y siete incorporan el diccionario de variables y la ficha de indicadores. Esta última incluye definiciones, fórmulas y universo de aplicación para su cálculo, que permite replicar los indicadores publicados en <https://censo2024.ine.gob.cl/resultados/>.

¹ R es un lenguaje de programación y entorno de software libre, especializado en el análisis estadístico y manipulación de datos. Fue creado por Robert Gentleman y Ross Ihaka. Para más información y su descarga, visitar <https://www.r-project.org/>.

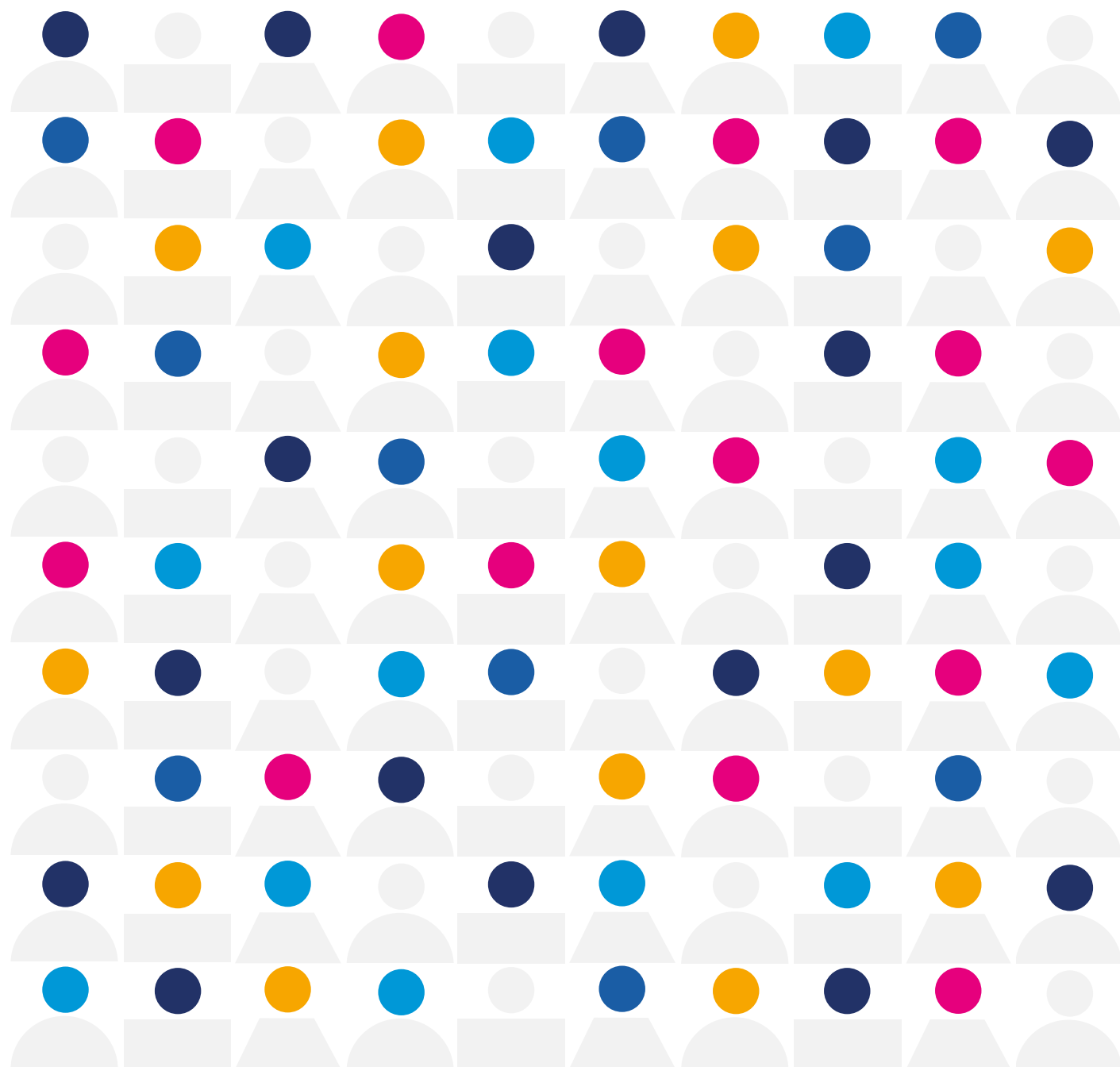
² El paquete *arrow* para R proporciona un *backend* de *dplyr* para trabajar con grandes conjuntos de datos. Para más información se recomienda consultar el siguiente enlace: <https://arrow.apache.org/>.

³ Este manual se escribió en base a los principios y la gramática de *tidyverse*, y se prioriza el uso de las funciones de los paquetes asociados a esta colección. Para saber más sobre como transformar, visualizar y modelar una estructura de datos utilizando R y Tidyverse en particular, se recomienda revisar el libro R para Ciencia de Datos disponible de manera gratuita en el siguiente enlace: <https://es.r4ds.hadley.nz/> Información específica sobre *tidyverse*, se encuentra disponible en el siguiente enlace: <https://www.tidyverse.org/packages/>.

⁴ El cuestionario publicado se puede descargar del siguiente enlace: https://censo2024.ine.gob.cl/wp-content/uploads/2025/03/Cuestionario-Viviendas-Particulares_CPV2024.pdf

2.

Antecedentes y características generales del Censo 2024



2.1. Objetivos del Censo 2024

El Censo de Población y Vivienda es una de las operaciones más complejas y masivas que puede realizar un país, siendo el Instituto Nacional de Estadísticas el organismo encargado de implementarlo.

El Censo 2024 tiene cuatro objetivos específicos:

1. Proveer información para la actualización y elaboración de las estimaciones y proyecciones de población.
2. Proporcionar información para la actualización del marco muestral para las encuestas de hogares.
3. Suministrar información estadística prioritaria para el diseño y seguimiento de políticas públicas.
4. Generar información estadística sobre áreas geográficas específicas y subgrupos de la población.

2.2. Metodología y unidades de análisis

El proceso de recolección del Censo de Población y Vivienda 2024 se realizó bajo una metodología de derecho. Esto quiere decir que las personas fueron contabilizadas en el lugar donde residen habitualmente, en vez del lugar donde pernoctaron la noche anterior, como se realizó en censos anteriores.

El Censo 2024 tuvo cobertura nacional y se aplicó en todo el territorio, incluyendo zonas urbanas y rurales, áreas de difícil acceso, localidades apartadas, personas en situación de calle y viviendas colectivas. Esto permitió que la información recolectada permitiera desagregaciones a distintos niveles geográficos, desde el país en su totalidad hasta regiones y comunas.

La población objetivo del Censo 2024 correspondió a todas las personas residentes habituales del territorio nacional al momento del censo, sin distinción de edad, sexo, nacionalidad, estatus migratorio u otra condición.

Las unidades de análisis, que también corresponden a las unidades censadas, son las viviendas, los hogares y las personas.

Los hogares permiten caracterizar la organización de las personas que residen habitualmente en una misma vivienda particular y que comparten un mismo presupuesto para alimentación. Las viviendas, en tanto, constituyen la unidad física de alojamiento para el registro de hogares y personas.

2.3. Modalidades de recolección

La modalidad de recolección se refiere a la forma mediante la cual se aplicó el cuestionario censal. Se definió por el medio empleado (físico o digital) y la forma de aplicación (presencial guiado por un o una censista o en línea autoadministrado). En el Censo 2024 se implementaron tres modalidades: la entrevista presencial con dispositivo móvil de captura (CAPI)¹, la entrevista autoaplicada en línea (CAWI)² y la entrevista presencial con cuestionario en papel (PAPI)³.

- 1. Entrevista presencial con dispositivo móvil de captura (CAPI):** fue la modalidad principal para realizar el registro de información. Consistió en una entrevista presencial realizada por censistas, utilizando el cuestionario censal por medio de la aplicación *Survey Solutions Interviewer* del Banco Mundial que se encontraba instalada en el dispositivo móvil de captura.
- 2. Entrevista autoaplicada en línea (CAWI):** fue una modalidad complementaria a la entrevista presencial con dispositivo móvil de captura. Esta refiere al uso de la plataforma web Censo en Línea para que las personas pudieran responder el cuestionario censal.
- 3. Entrevista presencial con cuestionario en papel (PAPI):** fue una modalidad diseñada para ser aplicada en casos de contingencia. En este caso, cada censista aplicaba el cuestionario censal a través del formulario impreso.

Tabla 2.1: Distribución de personas censadas por modalidad

Modalidad	Frecuencia	Porcentaje
CAPI	17.320.830	93,7 %
CAWI	1.154.397	6,2 %
PAPI	5.205	0,0 %
Total	18.480.432	100,0 %

Nota: dado el redondeo de los valores a un decimal estos pueden no sumar 100%.

¹ CAPI significa *Computer-Assisted Personal Interviewing*: método de recolección presencial en el que el encuestador lee las preguntas al encuestado y registra sus respuestas en dispositivos electrónicos.

² CAWI significa *Computer-Assisted Web Interviewing*: método de recolección que utiliza cuestionarios a través de internet, sin la necesidad de un entrevistador.

³ PAPI significa *Paper and Pencil Interviewing*: método de recolección presencial en el que el encuestador lee las preguntas al encuestado y registra sus respuestas en formulario en papel.

2.4. Tipos de operativo de recolección

La recolección correspondió al conjunto de procedimientos y acciones destinados a obtener la información censal en el territorio. Cada operativo se definió en función de la población objetivo y de las condiciones particulares de la recolección, lo que implicó una planificación diferenciada, con logística y recursos específicos para su ejecución.

Los tipos de operativos realizados en este censo fueron:

- Operativo de viviendas particulares.
- Operativo especial de viviendas colectivas.
- Operativo especial de personas en situación de calle.

2.4.1. Operativo de viviendas particulares

El operativo para la recolección de viviendas particulares fue el principal operativo de recolección. La mayor parte de los registros de viviendas particulares fueron capturados mediante la modalidad CAPI.

Tabla 2.2: Distribución de personas censadas por modalidad para operativo de viviendas particulares

Modalidad	Frecuencia	Porcentaje
CAPI	17.299.254	94,3 %
CAWI	1.035.317	5,6 %
PAPI	5.084	0,0 %
Total	18.339.655	100,0 %

Nota: dado el redondeo de los valores a un decimal estos pueden no sumar 100%.

2.4.2. Operativo de viviendas colectivas

El operativo para la recolección de viviendas colectivas fue un operativo especial considerado para la captura de la información de las personas que residen habitualmente en hoteles, pensiones, residenciales, hospitales, cárceles, internados, conventos, recintos militares, entre otros. La mayor parte de los registros de viviendas colectivas fueron capturados mediante la modalidad CAWI.

Tabla 2.3: Distribución de personas censadas por modalidad para operativo de viviendas colectivas

Modalidad	Frecuencia	Porcentaje
CAPI	2.267	1,9 %
CAWI	116.639	98,0 %
PAPI	121	0,1 %
Total	119.027	100,0 %

2.4.3. Operativo de personas en situación de calle

El operativo especial de personas en situación de calle fue diseñado para censar a aquellas personas sin residencia fija que pernoctan en espacios públicos o privados que, si bien cumplen funciones habitacionales, carecen de las características de una vivienda. En los alojamientos temporales, la recolección se realizó en función de cada territorio calle, capacitando a las y los encargados de estos lugares para que aplicaran las entrevistas mediante la plataforma de Censo en Línea. La mayor parte de los registros de personas en situación de calle se capturaron mediante la modalidad CAPI.

Tabla 2.4: Distribución de personas censadas por modalidad para Operativo de personas en situación de calle

Modalidad	Frecuencia	Porcentaje
CAPI	19.309	88,8 %
CAWI	2.441	11,2 %
PAPI	0	0,0 %
Total	21.750	100,0 %

2.5. Características del cuestionario censal

El cuestionario publicado en la página web tiene dos versiones: uno para el operativo de viviendas particulares y otro para los operativos especiales de viviendas colectivas o personas en situación de calle. La diferencia entre ambos cuestionarios radica en la inclusión de preguntas de vivienda y hogar, las que no se consideran en el cuestionario de operativos especiales.

La versión aplicada en viviendas particulares está conformada por 50 preguntas. Su estructura se organizó en cinco módulos temáticos:

- Datos de la vivienda (preguntas 1 a 10)
- Identificación de los hogares (preguntas 11 a 11.2)
- Datos del hogar (preguntas 12 a 15)
- Registro de personas del hogar (preguntas 16 a 18)
- Datos de las personas (19 a 50)

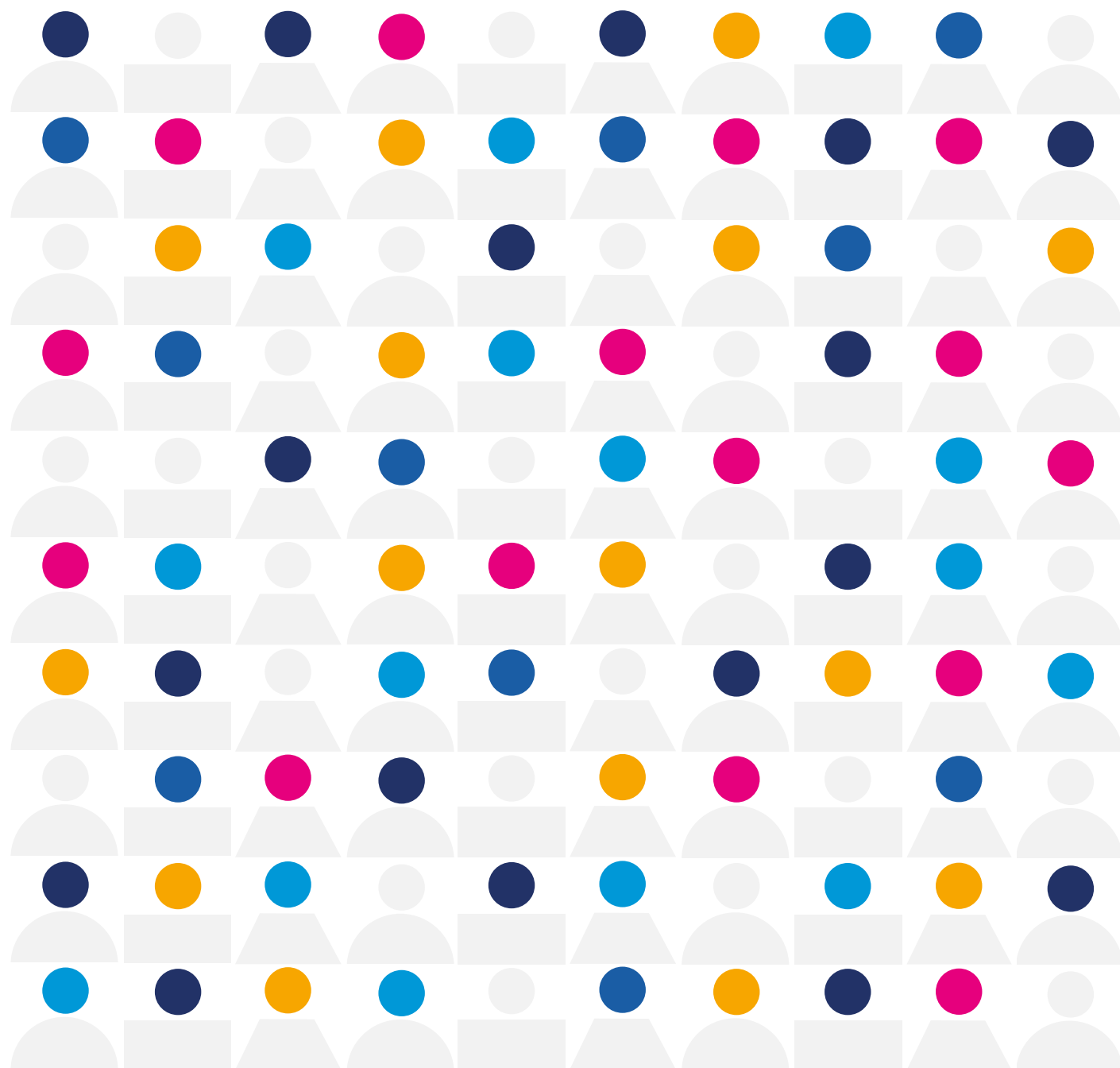
A su vez, el último módulo, datos de las personas, está conformado de ocho secciones:

- Datos de la persona (preguntas 19 a 23)
- Migración (preguntas 24 a 27)
- Pueblos y credos (preguntas 28 a 31)
- Discapacidad (pregunta 32)
- Educación (preguntas 33 a 37)
- Ocupación (preguntas 38 a 45)
- Fecundidad (preguntas 46 a 48)
- Género (preguntas 49 y 50)

La versión del cuestionario utilizada en viviendas colectivas y personas en situación de calle contiene solo las 32 preguntas correspondientes al módulo de datos de las personas (preguntas 20 a la 50).

3.

Acceso y solicitud de los datos



3.1. Descarga y lectura de las bases de datos

La base de datos del Censo 2024 se encuentra publicada y disponible para descarga en el portal web del Instituto Nacional de Estadísticas (INE) dentro de la sección “Resultados” (<https://censo2024.ine.gob.cl/resultados/>). Asimismo, las y los usuarios podrán realizar consultas y solicitudes relacionadas con la información entregada a través del portal del Sistema Integral de Información y Atención Ciudadana del INE (<https://atencionciudadana.ine.cl/>).

De ser requerida información adicional por parte de las y los usuarios, es importante considerar que los datos están sujetos a las siguientes condiciones:

- Disponibilidad de la información solicitada.
- Protección de la confidencialidad de la información, de acuerdo con los artículos 29° y 30° de Ley 17.374.

El formato disponible para la descarga es *csv* y *parquet*, organizado en tres niveles de información: vivienda, hogar y persona. Cada nivel se entrega en un archivo separado, los cuales pueden vincularse mediante “variables llave” que identifican la relación entre viviendas, hogares y personas.

Los archivos en formato *csv* constituyen un estándar abierto y tabular, ampliamente compatible con distintos programas, entre ellos el lenguaje de programación R, que será utilizado de manera exclusiva en este manual para todos los procedimientos asociados a la lectura de información, creación de variables y cálculo de indicadores del Censo 2024.

Por su parte, los archivos en formato *parquet* son un formato de almacenamiento en columnas de código abierto que se utiliza para almacenar, gestionar y analizar de manera eficiente grandes conjuntos de datos. A diferencia de los formatos de almacenamientos basados en filas, como *csv*, *parquet* organiza los datos en columnas para mejorar el rendimiento de las consultas. Cuando sea posible, se recomienda el uso del formato *parquet* por sobre el *csv*.

Tanto el archivo *parquet* como el archivo *csv* debe ser descomprimido para su uso. Después de descomprimirlo, puede ser cargado con funciones como `open_dataset` o `read_parquet` del paquete `arrow`.

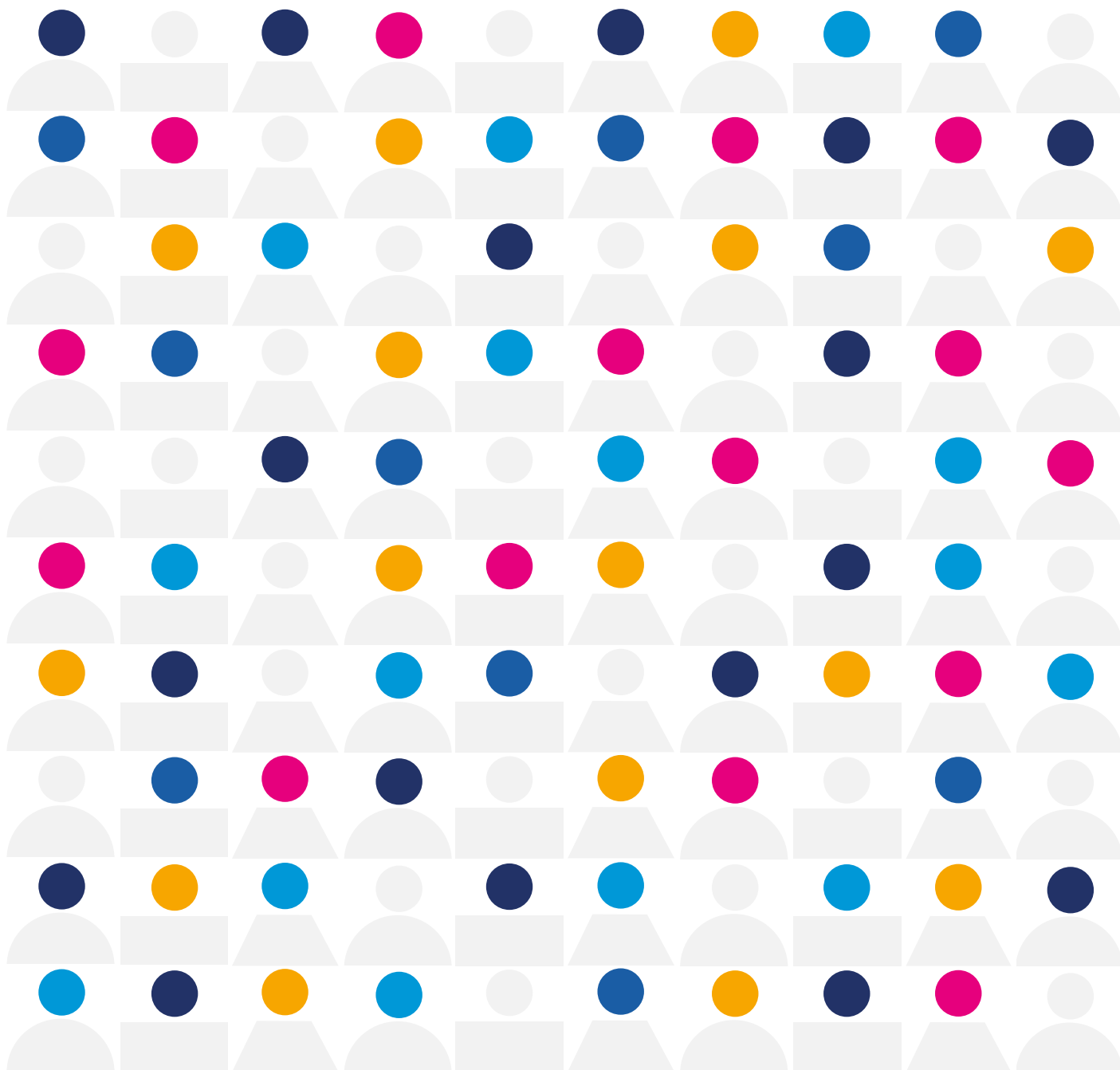
```
# archivo CSV
personas_censo2024 <- arrow::open_dataset("ruta/archivo/personas_censo2024.csv",
format = "csv", delim = ";")

# archivo Parquet
personas_censo2024 <- open_dataset("ruta/archivo/personas_censo2024.parquet")
```

El uso de `arrow` es recomendable porque se integra fácilmente con flujos de trabajo en `dplyr`, permitiendo trabajar con consultas sobre grandes volúmenes de datos.

4.

Estructura de las bases de datos



Este apartado presenta una guía para comprender la organización de la base de microdatos y su relación con las dos versiones del cuestionario disponibles en la página web. Para ello, en primer lugar, se describe la estructura general y los niveles jerárquicos de análisis de la base (vivienda, hogar y persona), junto con las llaves de vinculación que permiten integrar la información entre ellas. Asimismo, se incluye la correspondencia entre las preguntas del cuestionario y las variables de la base de datos, presentando también las variables adicionales creadas para facilitar el análisis, acompañadas de sus criterios de construcción.

Tabla 4.1: Cuestionarios disponibles en la web

Modalidad	Enlace
Viviendas particulares	https://censo2024.ine.gob.cl/wp-content/uploads/2025/03/Cuestionario-Viviendas-Particulares_CPV2024.pdf
Viviendas colectivas y personas en situación de calle	https://censo2024.ine.gob.cl/wp-content/uploads/2025/03/Cuestionario-Viviendas-Colectivas-y-Personas-en-Situacion-de-Calle_CPV2024.pdf

4.1. Organización de la base de datos

Las bases de datos del Censo 2024 fueron diseñadas en correspondencia con la estructura del cuestionario censal. Por lo tanto, antes de iniciar un análisis, se recomienda revisar el cuestionario con el fin de conocer las preguntas, las categorías de respuesta y su secuencia lógica.

La información se organiza en tres niveles de análisis: vivienda, hogar y persona. Cada nivel contiene variables específicas de la unidad que describe, las cuales pueden vincularse mediante identificadores comunes. Esta estructura permite realizar análisis de forma independiente en cada nivel o de manera integrada, según el objetivo de análisis.

4.1.1. Estructura general

La información censal se entrega en tres tablas relacionales, correspondientes a las unidades de análisis de vivienda, hogar y persona. En cada tabla están presentes registros u observaciones de los tres operativos: viviendas particulares, viviendas colectivas y personas en situación de calle¹.

¹ Con el objetivo de facilitar la conversión hacia otros formatos de publicación (ej. Redatam) en las tablas de viviendas y hogares se crea una observación para las entrevistas de viviendas colectivas y personas en situación de calle. No obstante, no hay información en las preguntas de vivienda y hogar para los registros de estos operativos, solo en las preguntas de personas.

La tabla de **vivienda** contiene un registro por cada vivienda censada en el caso de viviendas particulares y colectivas. Para el caso de personas en situación de calle, contiene un registro por cada entrevista realizada. No obstante, estos últimos no deben ser contabilizados como una unidad de vivienda.

- **Identificación:** se utiliza la variable `id_vivienda`, que corresponde al identificador de cada vivienda censada o de cada entrevista.
- **Contenido:** para el operativo de viviendas particulares, incluye variables que describen el estado de ocupación y las características de la vivienda particular. Para el caso de los operativos de viviendas colectivas y personas en situación de calle, la tabla de vivienda contiene la información geográfica asociada a cada entrevista y el tipo de operativo. Para las preguntas de vivienda particular estos operativos tienen valor *missing* (NA) dado que no aplican las preguntas de vivienda particular.

La tabla de **hogar** contiene un registro por cada hogar censado, pero contiene uno o más registros por cada vivienda particular censada.

- **Identificación:** se construye a partir de la combinación entre `id_vivienda` e `id_hogar`, lo que permite identificar a cada hogar dentro de una misma vivienda particular.
- **Contenido:** incluye variables que caracterizan a los hogares.

Los operativos de viviendas colectivas y personas en situación de calle también tienen una observación en esta tabla, pero como no se identifican hogares en estos operativos, el identificador `id_hogar` tiene valor cero (`id_hogar = 0`). Para las preguntas de características de los hogares los registros de estos operativos tienen valor *missing* (NA)

La tabla de **personas** contiene un registro por cada persona censada.

- **Identificación:** se construye a partir de la combinación de `id_vivienda`, `id_hogar` e `id_persona`, que permite identificar de forma única a cada persona.
- **Contenido:** agrupa todas las variables relacionadas con las características individuales de la población, como edad, sexo, educación, migración, ocupación, entre otras.

Dependiendo del operativo, cada registro de esta tabla representa:

- Una persona censada en cada hogar o vivienda en el caso de viviendas particulares.
- Una persona censada en cada vivienda colectiva.
- Una persona censada en el operativo de personas en situación de calle.

En los siguientes ejemplos podrá ver un extracto de cómo se ve la tabla de vivienda, hogar y personas para los tres operativos en un mismo grupo de entrevistas (`id_vivienda` = 123, 124 y 125). Considere que los valores para la variable (`tipo_operativo`) corresponden a:

- 1 = Personas en situación de calle
- 2 = Viviendas particulares
- 3 = Viviendas colectivas

Tabla 4.2: Ejemplo tabla de vivienda para los tres operativos

<code>id_vivienda</code>	<code>tipo_operativo</code>	<code>p4a_mat_paredes</code>
123	1	NA
124	2	1
125	3	NA

Tabla 4.3: Ejemplo tabla de hogar para los tres operativos

<code>id_vivienda</code>	<code>id_hogar</code>	<code>tipo_operativo</code>	<code>p12_tenencia_viv</code>
123	0	1	NA
124	1	2	3
125	0	3	NA

Tabla 4.4: Ejemplo tabla de personas para los tres operativos

<code>id_vivienda</code>	<code>id_hogar</code>	<code>id_persona</code>	<code>tipo_operativo</code>	<code>p24_lug_resid5</code>
123	0	1	1	3
124	1	1	2	2
124	1	2	2	-99
125	0	1	3	2
125	0	2	3	2

4.1.2. Llaves de vinculación

Para integrar la información de las tres tablas se deben considerar las siguientes variables que operan como llaves de vinculación:

- `id_vivienda`: Identificador único de la vivienda censada o identificador único de la entrevista en el caso de personas en situación de calle.
- `id_hogar`: Identificador de cada hogar dentro de la vivienda particular.
- `id_persona`: Identificador de cada persona.

El uso de estas llaves asegura la consistencia en la vinculación y permite realizar cruces de información entre viviendas, hogares y personas.

4.1.3. Ejemplo de vinculación en R

Para unir dos tablas en R, se recomienda utilizar el paquete `dplyr`. A continuación, se presenta un ejemplo que vincula las tablas de viviendas, hogares y personas para conocer cuántas personas de 15 años o más de la comuna de Santiago forman parte de hogares extensos, desglosado por sexo.

```
library(arrow)
library(tidyverse)

personas_censo2024 <- open_dataset("ruta/archivo/personas_censo2024.parquet")
hogares_censo2024 <- open_dataset("ruta/archivo/hogares_censo2024.parquet")
viviendas_censo2024 <- open_dataset("ruta/archivo/viviendas_censo2024.parquet")

# Vincular información de vivienda a nivel de persona
personas_por_sexo <- viviendas_censo2024 %>% select(id_vivienda, comuna) %>%
  filter(comuna == 13101) %>%
  left_join(hogares_censo2024 %>% select(id_vivienda, id_hogar, tipologia_hogar),
    by = c("id_vivienda")) %>%
  filter(tipologia_hogar == 5) %>%
  left_join(personas_censo2024 %>% select(id_vivienda, id_hogar, sexo, edad),
    by = c("id_vivienda", "id_hogar")) %>%
  filter(edad >= 15) %>%
  count(sexo) %>%
  collect()
```

4.1.4. Identificación geográfica

En términos de desagregación geográfica la base de microdatos contiene información a nivel de 1) División Política- Administrativa (en adelante DPA) y 2) la clasificación de las áreas como áreas urbanas o rurales.

La DPA corresponde a la organización territorial del país en unidades administrativas mayores y menores. Chile se estructura en 16 regiones, cada una de ellas subdividida en provincias (56 en total), y estas a su vez en comunas (346 en total).

Por otro lado, el área urbana o rural corresponde a la clasificación del territorio en entidades urbanas y rurales, de acuerdo con los criterios censales establecidos.

- **Entidad urbana:** se entiende como entidad urbana un asentamiento humano con continuidad² y concentración³ de construcciones en un amanzanamiento regular⁴ con población mayor a 2.000 habitantes, o entre 1.001 y 2.000 habitantes donde menos del 50% de la población que declara haber trabajado se dedica a actividades primarias⁵.
- **Entidad rural:** se entiende como entidad rural un asentamiento humano con población menor o igual a 1.000 habitantes, o entre 1.001 y 2.000 habitantes donde más del 50% de la población que declara haber trabajado se dedica a actividades primarias. Además, se define como rural a un conjunto menor de entidades que reúnen los criterios de población para ser definidas como urbanas, pero no los requisitos de amanzanamiento, continuidad o concentración de construcciones.

Tabla 4.5: Entidades territoriales de la base de microdatos

Entidad territorial	Variable	Observaciones
Región	region	Considera los códigos de cada región, desde la 1 a la 16.
Provincia	provincia	Considera el código de la región, concatenado con el número de provincia dentro de la región (RRP)
Comuna	comuna	Considera el código de la región, de la provincia concatenado con el número de la comuna (RRPCC).
Área urbana / rural	area	Considera el código 1 para el área urbana y 2 para el rural.

2 Continuidad: se relaciona con los dos anteriores, ya que la concentración debe presentarse de manera continua para delimitar a una entidad. En el área rural, suele detectarse concentración en torno a ejes viales estructurantes y tiende a dispersarse en los ejes viales menores.

3 Concentración: este principio se refiere a que la población y las edificaciones se encuentren concentradas. En el límite del área urbana las construcciones tienden a estar menos concentradas que las que están en el interior del urbano. Por consiguiente, en el área rural la dispersión se hace mayor, presentando características particulares que permiten diferenciar, a su vez, entre asentamientos rurales dispersos y concentrados.

4 Amanzanamiento: se caracteriza por tener una estructura conformada por manzanas, la que puede presentarse tanto en el área urbana como en el área rural, en asentamientos con mayor concentración de construcciones. Por su parte, la manzana es la unidad geográfica básica con fines estadísticos que conforman zonas censales en áreas urbanas y aldeas en el ámbito rural, contiene un grupo de viviendas contiguas o separadas, edificios, establecimientos y/o predios, delimitados por rasgos geográficos, culturales y naturales.

5 Esto se define a partir de la codificación de la pregunta 43 del cuestionario censal, asociada a la rama de actividad económica a la que se dedica la persona que declaró estar ocupada. Las actividades primarias son aquellas asociadas a la agricultura, ganadería, pesca, silvicultura y minería. Las secundarias, por otro lado, se relacionan con la industria.

4.2. Correspondencia entre preguntas del cuestionario y variables en la base de datos

A continuación, se presenta la correspondencia entre la versión del cuestionario de viviendas particulares con la base de microdatos, distinguiendo por tipo de unidad de observación (vivienda, hogar y persona).

4.2.1. Tabla de vivienda

Tabla 4.6: Correspondencia entre preguntas de cuestionario y variables de base viviendas

Pregunta cuestionario censal	Nombre variable
2. Tipo de vivienda particular	p2_tipo_vivienda
3a. Estado de ocupación de la vivienda	p3a_estado_ocupacion
3b. Estado de ocupación de la vivienda	p3b_estado_ocupacion
4a. ¿Cuál es el material de construcción en las paredes exteriores?	p4a_mat_paredes
4b. ¿Cuál es el material de construcción en la cubierta del techo?	p4b_mat_techo
4c. ¿Cuál es el material de construcción en el piso?	p4c_mat_piso
5. ¿Cuántas piezas de esta vivienda se usan exclusivamente como dormitorio?	p5_num_dormitorios
6. El agua que usa esta vivienda proviene principalmente de:	p6_fuente_agua
7. ¿Cuál es el sistema de distribución del agua en esta vivienda?	p7_distrib_agua
8. El servicio higiénico (WC) principal de esta vivienda es o está:	p8_serv_hig
9. La electricidad de esta vivienda proviene principalmente de:	p9_fuente_elect
10. ¿Cuál es el principal medio de eliminación de basura de esta vivienda?	p10_basura
11. ¿Cuántas personas residen habitualmente en esta vivienda?	p11a_num_personas
11.1. De las personas que residen habitualmente en esta vivienda, ¿todas comparten los gastos para alimentación?	p11b_comparte_gasto
11.2. ¿Cuántos grupos tienen gastos separados de alimentación?	p11c_num_hogar

4.2.2. Tabla de hogares

Tabla 4.7: Correspondencia entre preguntas de cuestionario y variables de base hogares

Pregunta cuestionario censal	Nombre variable
12. La vivienda que ocupa este hogar es:	p12_tenencia_viv
13. ¿Cuál es la principal fuente de energía o combustible que utilizan para cocinar?	p13_comb_cocina
14. ¿Cuál es la principal fuente de energía o combustible que utilizan para calefaccionar?	p14_comb_calefaccion
15. Este hogar, ¿dispone de teléfono móvil, celular o smartphone?	p15a_serv_tel_movil
15. Este hogar, ¿dispone de computador (escritorio o portátil)?	p15b_serv_compu
15. Este hogar, ¿dispone de tablet?	p15c_serv_tablet
15. Este hogar, ¿dispone de internet fija?	p15d_serv_internet_fija
15. Este hogar, ¿dispone de internet móvil desde un celular, tablet o BAM?	p15e_serv_internet_movil
15. Este hogar, ¿dispone de internet por conexión satelital?	p15f_serv_internet_satelital

4.2.3. Tabla de personas

Tabla 4.8: Correspondencia entre preguntas de cuestionario y variables de base personas

Pregunta cuestionario censal	Nombre variable
19. ¿Qué relación de parentesco tiene con la jefa o jefe de hogar?	parentesco
20. ¿Cuál es su sexo?	sexo
21. ¿Cuántos años cumplidos tiene?	edad
22. ¿En qué fecha nació?	
23. ¿Cuál es su estado conyugal o civil actual?	p23_est_civil
24. ¿En qué comuna o país vivía en abril de 2019?	p24_lug_resid5
	p24_lug_resid5_esp
25. Cuándo nació, ¿en qué comuna o país vivía su madre?	p25_lug_nacimiento
	p25_lug_nacimiento_esp
	p25_lug_nacimiento_rec

Pregunta cuestionario censal	Nombre variable
26. ¿En qué año llegó a vivir a Chile?	p26_llegada_periodo
27. ¿Cuál es su nacionalidad?	p27_nacionalidad p27_pais_nacionalidad_esp p27_nacionalidad_rec
28. ¿Es o se considera perteneciente a algún pueblo indígena u originario?	p28_autoid_pueblo
28. ¿A cuál?	p28_pueblo_pert
29. De acuerdo con sus antepasados, tradiciones y cultura, es o se considera:	p29_afrodescendencia p29_afrodescendencia_rec
30. ¿Habla o entiende una de las siguientes lenguas indígenas u originarias?	p30_leng_indigena p30_lengua_indigena_rec
31. ¿Cuál es su religión o credo?	p31_religion p31_religion_rec
32. ¿Tiene dificultad para ver, aun usando anteojos o lentes?	p32a_dificultad_ver
32. ¿Tiene dificultad para oír, aun usando un dispositivo auditivo	p32b_dificultad_oir
32. ¿Tiene dificultad para caminar o subir escaleras?	p32c_dificultad_mover
32. ¿Tiene dificultad para recordar o concentrarse?	p32d_dificultad_cogni
32. ¿Tiene dificultad para realizar tareas de cuidado personal, como bañarse o vestirse?	p32e_dificultad_cuidado
32. ¿Tiene dificultad para hablar o comunicarse usando su idioma habitual, por ejemplo, entender o ser entendido por otras personas?	p32f_dificultad_comunic
33. ¿Asiste actualmente a la educación formal? Incluye educación parvularia, especial, básica, media y superior	p33_edu_asiste
34. ¿Cuál es su nivel educativo más alto alcanzado?	
35. ¿Dentro del nivel declarado anteriormente, ¿cuál es el último curso o año que se aprobó?	cine11
36. ¿Terminó el nivel declarado anteriormente?	
34. ¿Cuál es su nivel educativo más alto alcanzado?	
35. ¿Dentro del nivel declarado anteriormente, ¿cuál es el último curso o año que se aprobó?	escolaridad

Pregunta cuestionario censal	Nombre variable
37. ¿Sabe leer y escribir?	p37_alfabet
38. La semana pasada:	sit_fuerza_trabajo
39. Independiente de lo anterior, la semana pasada:	
40. ¿Trabajó cómo...?	p40_cise_rec
41. ¿Cuál es el oficio, labor u ocupación principal que realizó la semana pasada?	cod_ciuo
42. ¿Qué tareas realizó en ese trabajo?	
43. En ese trabajo, ¿a qué se dedica la empresa, negocio, institución o actividad por cuenta propia?	cod_caenes
44. ¿En qué comuna o país se ubica su trabajo?	p44_lug_trab
	p44_lug_trab_esp
45. ¿Cuál es el medio de transporte principal que utiliza para dirigirse a su lugar de trabajo?	p45_medio_transporte
46. ¿Cuántas hijas e hijos nacidos vivos ha tenido en total?	p46a_tot_hijs_nac
46. Del total de hijas e hijos nacidos vivos, ¿cuántas son mujeres?	p46b_hijas_nac
46. Del total de hijas e hijos nacidos vivos, ¿cuántos son hombres?	p46c_hijos_nac
47. ¿Cuántas hijas e hijos están vivos actualmente?	p47a_tot_hijs_sobrev
47. Del total de hijas e hijos que se encuentran vivos, ¿cuántas son mujeres?	p47b_hijas_sobrev
47. Del total de hijas e hijos que se encuentran vivos, ¿cuántos son hombres?	p47c_hijos_sobrev
48. ¿En qué mes y año nació su última hija o hijo nacido vivo? Año	p48_anio_nac_uh
48. ¿En qué mes y año nació su última hija o hijo nacido vivo? Mes	p48_mes_nac_uh
49. ¿Con cuál género se identifica?	div_genero
50. ¿Se identifica como trans?	

4.3. Variables calculadas

La base de datos del Censo 2024 incluye un conjunto de variables adicionales creadas para facilitar el análisis de las y los usuarios. Estas variables fueron construidas a partir de una o más preguntas del cuestionario.

A continuación, se presentan los criterios para el cálculo de estas variables.

4.3.1. Cantidad de personas censadas (`cant_per`)

La variable `cant_per` tiene por objetivo contabilizar la cantidad de personas censadas en cada vivienda colectiva, particular o por cada entrevista del operativo de personas en situación de calle.

Esta variable se encuentra en la tabla de viviendas y tiene un rango que va desde 0 hasta 5.806 personas censadas. Las viviendas desocupadas o donde no fue posible realizar la entrevista tienen valor cero para esta variable. Los valores más altos de `cant_per` los alcanzan las viviendas colectivas. Para el operativo de personas en situación de calle esta variable siempre tiene valor 1.

La variable `cant_per` no debe confundirse con la variable `p11a_num_persona` que contabiliza a las personas que residen en la vivienda particular de acuerdo con la declaración de las y los informantes.

4.3.2. Cantidad de hogares censados (`cant_hog`)

La variable `cant_hog` tiene por objetivo contabilizar la cantidad de hogares censados en cada vivienda particular.

Esta variable se encuentra en la tabla de viviendas y tiene un rango que va desde 0 hasta 26 hogares. Las viviendas con cero hogares corresponden a viviendas particulares desocupadas o donde no fue posible realizar la entrevista. Para los operativos de viviendas colectivas y personas en situación de calle esta variable tiene valor *missing* (NA).

La variable `cant_hog` no debe confundirse con la variable `p11c_num_hog` que contabiliza a los hogares que residen en la vivienda de acuerdo con la declaración de las y los informantes.

4.3.3. Identificador de comunas con menos de 6.000 personas censadas (`comuna_bajo_umbral`)

La variable `comuna_bajo_umbral` tiene por objetivo identificar a las comunas que tienen menos de 6.000 personas censadas. El umbral de 6.000 personas es utilizado para definir distintos tratamientos de anonimización. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

Para conocer las etiquetas de esta variable, ver **Diccionario de variables**.

4.3.4. Edad quinquenal (`edad_quinquenal1`)

La variable `edad_quinquenal1` es una recodificación de la variable `edad`⁶ en tramos de 5 años desde los 0 hasta los 84 años, y un tramo abierto que considera “85 años o más”.

Las categorías de esta variable representan el límite inferior de la edad quinquenal, por ejemplo, para el tramo entre 0-4 años de edad, el valor se representa con un 0.

Para conocer las etiquetas, ver **Diccionario de variables**.

4.3.5. Lugar específico de residencia hace cinco años (`p24_lug_resid5_esp`)

La variable `p24_lug_resid5_esp`⁷ contiene la información específica del código de la comuna o el país de residencia de las personas hace cinco años. Se construye a partir de la pregunta `p24_lug_resid5`, que tiene cinco categorías:

1. Aún no nacía
2. En esta comuna
3. En otra comuna
4. En otro país
- 99. No responde

Cuando la persona declara que hace cinco años residía en esta misma comuna (categoría 2), `p24_lug_resid5_esp` toma el código de la comuna de residencia habitual de la persona. Si declara “En otra comuna” (categoría 3), se asigna el código de la comuna de residencia de hace cinco años declarado. Cuando no se dispone de comuna, pero sí de la región de residencia hace cinco años, se asigna el código de región correspondiente. Si la declaración es “En otro país” (categoría 4) y el país declarado es Venezuela, Perú, Haití, Colombia, Bolivia o Argentina, se asigna el código del país según el estándar UN-M49⁸. Sin embargo, si se responde un país distinto a los mencionados, por motivos de control de divulgación se utiliza el código del continente al cual pertenece según el estándar UN-M49. En el caso de las personas que responden “Aún no nacía” (categoría 1) en `p24_lug_resid5_esp`, se asigna NA.

Para conocer el código de cada entidad territorial, ver Diccionario de variables territoriales disponible en: <https://censo2024.ine.gob.cl/resultados/>

⁶ La variable `edad` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. La construcción de la variable `edad_quinquenal` se realizó con la variable `edad` que maneja internamente la institución. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

⁷ La variable `p24_lug_resid5_esp` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para el resto de las comunas a partir de la variable `p24_lug_resid5` se puede identificar si la persona residía hace cinco años antes del momento del Censo en Chile o en el extranjero. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

⁸ Listado de códigos normalizados de países para uso estadístico publicado por la División de Estadística de Naciones Unidas.

4.3.6. Lugar específico de nacimiento (p25_lug_nacimiento_esp)

La variable `p25_lug_nacimiento_esp`⁹ contiene la información específica del código de la comuna o el país de las personas al momento de su nacimiento. Se construye a partir de la pregunta `p25_lug_nacimiento`, que tiene cuatro categorías:

1. En esta comuna
2. En otra comuna
3. En otro país
- 99. No responde

Cuando la persona declara que nació en esta misma comuna (categoría 1), `p25_lug_nacimiento_esp` toma el código de la comuna de residencia habitual de la persona. Si declara “En otra comuna” (categoría 2), se asigna el código de la comuna de nacimiento declarado; cuando no se dispone de comuna, pero sí de la región, se asigna el código de la región correspondiente. Si la declaración es “En otro país” (categoría 3) y el país declarado es Venezuela, Perú, Haití, Colombia, Bolivia o Argentina, se asigna el código del país según el estándar UN-M49. Sin embargo, si se responde otro país, por motivos de control de divulgación se asigna el código del continente al cual pertenece según el estándar UN-M49.

Para conocer el código de cada entidad territorial, ver el Diccionario de variables territoriales disponible en: <https://censo2024.ine.gob.cl/resultados/>

4.3.7. Nacionalidad

La base de datos de personas contiene tres variables referidas al país de nacionalidad de las personas. La primera es la variable `p27_nacionalidad` que tiene tres categorías¹⁰:

1. Chilena (exclusivamente)
2. Chilena y otra (doble nacionalidad)
3. Otra nacionalidad
- 99. No responde

La segunda identifica si las personas tienen nacionalidad chilena o no (`p27_nacionalidad_rec`). Se construye a partir de la pregunta 27 del cuestionario censal, agrupando sus categorías en: “Sí” para las personas que tengan nacionalidad chilena (agrupación de categorías “1. Chilena (exclusivamente)” y “2. Chilena y otra (doble nacionalidad)” de la pregunta “No” para las personas que no tienen nacionalidad chilena (categoría “3. Otra nacionalidad” de la pregunta 27).

⁹ La variable `p25_lug_nacimiento_esp` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para el resto de las comunas está disponible la variable `p25_lug_nacimiento_rec` que especifica si la persona nació en Chile o en el extranjero. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

¹⁰ La variable `p27_nacionalidad` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para el resto de las comunas está disponible la variable `p27_nacionalidad_rec` que especifica si la persona tiene nacionalidad chilena o no. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

La tercera variable, `p27_nacionalidad_esp`¹¹, especifica el país de la otra nacionalidad. Si el país declarado es Venezuela, Perú, Haití, Colombia, Bolivia o Argentina, se asigna el código de país según el estándar UN-M49. Para el resto de los países por motivos de control de divulgación se asigna el código del continente al que pertenecen según UN-M49.

Para conocer el código de cada entidad territorial, ver el Diccionario de variables territoriales disponible en: <https://censo2024.ine.gob.cl/resultados/>

4.3.8. Afrodescendencia recodificada (`p29_afrodescendencia_rec`)

La variable `p29_afrodescendencia_rec` permite identificar si la persona es o se considera afrodescendiente. Se compone de tres categorías:

1. Sí
2. No
- 99. No responde

La variable se construye a partir de la pregunta 29 del cuestionario censal.

Tabla 4.9: Afrodescendencia recodificada

Categoría	Etiqueta categoría	Afrodescendencia
1	Afrodescendiente	1
2	Afrochileno/a	1
3	Negro/a	1
4	Del Pueblo Tribal Afrodescendiente Chileno	1
5	Moreno/a de Azapa	1
6	Moreno/a de la Chimba	1
7	Ninguna de las anteriores	2
-99	No responde	-99

¹¹ La variable `p27_nacionalidad_esp` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para el resto de las comunas está disponible la variable `p27_nacionalidad_rec` que especifica si la persona tiene nacionalidad chilena o no. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

4.3.9. Manejo de lengua indígena u originaria (p30_lengua_indigena_rec)

La variable `p30_lengua_indigena_rec` permite identificar si la persona habla o entiende alguna lengua indígena u originaria. Se compone por tres categorías:

1. Sí
2. No
- 99. No respuesta

La variable se construye a partir de la pregunta 30 del cuestionario censal.

Tabla 4.10: Manejo de lengua indígena u originaria

Categoría	Etiqueta categoría	Categoría
1	Mapuzugun (lengua mapuche)	1
2	Aymara	1
3	Quechua	1
4	Rapa Nui	1
5	Ckunza	1
6	Kawéswar	1
7	Yagán	1
8	Otra lengua indígena de Chile	1
9	No habla ni entiende ninguna lengua indígena de Chile	2
-99	No responde	-99

4.3.10. Religión o credo recodificada (p31_religion_rec)

La variable `p31_religion_rec` permite identificar si la persona profesa alguna religión o credo. Se compone por tres categorías:

1. Sí
2. No
- 99. No responde

La variable se construye a partir de la pregunta 31 del cuestionario censal.

Tabla 4.11: Religión o credo recodificada

Categoría	Etiqueta categoría	Categoría
1	Católica	1
2	Evangélica o protestante	1
3	Judía	1
4	Musulmana	1
5	Iglesia de Jesucristo de los Santos de los Últimos Días	1
6	Católica Ortodoxa	1
7	Budista	1
8	Hinduista	1
9	Fe Bahá'í	1
10	Testigo de Jehová	1
11	Otra religión o credo	1
12	Ninguna	2
-99	No responde	-99

4.3.11. Años de escolaridad (escolaridad)

Para la construcción de los años de estudio de cada persona se utilizan las siguientes preguntas:

34. ¿Cuál es su nivel educativo más alto alcanzado?

35. Dentro del nivel declarado anteriormente, ¿cuál es el último curso o año que aprobó?

La pregunta 34 toma valores entre 1 a 16, considerando los distintos niveles educacionales. La pregunta 35 toma valores entre 1 a 8, identificando el curso del nivel declarado.

El cálculo de los años de escolaridad sigue una lógica en la que se acumulan los años de estudio aprobados, comenzando desde los niveles más básicos y sumando los años correspondientes a los niveles superiores según el último curso aprobado.

Por ejemplo, si una persona declara que su nivel educativo es “9. Media científico humanista o artística” y dentro del nivel declarado el último curso aprobado fue 3ro medio, el cálculo de la escolaridad considera los 8 años de enseñanza básica y los 3 años de educación media aprobados (11 años de escolaridad).

La construcción de la variable considera lo siguiente:

Tabla 4.12: Construcción de variable años de escolaridad

Categoría	Nivel educativo	Años de escolaridad
1	Nunca asistió	0
2	Sala cuna	0
3	Jardín infantil	0
4	Prekínder	0
5	Kinder	0
6	Educación especial o diferencial	0
7	Educación básica	pregunta 35 + 0
8	Primaria (sistema antiguo)	pregunta 35 + 0
9	Media científico-humanista o artística	pregunta 35 + 8
10	Media técnico-profesional	pregunta 35 + 8
11	Humanidades (sistema antiguo)	pregunta 35 + 6
12	Técnico comercial, industrial, normalista (sistema antiguo)	pregunta 35 + 6
13	Técnico nivel superior (1 a 3 años) (incluye suboficial FFAA)	pregunta 35 + 12
14	Profesional (4 años o más) (incluye oficial FFAA)	pregunta 35 + 12
15	Magíster	pregunta 35 + 17
16	Doctorado	pregunta 35 + 19
-99	Sin respuesta en nivel educativo	-99

Para los niveles 1 a 6 de la pregunta 35 no se contabilizan años de escolaridad, por lo que esta variable toma valor cero. Para el resto de los niveles educativos, el cálculo incorpora los años de enseñanza básica, media y superior, sumados al curso aprobado dentro del nivel declarado, según corresponda. Esta lógica busca mantener la coherencia para los años de escolaridad, así como la comparabilidad entre personas que cursaron distintos sistemas educativos.

4.3.12. CINE 11 (cine11)

La variable `cine11` homologa los niveles educativos del Censo 2024 al Clasificador Internacional Normalizado de la Educación (CINE 11) desarrollado por UNESCO. El propósito de este estándar es establecer equivalencias entre los niveles educativos nacionales, permitiendo comparabilidad internacional y coherencia metodológica en la medición del logro educativo entre países.

El logro educativo corresponde al nivel más alto de educación que una persona ha completado exitosamente, es decir, aquel en el cual ha cumplido con todos los requisitos académicos formales establecidos por el programa o institución educativa, incluyendo la certificación correspondiente si aplica (UNESCO, 2012). El CINE 2011 organiza la educación en nueve niveles jerárquicos, desde el nivel 0 (educación preprimaria) hasta el nivel 8 (doctorado o equivalente). Emplea un sistema de codificación a tres dígitos que combina el nivel educativo alcanzado, la orientación del programa (general, vocacional, académica o profesional) y la completitud del nivel.

La construcción de la variable considera el nivel educativo, el curso y el término de nivel de la persona como se detalla en la siguiente tabla.

Tabla 4.13: Homologación a CINE 11 en base a combinaciones de nivel más alto alcanzado, último curso aprobado y completitud de nivel

Nivel educativo	Curso	Término de nivel	Código CINE 11	Etiqueta
Nunca asistió	-	-	01	Nunca cursó un programa educativo
Sala cuna	-	-	02	Educación de la primera infancia (incluye la forma parcial)
Jardín infantil	-	-	02	Educación de la primera infancia (incluye la forma parcial)
Prekinder	-	-	02	Educación de la primera infancia (incluye la forma parcial)
Kínder	-	-	02	Educación de la primera infancia (incluye la forma parcial)
Educación especial o diferencial	-	-	98	Educación especial o diferencial
Básica	1 a 5	No	03	Primaria en forma parcial (sin conclusión del nivel)
Básica	6 a 7	No	10	Educación primaria (nivel 1)

Nivel educativo	Curso	Término de nivel	Código CINE 11	Etiqueta
Básica	8	Sí	14	Educación primaria (nivel 2), con orientación general
Primaria (sistema antiguo)	1 a 5; -99	No	03	Primaria en forma parcial (sin conclusión del nivel)
Primaria (sistema antiguo)	6	Sí	10	Educación primaria (nivel 1)
Media científico humanista o artística	1 a 3; -99	No	14	Educación primaria (nivel 2), con orientación general
Media científico humanista o artística	4	Sí	24	Educación secundaria, con orientación general
Media técnico profesional	1 a 4; -99	No	14	Educación primaria (nivel 2), con orientación general
Media técnico profesional	4 a 5; -99	Sí	25	Educación secundaria, con orientación vocacional
Humanidades (sistema antiguo)	1	No	10	Educación primaria (nivel 1)
Humanidades (sistema antiguo)	2 a 5	No	14	Educación primaria (nivel 2), con orientación general
Humanidades (sistema antiguo)	6	Sí	24	Educación secundaria, con orientación general
Técnico comercial, industrial, normalista (sistema antiguo)	1	No	10	Educación primaria (nivel 1)
Técnico comercial, industrial, normalista (sistema antiguo)	2 a 5	No	14	Educación primaria (nivel 2), con orientación general
Técnico comercial, industrial, normalista (sistema antiguo)	6; -99	Sí	25	Educación secundaria, con orientación vocacional
Técnico nivel superior (1 a 3 años) (incluye suboficial FFAA)	1 a 3; -99	No	25	Educación secundaria, con orientación vocacional
Técnico nivel superior (1 a 3 años) (incluye suboficial FFAA)	1 a 3; -99	Sí	35	Educación terciaria de ciclo corto, con orientación vocacional

Nivel educativo	Curso	Término de nivel	Código CINE 11	Etiqueta
Profesional (4 años o más) (incluye oficial FFAA)	1 a 6; -99	No	24	Educación secundaria, con orientación general
Profesional (4 años o más) (incluye oficial FFAA)	4 a 7; -99	Sí	46	Grado de educación terciaria o nivel equivalente, con orientación no especificada
Magíster	1 a 2; -99	No	46	Grado de educación terciaria o nivel equivalente, con orientación no especificada
Magíster	1 a 3; -99	Sí	56	Nivel de maestría, especialización o equivalente, con orientación no especificada
Doctorado	1 a 5; -99	No	56	Nivel de maestría, especialización o equivalente, con orientación no especificada
Doctorado	3 a 5; -99	Sí	64	Nivel de doctorado o equivalente, con orientación académica
Sin respuesta	-	-	-99	No sabe o no responde en variables relevantes para la clasificación

4.3.13. Asistencia a la educación parvularia (asistencia_parv)

La variable `asistencia_parv`¹² clasifica a las personas entre 0 a 5 años dependiendo de si asisten o no a la educación parvularia. Para construir esta variable se utiliza la información de las variables de edad, asistencia a la educación formal (`p33_edu_asiste`) y la pregunta de nivel educativo (pregunta 34). El principal uso de la variable `asistencia_parv` es permitir el cálculo de la tasa de asistencia neta a la educación parvularia. Por lo tanto, esta variable solo tiene valores para las personas de 0 a 5 años que tienen respuesta válida en las preguntas de asistencia y nivel, en tanto se espera que las personas en estas edades se encuentren cursando este nivel educativo.

Para identificar a las personas que **asisten** a la educación parvularia se utiliza la variable de asistencia a la educación formal (`p33_edu_asiste`) y los niveles asociados a la educación parvularia: “Nunca asistió”¹³, “Sala cuna”, “Jardín infantil” y “Prekínder”¹⁴.

¹² La variable `asistencia_parv` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

¹³ Debido a la diferencia de temporalidades de las preguntas de nivel educativo más alto alcanzado y asistencia a la educación formal, existe una dificultad para determinar a qué nivel educativo asiste actualmente una persona. Por esta razón, se asumió que las personas que responden que se encuentran asistiendo actualmente a la educación formal (`p33_edu_asiste` = 1) y que responden en nivel “1. Nunca asistió” corresponden a aquellas personas que están iniciando su educación formal en el sistema educativo. Para el caso de las personas entre 0 a 5 años, se asume que esto corresponde a la educación parvularia considerando las edades de ingreso a los niveles del sistema educacional chileno.

¹⁴ Debido a la diferencia de temporalidades de las preguntas de nivel educativo más alto alcanzado y asistencia a la educación formal, existe una dificultad para determinar a qué nivel educativo asiste actualmente una persona. Por esta razón, se asumió que las personas que responden que se encuentran asistiendo actualmente a la educación formal (`p33_edu_asiste` = 1) y cuyo último nivel más alto alcanzado es “5. Kinder” se encuentran asistiendo a la educación básica (`asistencia_basica` = 1).

4.3.14. Asistencia a la educación básica (asistencia_basica)

La variable `asistencia_basica`¹⁵ clasifica a las personas entre 6 a 13 años dependiendo de si asisten o no a la educación básica. Para construir esta variable se utiliza la información de las variables de edad, asistencia a la educación formal (`p33_edu_asiste`), último nivel alcanzado (pregunta 34), curso (pregunta 35) y la pregunta de término de nivel (pregunta 36). El principal uso de la variable `asistencia_basica` es permitir el cálculo de la tasa de asistencia neta a la educación básica. Por lo tanto, esta variable solo tiene valores para las personas de 6 a 13 años que tienen respuesta válida en las preguntas de asistencia, nivel y término de nivel, en tanto se espera que las personas en estas edades se encuentren cursando este nivel educativo.

Para identificar a las personas que **asisten** a este nivel, se debe considerar la asistencia de las personas que declaran la siguiente combinación de variables¹⁶.

Tabla 4.14: Combinación entre asistencia, nivel, curso y término de nivel para identificar a las personas que asisten a educación básica

Edad	Asiste a la educación formal	Nivel educativo	Curso	Término de nivel	Asistencia Educación Básica
6 a 13 años	1. Sí	5. Kinder	-	-	1. Sí
6 a 13 años	1. Sí	7. Básica	1ro a 7mo; -99 (no respuesta)	2. No completa	1. Sí

4.3.15. Asistencia a la educación media (asistencia_media)

La variable `asistencia_media`¹⁷ clasifica a las personas entre 14 a 17 años dependiendo de si asisten o no a la educación media. Para construir esta variable se utiliza la información de las variables de edad, asistencia a la educación formal (`p33_edu_asiste`), la pregunta de nivel educativo (pregunta 34), curso (pregunta 35) y la pregunta de término de nivel (pregunta 36). El principal uso de la variable `asistencia_media` es permitir el cálculo de la tasa de asistencia neta a la educación media. Por lo tanto, esta variable solo tiene valores para las personas de 14 a 17 años que tienen respuesta válida en las preguntas de asistencia y nivel, en tanto se espera que las personas en estas edades se encuentren cursando este nivel educativo.

Para identificar a las personas que asisten a este nivel, se debe considerar la asistencia de las personas que declaran la siguiente combinación de variables¹⁸.

¹⁵ La variable `asistencia_basica` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

¹⁶ Esta combinación permite identificar si la asistencia reportada corresponde al nivel básico, ya que una persona que reporta que asiste actualmente a la educación formal, señala que su último nivel alcanzado es "6. Básica" y además indica que terminó sus estudios de ese nivel, se asume que está asistiendo al nivel siguiente, es decir, educación media.

¹⁷ La variable `asistencia_media` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

¹⁸ Esta combinación permite identificar si la asistencia reportada corresponde efectivamente al nivel educativo de enseñanza media, ya que una persona que reporta que asiste actualmente a la educación formal, señala que su último nivel alcanzado es "9. Media científico humanista o artística" o "10. Media técnico profesional" y además indica que terminó sus estudios de ese nivel, se asume que está asistiendo al nivel siguiente, es decir, educación superior. Debido a la diferencia de temporalidades de las preguntas de nivel educativo más alto alcanzado y asistencia a la educación formal, existe una dificultad para determinar a qué nivel educativo asiste actualmente una persona. Por esta razón, se tuvo que tomar este supuesto.

Tabla 4.15: Combinación entre asistencia, nivel, curso y término de nivel para identificar a las personas que asisten a la educación media

Edad	Asiste a la educación formal	Nivel educativo	Curso	Término de nivel	Asistencia Educación Media
14 a 17 años	1. Sí	6. Básica	8vo	1. Sí	1. Sí
14 a 17 años	1. Sí	9. Media científico humanista o artística	1ro a 3ro; -99 (no respuesta)	2. No	1. Sí
14 a 17 años	1. Sí	10. Media técnico profesional	1ro a 4to; -99 (no respuesta)	2. No	1. Sí

4.3.16. Asistencia a la educación superior (asistencia_superior)

La variable `asistencia_superior`¹⁹ clasifica a las personas entre 18 a 24 años dependiendo de si asisten o no a la educación superior. Para construir esta variable se utiliza la información de las variables de edad, asistencia a la educación formal (`p33_edu_asiste`), último nivel alcanzado (pregunta 34), curso (pregunta 35) y la pregunta de término de nivel (pregunta 36). El principal uso de la variable `asistencia_superior` es permitir el cálculo de la tasa de asistencia neta a la educación superior. Por lo tanto, esta variable solo tiene valores para las personas de 18 a 24 años que tienen respuesta válida en las preguntas de asistencia y nivel, en tanto se espera que las personas en estas edades se encuentren cursando este nivel educativo.

Para identificar a las personas que asisten a este nivel, se debe considerar la asistencia de las personas que declaran la siguiente combinación de variables²⁰.

Tabla 4.16: Combinación entre asistencia, nivel, curso y término de nivel para identificar a las personas que asisten a educación superior

Edad	Asiste a la educación formal	Nivel educativo	Curso	Término de nivel	Asistencia Educación Superior
18 a 24 años	1. Sí	9. Media científico humanista o artística	4to	1. Sí	1. Sí
18 a 24 años	1. Sí	10. Media técnico profesional	4to-5to; -99 (no respuesta)	1. Sí	1. Sí

¹⁹ La variable `asistencia_superior` en la base publicada se encuentra disponible para las comunas de 6.000 personas censadas o más. Para más detalles se sugiere revisar el capítulo de **Control de divulgación estadística** de este manual.

²⁰ Esta combinación permite identificar si la asistencia reportada corresponde efectivamente al nivel educativo de enseñanza superior, ya que una persona que reporta que asiste actualmente a la educación formal, señala que su último nivel alcanzado tiene valor "13. Técnico nivel superior" o "14. Profesional" y además indica que terminó sus estudios de ese nivel, se asume que está asistiendo al nivel siguiente, es decir, magister o doctorado. Debido a la diferencia de temporalidades de las preguntas de nivel educativo más alto alcanzado y asistencia a la educación formal, existe una dificultad para determinar a qué nivel educativo asiste actualmente una persona. Por esta razón, se tuvo que tomar este supuesto.

Edad	Asiste a la educación formal	Nivel educativo	Curso	Término de nivel	Asistencia Educación Superior
18 a 24 años	1. Sí	13. Técnico nivel superior (1 a 3 años) (incluye suboficial FFAA)	1ro-3ro; -99 (no respuesta)	2. No	1. Sí
18 a 24 años	1. Sí	14. Profesional (4 años o más) (incluye oficial FFAA)	1ro-6to; -99 (no respuesta)	2. No	1. Sí

4.3.17. Situación de la fuerza de trabajo (sit_fuerza_trabajo)

La variable `sit_fuerza_trabajo` corresponde a la clasificación de las personas en edad de trabajar según su respuesta en las preguntas 38 y 39 del cuestionario censal. Las categorías de esta variable son:

- Población ocupada:** personas en edad de trabajar, que durante la semana de referencia dedicaron al menos una hora a alguna actividad para producir bienes o servicios a cambio de una remuneración o beneficios.
 - Población desocupada:** personas en edad de trabajar sin trabajo en la ocupación, que durante la semana de referencia estuvieron disponibles para trabajar y/o en busca de trabajo.
 - Población fuera de la fuerza de trabajo:** personas en edad de trabajar que no tuvieron una ocupación y tampoco buscaron y/o estuvieron disponibles para trabajar. En otras palabras, corresponde a personas de 15 años o más que no quedan clasificadas como ocupadas ni desocupadas durante la semana de referencia.
- 99. No responde:** personas en edad de trabajar que no responden las preguntas 38 y 39 del cuestionario censal.

Las categorías de `sit_fuerza_trabajo` se construyen a partir de las siguientes respuestas entre las preguntas 38 y 39:

Tabla 4.17: Construcción de variable situación en la fuerza de trabajo

Pregunta 38	Pregunta 39	Categoría
1. Trabajó por un pago en dinero o especies	No se habilita	1
2. Trabajó sin pago para un familiar	No se habilita	1
3. Tenía empleo, pero estuvo de vacaciones, con licencia, en descanso laboral, etc	No se habilita	1

Pregunta 38	Pregunta 39	Categoría
4. Se encontraba buscando empleo y disponible para trabajar	1. Realizó alguna actividad de al menos una hora para generar ingresos o 2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia).	1
5. Estaba estudiando	1. Realizó alguna actividad de al menos una hora para generar ingresos o 2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia).	1
6. Es jubilado/a, pensionado/a o rentista	1. Realizó alguna actividad de al menos una hora para generar ingresos o 2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia).	1
7. Realizó quehaceres de su hogar	1. Realizó alguna actividad de al menos una hora para generar ingresos o 2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia).	1
8. Otra situación	1. Realizó alguna actividad de al menos una hora para generar ingresos o 2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia).	1
No responde (-99)	1. Realizó alguna actividad de al menos una hora para generar ingresos o 2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia).	1
4. Se encontraba buscando empleo y disponible para trabajar	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	2
5. Estaba estudiando	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	3
6. Es jubilado/a, pensionado/a o rentista	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	3
7. Realizó quehaceres de su hogar	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	3

Pregunta 38	Pregunta 39	Categoría
8. Otra situación	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	3
No responde (-99)	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	3

4.3.18. Dependencia económica (utilizada en el cálculo del déficit habitacional cuantitativo) (depend_econ_deficit_hab)

La variable `depend_econ_deficit_hab` permite identificar si, de acuerdo con las definiciones metodológicas del Ministerio de Vivienda y Urbanismo (Minvu) para el cálculo del déficit, las personas son económicamente independientes (perciben ingresos) o independientes (no perciben ingresos). Se recomienda usar esta variable solo para el cálculo del déficit habitacional.

Las categorías de esta variable son las siguientes:

1. Económicamente independiente
2. Económicamente dependiente
- 99. No responde

Tabla 4.18: Construcción de variable dependencia

Pregunta 38	Pregunta 39	Categoría
1. Trabajó por un pago en dinero o especies	No se habilita	1
2. Trabajó sin pago para un familiar	No se habilita	1
3. Tenía empleo, pero estuvo de vacaciones, con licencia, en descanso laboral, etc.	No se habilita	1
4. Se encontraba buscando empleo y disponible para trabajar	1. Realizó alguna actividad de al menos una hora para generar ingresos	1
5. Estaba estudiando	1. Realizó alguna actividad de al menos una hora para generar ingresos	1

Pregunta 38	Pregunta 39	Categoría
6. Es jubilado/a, pensionado/a o rentista	1. Realizó alguna actividad de al menos una hora para generar ingresos	1
7. Realizó quehaceres de su hogar	1. Realizó alguna actividad de al menos una hora para generar ingresos	1
8. Otra situación	1. Realizó alguna actividad de al menos una hora para generar ingresos	1
No responde (-99)	1. Realizó alguna actividad de al menos una hora para generar ingresos	1
4. Se encontraba buscando empleo y disponible para trabajar	2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia)	2
5. Estaba estudiando	2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia)	2
6. Es jubilado/a, pensionado/a o rentista	2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia)	1
7. Realizó quehaceres de su hogar	2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia)	2
8. Otra situación	2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia)	2
No responde (-99)	2. Ayudó o realizó algún servicio sin remuneración para la empresa o negocio familiar (de un integrante de su familia)	2
4. Se encontraba buscando empleo y disponible para trabajar	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	2
5. Estaba estudiando	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	2
6. Es jubilado/a, pensionado/a o rentista	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	1

Pregunta 38	Pregunta 39	Categoría
7. Realizó quehaceres de su hogar	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	2
8. Otra situación	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	2
No responde (-99)	3. No realizó ninguna actividad por ingresos, ni siquiera por una hora.	2
No responde (-99)	No responde (-99)	-99

Cabe destacar que las personas de 14 años o menos quedan clasificadas bajo la categoría “2. Económicamente dependiente”.

4.3.19. Situación en el empleo (p40_cise_rec)

La variable p40_cise_rec corresponde a la clasificación de situación en el empleo de las personas ocupadas según la respuesta en la pregunta 40 del cuestionario censal.

Las categorías de la variable p40_cise_rec son:

1. Independiente
2. Dependiente
3. Trabajador/a familiar o personal no remunerado en un negocio de un integrante de la familia
- 99. No responde

Las categorías de p40_cise_rec se construyen a partir de las siguientes respuestas:

Tabla 4.19: Construcción de variable situación en el empleo

Pregunta 40	Categoría
1. Empleador/a patrón/a: con empleados/as contratados/as	1
2. Trabajador/a independiente o por cuenta propia: sin empleados/as contratados/as	1
3. Asalariado/a, empleado/a u obrero/a para un patrón/a	2
4. Trabajador/a de casa particular o servicio doméstico	2

Pregunta 40	Categoría
5. Trabajador/a familiar o personal no remunerado en un negocio de un integrante de la familia	3
-99. No respuesta	-99

4.3.20. Grupo ocupacional - CIUO-08.CL (cod_ciuo)

Corresponde a la codificación de la ocupación principal de la población ocupada. Las respuestas se codificaron utilizando el Clasificador Chileno de Ocupaciones (CIUO-08.cl), correspondiente a una adaptación nacional del Clasificador Internacional Uniforme de Ocupaciones de 2008 (CIUO-08). La codificación se realizó a 1 dígito, mediante una combinación de codificación manual y codificación automática, utilizando modelos de redes neuronales.

Para conocer las etiquetas, ver **Diccionario de variables** de personas.

4.3.21. Actividad económica - CAENES (cod_caenes)

Corresponde a la codificación de actividad económica de la empresa, negocio, institución o actividad por cuenta propia en que trabajan las personas ocupadas. La codificación se realizó utilizando el Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES), correspondiente a una adaptación del Clasificador Chileno de Actividades Económicas CIIU4.CL 2012. La codificación se realizó a 1 dígito, mediante una combinación de codificación manual y codificación automática, utilizando modelos de redes neuronales.

Para conocer las etiquetas, ver **Diccionario de variables** de personas.

4.3.22. Lugar de trabajo (p44_lug_trab_esp)

La variable p44_lug_trab_esp contiene la información específica del código de la comuna o continente del país donde trabajan las personas ocupadas. Se construye a partir de la pregunta p44_lug_trab, que tiene las siguientes categorías:

1. En su comuna, en la vivienda donde reside
2. En su comuna, fuera de la vivienda donde reside
3. En otra comuna
4. En otro país
5. Varias comunas o países
- 99. No responde

La construcción de la variable es la siguiente: cuando la persona declara que trabaja en su comuna, ya sea en la vivienda donde reside (categoría 1) o fuera de la vivienda donde reside (categoría 2), `p44_lug_trab_esp` toma el código de la comuna de residencia habitual. Si declara “En otra comuna” (categoría 3), se asigna el código de la comuna donde trabaja según lo declarado; cuando no se dispone de comuna, pero sí de región, se asigna el código de la región correspondiente. Si la declaración es “En otro país” (categoría 4) por motivos de control de divulgación, se asigna el código del continente al cual pertenece el país, según el estándar UN-M49. En la categoría “Varias comunas o países” (categoría 5) se asigna el valor 999. En los casos de no respuesta se asigna -99 en `p44_lug_trab_esp`.

Para conocer el código de cada entidad territorial, ver Diccionario de variables territoriales disponible en: <https://censo2024.ine.gob.cl/resultados/>

4.3.23. Diversidades de género (`div_genero`)

Corresponde a la identificación de la población de género diverso, es decir, personas trans, no binarias y de otras identidades de género, respecto al total de la población.

Para su construcción se utilizan las variables sexo, género y clasificación trans:

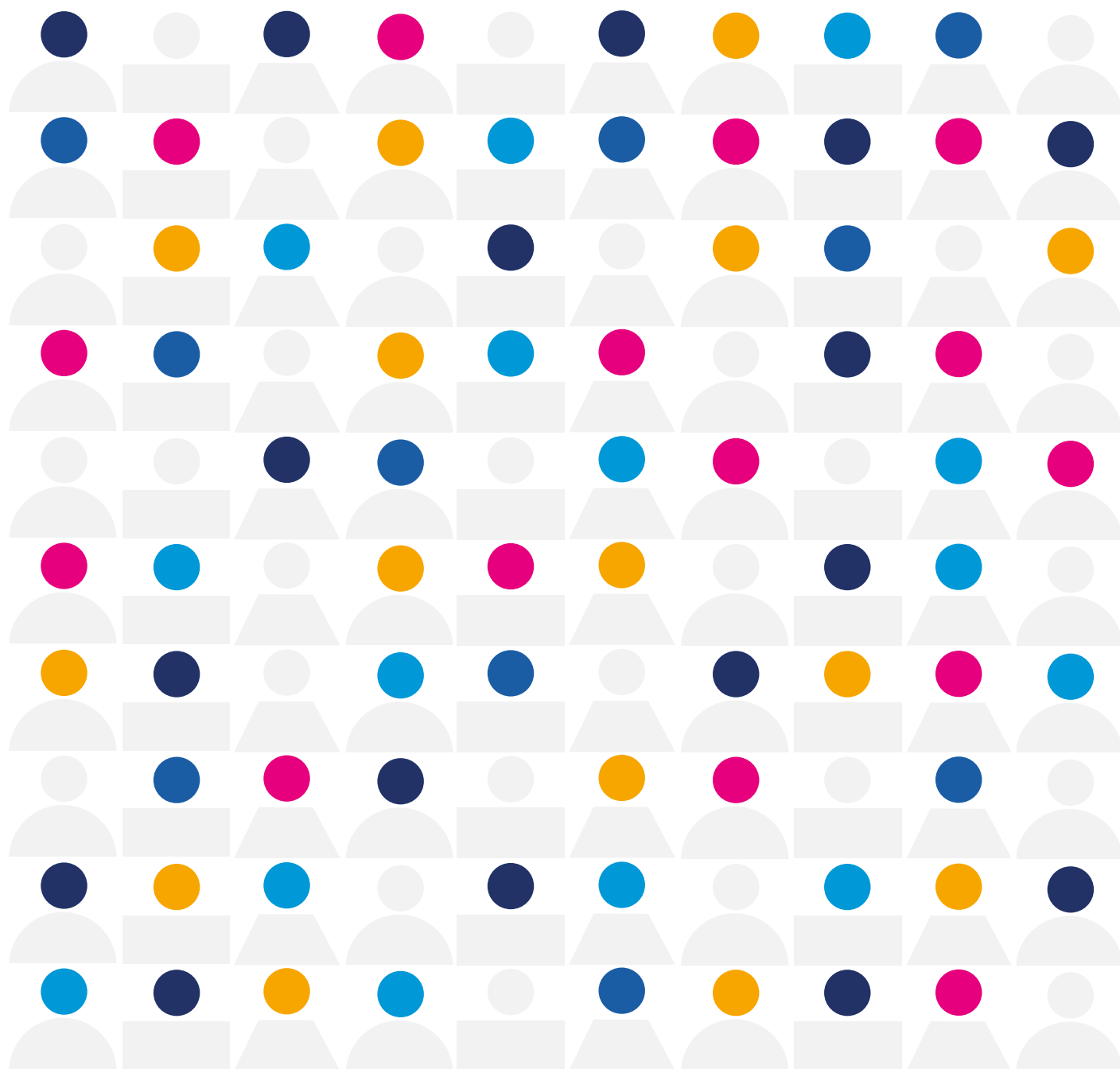
Tabla 4.20: Construcción de variable diversidades de género

Sexo	Género	Identificación trans	Diversidades de género
1. Hombre	2. Femenino	1. Sí	1. Sí
1. Hombre	3. Transmasculino	No se habilita	1. Sí
1. Hombre	4. Transfemenino	No se habilita	1. Sí
1. Hombre	5. No binario	Cualquier valor válido (1. Sí, 2. No, -99)	1. Sí
1. Hombre	6. Otro	Cualquier valor válido (1. Sí, 2. No, -99)	1. Sí
1. Hombre	1. Masculino	No se habilita	2. No
1. Hombre	7. Prefiere no responder	No se habilita	3. No clasificable
1. Hombre	8. No sabe	No se habilita	3. No clasificable
1. Hombre	2. Femenino	-99. No responde	3. No clasificable
1. Hombre	2. Femenino	2. No	3. No clasificable

Sexo	Género	Identificación trans	Diversidades de género
2. Mujer	1. Masculino	1. Sí	1. Sí
2. Mujer	3. Transmasculino	No se habilita	1. Sí
2. Mujer	4. Transfemenino	No se habilita	1. Sí
2. Mujer	5. No binario	Cualquier valor válido (1. Sí, 2. No, -99)	1. Sí
2. Mujer	6. Otro	Cualquier valor válido (1. Sí, 2. No, -99)	1. Sí
2. Mujer	2. Femenino	No se habilita	2. No
2. Mujer	1. Masculino	-99. No responde	3. No clasificable
2. Mujer	1. Masculino	2. No	3. No clasificable
2. Mujer	7. Prefiere no responder	No se habilita	3. No clasificable
2. Mujer	8. No sabe	No se habilita	3. No clasificable

5.

Control de divulgación estadística



En la actualidad, las oficinas de estadísticas enfrentan importantes desafíos en materia de protección de datos asociados a sus publicaciones. La constante evolución de las técnicas de reidentificación, impulsada por los avances tecnológicos y la proliferación de fuentes de información no tradicionales, plantea nuevos retos para resguardar la confidencialidad de la información y la protección de los datos que las personas entregan con fines de uso estadístico.

La base de microdatos del Censo de Población y Vivienda 2024 fue sometida a un tratamiento de anonimización estadística con el propósito de proteger la confidencialidad de las personas y hogares censados. Este procedimiento se enmarca en las disposiciones establecidas en la Ley del Secreto Estadístico (Ley N° 17.374) y en la Ley sobre protección de la vida privada (Ley N° 19.628), las cuales mandatan al Instituto Nacional de Estadísticas a resguardar la identidad de las personas informantes y evitar la divulgación de información que permita su identificación directa o indirecta, o incluso eventuales procesos de reidentificación.

Los marcos legales que rigen la protección de la confidencialidad definen qué tipo de publicaciones pueden realizar las oficinas de estadística de cada país. La literatura sobre el control de divulgación distingue entre tres tipos: 1) archivo de uso público, 2) archivo de uso científico y 3) enclave de datos o centro de investigación controlado. Para el caso del INE, el archivo de uso público es el único tipo de liberación de microdatos que es aplicable de acuerdo con el marco legal actual (INE, 2021)¹. Un archivo de uso público es un tipo de publicación que debe estar disponible para cualquier usuario, por ejemplo, en el sitio web del INE.

La anonimización aplicada se sustenta en el marco de control de divulgación estadística, que tiene como objetivo reducir el riesgo de reidentificación manteniendo la utilidad de la información para fines analíticos y de investigación. Para ello, se evalúan combinaciones de variables que podrían permitir la identificación indirecta de las personas, considerando además la estructura jerárquica de los datos por viviendas, hogares y personas.

El enfoque implementado privilegia el uso de métodos no perturbativos, que reducen el detalle de las variables sin alterar su coherencia interna ni las relaciones entre ellas.

Por lo tanto, para reducir el riesgo de reidentificación, se aplicaron los siguientes tratamientos a las variables:

- **Recodificación global:** agrupación de categorías detalladas en niveles más amplios de clasificación.
- **Codificación superior:** transformación de variables continuas u ordinales en rangos o categorías resumidas.
- **Supresión parcial de variables:** reducción del detalle en variables específicas para comunas o dominios con menor tamaño poblacional. Para el caso de las variables consideradas como datos personales sensibles, se suprime la información para categorías con menos de 50 casos a nivel de comuna².

¹ El INE publicó el 2021 una guía para el control de divulgación estadística de microdatos. Esta guía está basada en la metodología del Banco Mundial para los procesos de anonimización y control de divulgación. Para más información sobre la guía, se sugiere consultar el siguiente enlace: <https://www.ine.gob.cl/docs/default-source/buenas-practicas/estandares/estandar/documento/guia-control-divulgacion-estadistica-microdatos.pdf>

² La elección del umbral de 50 casos se realizó tras la simulación de escenarios para cada variable sensible, junto con otras características de la persona, con el objetivo de suprimir aquellos casos en donde la combinación de las características presente una baja prevalencia.

La definición de los métodos aplicados a cada variable fue realizada en función del riesgo de reidentificación medido a partir del k-anonimato³, así como de su clasificación en el marco legal chileno como variable sensible.

Las variables sensibles corresponden a datos que, sin ser necesariamente identificadores directos o indirectos, pertenecen al dominio de la vida privada de los informantes y cuya divulgación podría afectar esta.

Siguiendo las definiciones de la Ley 19.628 sobre protección de la vida privada y la Ley 21.719 que entrará en vigor en diciembre de 2026, se consideran como variables sensibles a aquellos datos que refieran al origen étnico de las personas, las creencias religiosas y el género de las personas⁴.

La siguiente tabla describe el tratamiento aplicado a las variables:

Tabla 5.1: Tratamientos de control de divulgación aplicados a variables del Censo 2024

Variable	Tratamiento aplicado	Descripción
Desagregación territorial	Recodificación global	Se privilegia la desagregación a nivel de comuna-área debido a que mantiene un equilibrio adecuado entre riesgo y utilidad de la información.
Tipo de vivienda particular (pregunta 2)	Recodificación global	<p>Agrupación en categorías amplias: se agrupan las categorías “Casa con acceso directo” y “Casa en condominio cerrado” en la categoría “Casa”. Además, se agrupa la categoría “Departamento en edificio con ascensor” y “Departamento en edificio sin ascensor” en la categoría “Departamento”.</p> <p>El resto de las categorías se mantiene igual a como se presentan en el cuestionario.</p>
Edad	Codificación superior y supresión parcial de variables	<p>Recodificación de edades superiores a 85 años o más.</p> <p>En comunas con menos de 6.000 habitantes, la edad se reemplaza por edad quinquenal.</p>

³ El k-anonimato se basa en el principio de que en un conjunto de datos seguro el número de informantes que comparten la misma combinación de valores de identificadores indirectos debe ser mayor que el umbral k. Es decir, cada registro debe tener al menos otros k-registros idénticos, según la concatenación de los valores en los identificadores indirectos presentes en los microdatos.

⁴ La Ley 19.628 define como datos personales sensibles a “aquellos datos personales que se refieren a las características físicas o morales de las personas o a hechos o circunstancias de su vida privada o intimidad, tales como los hábitos personales, el origen racial, las ideologías y opiniones políticas, las creencias o convicciones religiosas, los estados de salud físicos o psíquicos y la vida sexual”. Por otro lado, la Ley 21.719 define como datos personales sensibles a “aquellos datos personales que se refieren a las características físicas o morales de las personas o a hechos o circunstancias de su vida privada o intimidad, que revelen el origen étnico o racial, la afiliación política, sindical o gremial, la situación socioeconómica, las convicciones ideológicas o filosóficas, las creencias religiosas, los datos relativos a la salud, al perfil biológico humano, los datos biométricos, y la información relativa a la vida sexual, a la orientación sexual y a la identidad de género de una persona natural”.

Variable	Tratamiento aplicado	Descripción
Residencia hace 5 años (país) (pregunta 24)	Recodificación global y supresión parcial de variables	<p>Recodificación de países en categorías amplias: se presentan países con mayor representación dentro del Censo 2024 -Argentina, Perú, Bolivia, Colombia, Venezuela y Haití- y el resto de los países es agrupado por continente.</p> <p>En comunas con menos de 6.000 habitantes, es posible identificar si la persona residía hace cinco años antes del momento del Censo en Chile o en el extranjero, siendo suprimida la información relativa al país específico.</p>
Lugar de nacimiento (país) (pregunta 25)	Recodificación global y supresión parcial de variables	<p>Recodificación de países en categorías amplias: se presentan países con mayor representación -Argentina, Perú, Bolivia, Colombia, Venezuela y Haití- y el resto de los países es agrupado por continente.</p> <p>En comunas con menos de 6.000 habitantes, se recodifica para indicar únicamente si la persona nació en Chile o en otro país, siendo suprimida la información específica del lugar de nacimiento para estos casos.</p>
Año de llegada al país (pregunta 26)	Codificación superior	<p>Recodificación de años en categorías más amplias: los años se recodifican en periodos de llegada al país: entre 2023 y 2024, entre 2020 y 2022, entre 2017 y 2019, entre 2014 y 2016, entre 2010 y 2013, entre 2000 y 2009, entre 1990 y 1999, antes de 1990.</p>
País de nacionalidad (pregunta 27)	Recodificación global y supresión parcial de variables	<p>Recodificación de países en categorías amplias: se presentan países con mayor representación -Argentina, Perú, Bolivia, Colombia, Venezuela y Haití- y el resto de los países es agrupado por continente.</p> <p>En comunas con menos de 6.000 habitantes, se recodifica para indicar únicamente si la persona tiene o no nacionalidad chilena, siendo suprimido el país de nacionalidad para estos casos.</p>
Pertenencia pueblos originarios (pregunta 28)	Supresión parcial de variables	<p>Se identifica como variable sensible de acuerdo con las leyes 19.628 y 21.719. Las categorías de respuesta solo se presentan cuando tengan 50 o más casos dentro de la comuna. En caso contrario, son suprimidas.</p>

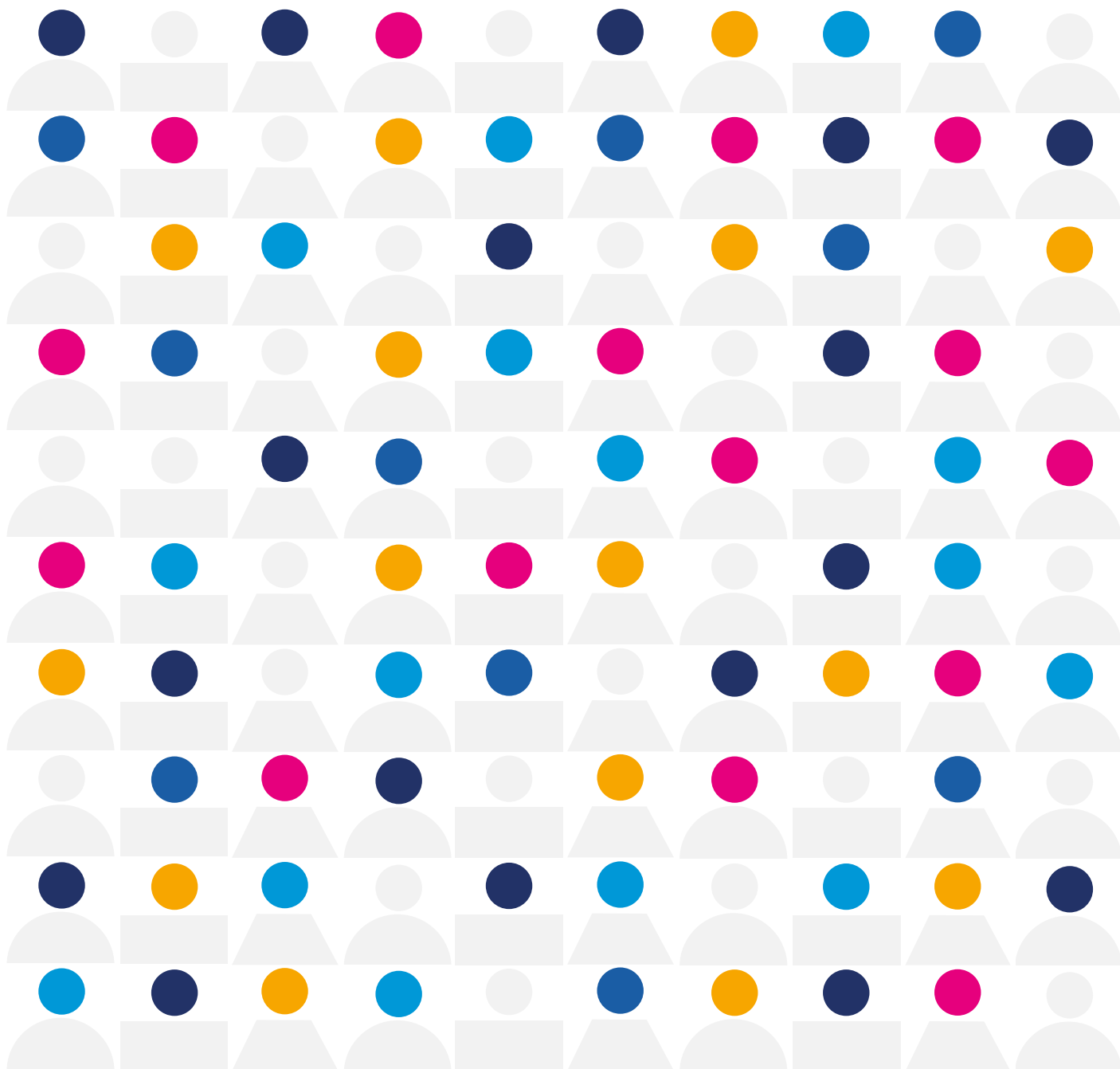
Variable	Tratamiento aplicado	Descripción
Afrodescendencia (pregunta 29)	Supresión parcial de variables	Se identifica como variable sensible de acuerdo con las leyes 19.628 y 21.719. Las categorías de respuesta solo se presentan cuando tengan 50 o más casos dentro de la comuna. En caso contrario, son suprimidas.
Manejo de lengua indígena (pregunta 30)	Supresión parcial de variables	Esta variable debe ser tratada por su relación con la variable de pertenencia a pueblos originarios, en tanto una persona que habla alguna de estas lenguas tendrá una probabilidad muy alta de ser o pertenecer al pueblo asociado a esa lengua. La supresión en la pregunta de pueblo originario implica la supresión en la pregunta sobre manejo de la lengua.
Religión o credo (pregunta 31)	Recodificación global y supresión parcial de variables	Se identifica como variable sensible de acuerdo con las leyes 19.628 y 21.719. Se agrupan las categorías “Judía”, “Musulmana”, “Católica Ortodoxa”, “Budista”, “Hinduista”, “Fe Bahá’í”, “Otra religión o credo” en una única categoría de “Otra religión o credo”. Además, las categorías de respuesta solo se presentan cuando tengan 50 o más casos dentro de la comuna. En caso contrario, son suprimidas.
Nivel educativo (pregunta 34), cursos/ grados aprobados (pregunta 35) y finalización de nivel (pregunta 36)	Recodificación global	<p>Recodificación según la Clasificación Internacional Normalizada de la Educación (CINE 2011) permitiendo comparabilidad internacional entre los niveles educacionales.</p> <p>Para más detalles de cómo fue recodificado, revisar la sección “4.3 Variables calculadas”.</p>
Situación laboral (pregunta 38 y pregunta 39)	Recodificación global	Agrupación en tres categorías: ocupado, desocupado, y fuera de la fuerza de trabajo. Además, para efectos del cálculo del déficit habitacional cuantitativo, en particular para el componente de núcleos allegados hacinados, también se agrupan las categorías para identificar lo que la metodología de cálculo define como personas con independencia económica y personas con dependencia económica.
Situación en el empleo (pregunta 40)	Recodificación global	Agrupación en tres categorías: trabajador dependiente, independiente y familiar no remunerado.

Variable	Tratamiento aplicado	Descripción
Lugar de trabajo (país) (pregunta 44)	Recodificación global	Recodificación de países extranjeros por continentes.
Género e identificación trans (pregunta 49 y pregunta 50)	Recodificación global y supresión parcial de variables	Se identifica como variable sensible de acuerdo con las leyes 19.628 y 21.719. A partir de las respuestas a las preguntas 20, 49 y 50 se crea una variable que identifica si la persona se identifica dentro de las categorías de diversidades de género (trans, no binario, otro), si no lo hace o si no es posible clasificar. La forma de cálculo es descrita en la sección de "Variables calculadas". Además, la variable se presenta solo cuando la categorías "sí" tenga más de 50 observaciones. En caso contrario, la variable es suprimida.

El nivel de desagregación de la información publicada en la base de microdatos corresponde al máximo detalle posible que resguarda la confidencialidad de los informantes. Así, no es posible entregar datos con un mayor nivel de desagregación que el oficialmente difundido, en tanto esto podría implicar un riesgo de identificación directa o indirecta de personas, hogares o viviendas, aun cuando la información no contenga nombres ni otros identificadores directos. Esta restricción responde al cumplimiento estricto de los principios de confidencialidad y secreto estadístico establecidos en las Leyes N° 17.374 y N° 19.628, que rigen todas las operaciones estadísticas realizadas por el INE.

6.

Diccionario de variables



Las bases de datos del Censo 2024 contienen 7.664.466 observaciones para la tabla de vivienda, 6.622.597 para la tabla de hogares y 18.480.432 observaciones para la tabla de personas.

Tabla 6.1: Formato y número de observaciones por tablas de datos

Nombre archivo	Formato disponible	N de variables	N de observaciones	Identificadores
viviendas_censo2024	csv y parquet	25	7.664.466	id_vivienda
hogares_censo2024	csv y parquet	18	6.622.597	id_vivienda, id_hogar
personas_censo2024	csv y parquet	63	18.480.432	id_vivienda, id_hogar, id_persona

Para cada tabla, a continuación, se presentan los nombres de variables, su descripción, etiquetas de categorías, valores, rangos y un conteo de los valores. Esta misma información puede ser encontrada en el archivo descargable del Diccionario de variables del Censo 2024 (<https://censo2024.ine.gob.cl/resultados/>).

6.1 Valores especiales

Para homologar, se establecieron los siguientes valores para determinados casos y que son transversales a todas las tablas de datos:

- **-66**: indica que el valor de este registro fue suprimido por anonimización¹.
- **-99**: valor asignado para los casos en que las personas informantes no entregaron una respuesta a una pregunta.
- **NA**: Corresponde al valor *missing*. Este valor se utiliza cuando una pregunta no aplica a una unidad (por ejemplo: las personas de 14 años o menos tienen valor NA en la variable que identifica si la persona está ocupada, desocupada o inactiva (*sit_fuerza_trabajo*). En algunos softwares es representado por un ".", como es el caso de STATA.

¹ Para revisar información sobre este proceso se sugiere ir al capítulo de **Control de divulgación estadística**.

6.2 Variables de viviendas

Tabla 6.2: Diccionario de Variables - Vivienda

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
id_vivienda	Llave identificadora de la vivienda o de la entrevista	1: 7664466	Válidos	1: 7664466	Todos los registros	7.664.466
region	Región	1:16	Ver diccionario códigos territoriales	1:16	Todos los registros	7.664.466
provincia	Provincia	11:163	Ver diccionario códigos territoriales	11:163	Todos los registros	7.664.466
comuna	Comuna	1101:16305	Ver diccionario códigos territoriales	1101:16305	Todos los registros	7.664.466
area	Área urbana o rural	1	Urbano	1:2	Todos los registros	6.365.998
		2	Rural			1.298.468
tipo_operativo	Tipo de operativo	1	Persona en situación de calle	1:3	Todos los registros	21.750
		2	Vivienda particular			7.638.396
		3	Vivienda colectiva			4.320
p2_tipo_vivienda	2. Tipo de vivienda particular ²	1	Casa	1:9; NA (no aplica)	Todas las viviendas particulares	5.937.284
		3	Departamento			1.579.970
		5	Vivienda tradicional indígena (ruka u otras)			1.967
		6	Pieza en casa antigua o en conventillo			51.305
		7	Mediagua, mejora, vivienda de emergencia, rancho o choza			27.490
		8	Móvil (carpa, casa rodante o similar)			1.557
		9	Otro tipo de vivienda particular			38.823
		NA	No aplica			26.070
p3a_estado_ocupacion	3a. Estado de ocupación de la vivienda	1	Ocupada	1:2; NA (no aplica)	Todas las viviendas particulares	6.730.463
		2	Desocupada			907.933
		NA	No aplica			26.070

² Los números de las categorías no son correlativos porque se mantiene los valores de la versión impresa del cuestionario de viviendas particulares.

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p3b_estado_ocupacion	3b. Estado de ocupación de la vivienda (detalle) ³	1	Moradores presentes	1:10; NA (no aplica)	Todas las viviendas particulares	6.408.172
		2	Moradores ausentes			208.084
		6	Otras razones de no entrevista			114.207
		7	En venta o arriendo			173.315
		8	Vivienda de temporada (veraneo o de uso temporal)			476.135
		9	Abandonada			97.052
		10	Otra razón de desocupación			161.431
		NA	No aplica			26.070
p4a_mat_paredes	4a. ¿Cuál es el material de construcción en las paredes exteriores?	1	Hormigón armado	1:6; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	1.499.949
		2	Albañilería: bloque de cemento, ladrillo o piedra			2.754.185
		3	Tabique forrado por ambas caras (madera o acero)			1.854.738
		4	Tabique sin forro interior (madera u otro)			176.313
		5	Adobe, barro, pirca, quincha u otro material artesanal			106.790
		6	Materiales precarios o de desecho: cartón, sacos, trozos de latas o plásticos, etc.			14.752
		-99	No respuesta			1.445
		NA	No aplica			1.256.294
p4b_mat_techo	4b. ¿Cuál es el material de construcción en la cubierta del techo?	1	Tejas o tejuelas de arcilla, metálicas, de cemento, de madera, asfálticas o plásticas	1:8; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	1.069.613
		2	Losa hormigón			958.191
		3	Planchas metálicas de zinc, cobre, etc.			3.658.013
		4	Planchas de fibrocemento tipo pizarreño			699.395
		5	Fonolita o plancha de fieltro embreado			10.847
		6	Paja, coirón, totora o caña			1.090
		7	Materiales precarios o de desecho: cartón, sacos, trozos de latas o plásticos, etc.			6.800
		8	Sin cubierta sólida de techo			1.872
		-99	No respuesta			2.351
		NA	No aplica			1.256.294
p4c_mat_piso	4c. ¿Cuál es el material de construcción en el piso?	1	Parquet, piso flotante, cerámico, madera, alfombra, flexit, cubrepiso u otro similar; sobre radier o vigas de madera	1:5; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	6.036.357
		2	Radier sin revestimiento			145.177
		3	Baldosa de cemento			143.611
		4	Capa de cemento sobre tierra			63.853
		5	Tierra			17.010
		-99	No respuesta			2.164
		NA	No aplica			1.256.294

3 Agrupa las categorías de "Informante rechaza la entrevista", "Informante tiene un impedimento físico y/o mental para contestar", "Informante no habla español" y "Otra razón de no entrevista".

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p5_num_dormitorios	5. ¿Cuántas piezas de esta vivienda se usan exclusivamente como dormitorio?	0	0	0;6; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	36.835
		1	1			1.575.168
		2	2			2.419.639
		3	3			1.785.285
		4	4			451.532
		5	5			100.036
		6	6 o más			38.643
		-99	No respuesta			1.034
		NA	No aplica			1.256.294
p6_fuente_agua	6. El agua que usa esta vivienda proviene principalmente de:	1	Red pública	1;4; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	5.914.136
		2	Pozo o noria			263.309
		3	Camión aljibe			131.141
		4	Río, vertiente, estero, canal, lago, agua lluvia, etc.			98.243
		-99	No respuesta			1.343
		NA	No aplica			1.256.294
p7_distrib_agua	7. ¿Cuál es el sistema de distribución del agua en esta vivienda?	1	Con llave dentro de la vivienda	1;3; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	6.196.815
		2	Con llave dentro del sitio, pero fuera de la vivienda			160.098
		3	No tiene sistema, la acarrea			49.617
		-99	No respuesta			1.642
		NA	No aplica			1.256.294
p8_serv_hig	8. El servicio higiénico (WC) principal de esta vivienda es o está:	1	Dentro de la vivienda, conectado a una red de alcantarillado	1;9; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	5.344.651
		2	Fuera de la vivienda, conectado a una red de alcantarillado			104.723
		3	Conectado a una fosa séptica			767.214
		4	Conectado a pozo negro (letrina sanitaria o cajón)			164.499
		5	En un cajón sobre acequia o canal			3.868
		6	En un cajón conectado a otro sistema			2.310
		7	Baño químico			2.141
		8	Conectado a baño seco			2.443
		9	No tiene servicio higiénico			14.794
		-99	No respuesta			1.529
		NA	No aplica			1.256.294
p9_fuente_elect	9. La electricidad de esta vivienda proviene principalmente de:	1	Red pública	1;6; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	6.316.147
		2	Generador con diésel o bencina			16.021
		3	Placa solar			32.343
		4	Energía eólica (viento)			695
		5	Otro			16.430
		6	No tiene energía eléctrica			24.917
		-99	No respuesta			1.619
		NA	No aplica			1.256.294

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p10_basura	10. ¿Cuál es el principal medio de eliminación de basura de esta vivienda?	1	La recogen los servicios de aseo	1:5; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	6.221.171
		2	La entierra y/o quema			111.904
		3	La deja en terreno eriazos, quebrada o zanja			14.058
		4	La tira al río, laguna o mar			1.009
		5	Otro			58.798
		-99	No respuesta			1.232
		NA	No aplica			1.256.294
p11a_num_personas	11. ¿Cuántas personas residen habitualmente en esta vivienda?	1:60		1:60; NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	6.408.172
		NA	No aplica			1.256.294
p11b_comparte_gasto	11.1. De las personas que residen habitualmente en esta vivienda, ¿todas comparten los gastos para alimentación?	1	Sí	1:2; NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	4.965.946
		2	No			153.693
		NA	No aplica			2.544.827
p11c_num_hogar	11.2. ¿Cuántos grupos tienen gastos separados de alimentación?	2:26		2:26; NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	153.693
		NA	No aplica			7.510.773
cant_per	Número de personas censadas en la vivienda	0:5806		0:5806	Todos los registros	7.664.466
cant_hog	Número de hogares censados en la vivienda particular	0:26		0:26; NA (no aplica)	Todas las viviendas particulares	7.638.396
		NA	No aplica			26.070
indice_hacinamiento	Índice de hacinamiento	1	Vivienda sin hacinamiento	1:3; -99 (no respuesta); NA (no aplica)	Todas las viviendas particulares ocupadas con moradores presentes	6.039.952
		2	Vivienda con hacinamiento medio			330.955
		3	Vivienda con hacinamiento crítico			36.231
		-99	No respuesta			1.034
		NA	No aplica			1.256.294
comuna_bajo_umbral	Comuna con menos de 6.000 personas censadas	1	Sí	1:2	Todos los registros	99.321
		2	No			7.565.145

6.3 Variables de hogar

Tabla 6.3: Diccionario de Variables - Hogar

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
id_vivienda	Llave identificadora de la vivienda o de la entrevista	1:7664466	Válidos	1:7664466	Todos los registros	6.622.597
id_hogar	Llave identificadora de hogar	0:26	Válidos	0:26	Todos los registros	6.622.597
region	Región	1:16	Ver diccionario códigos territoriales	1:16	Todos los registros	6.622.597
provincia	Provincia	11:163	Ver diccionario códigos territoriales	11:163	Todos los registros	6.622.597
comuna	Comuna	1101:16305	Ver diccionario códigos territoriales	1101:16305	Todos los registros	6.622.597

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
area	Área urbana o rural	1	Urbano	1:2	Todos los registros	5.733.950
		2	Rural			888.647
tipo_operativo	Tipo de operativo	1	Persona en Situación de Calle	1:3	Todos los registros	21.750
		2	Vivienda Particular			6.596.527
		3	Vivienda Colectiva			4.320
pi2_tenencia_viv	12. La vivienda que ocupa este hogar es:	1	Propia pagada	1:9; -99 (no respuesta), NA (no aplica)	Todos los hogares	3.120.238
		2	Propia pagándose			910.247
		3	Arrendada con contrato			1.229.328
		4	Arrendada sin contrato			498.871
		5	Cedida por trabajo o servicio			90.697
		6	Cedida por familiar u otro			511.118
		7	Usufructo: solo uso y goce			61.019
		8	Ocupada de hecho			57.788
		9	Propiedad en sucesión o litigio			116.125
		-99	No respuesta			1.096
		NA	No aplica			26.070
pi3_comb_cocina	13. ¿Cuál es la principal fuente de energía o combustible que utilizan para cocinar?	1	Gas	1:8; -99 (no respuesta); NA (no aplica)	Todos los hogares	5.780.274
		2	Parafina o petróleo			3.832
		3	Leña			339.220
		4	Pellet			1.960
		5	Carbón			1.520
		6	Electricidad			439.520
		7	Energía solar (ej. cocina u horno solar)			639
		8	No utiliza fuente de energía o combustible para cocinar			28.860
		-99	No respuesta			702
		NA	No aplica			26.070
pi4_comb_calefaccion	14. ¿Cuál es la principal fuente de energía o combustible que utilizan para calefaccionar?	1	Gas	1:8; -99 (no respuesta); NA (no aplica)	Todos los hogares	1.688.313
		2	Parafina o petróleo			680.755
		3	Leña			1.606.997
		4	Pellet			137.853
		5	Carbón			13.426
		6	Electricidad			1.283.687
		7	Otra			24.691
		8	No utiliza fuente de energía o combustible para calefaccionar			1.159.827
		-99	No respuesta			978
		NA	No aplica			26.070
pi5a_serv_tel_movil	15. Este hogar, ¿dispone de teléfono móvil, celular o smartphone?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Todos los hogares	6.475.071
		2	No			119.949
		-99	No respuesta			1.507
		NA	No aplica			26.070

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p15b_serv_compu	15. Este hogar, ¿dispone de computador (escritorio o portátil)?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Todos los hogares	3.621.619
		2	No			2.964.123
		-99	No respuesta			10.785
		NA	No aplica			26.070
p15c_serv_tablet	15. Este hogar, ¿dispone de tablet?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Todos los hogares	1.645.321
		2	No			4.939.222
		-99	No respuesta			11.984
		NA	No aplica			26.070
p15d_serv_internet_fija	15. Este hogar, ¿dispone de internet fija?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Todos los hogares	4.250.122
		2	No			2.338.834
		-99	No respuesta			7.571
		NA	No aplica			26.070
p15e_serv_internet_movil	15. Este hogar, ¿dispone de internet móvil desde un celular, tablet o BAM?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Todos los hogares	5.734.118
		2	No			854.837
		-99	No respuesta			7.572
		NA	No aplica			26.070
p15f_serv_internet_satelital	15. Este hogar, ¿dispone de internet por conexión satelital?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Todos los hogares	288.473
		2	No			6.293.135
		-99	No respuesta			14.919
		NA	No aplica			26.070
tipologia_hogar	Tipología de hogar	1	Unipersonal	1:7; NA (no aplica)	Todos los hogares	1.437.648
		2	Nuclear monoparental			985.912
		3	Nuclear pareja sin hijos			940.030
		4	Nuclear pareja con hijos			1.796.839
		5	Extenso			945.125
		6	Compuesto			112.706
		7	Sin núcleo			378.267
		NA	No aplica			26.070
comuna_bajo_umbral	Comuna con menos de 6.000 personas censadas	1	Sí	1:2	Todos los registros	60.761
		2	No			6.561.836

6.4 Variables de personas

Tabla 6.4: Diccionario de Variables – Persona

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
id_vivienda	Llave identificadora de la vivienda o de la entrevista	1:7664466	Válidos	1:7664466	Todos los registros	18.480.432
id_hogar	Llave identificadora de hogar	0:26	Válidos	0:26	Todos los registros	18.480.432
id_persona	Llave identificadora de persona	1:5806	Válidos	1:5806	Todos los registros	18.480.432
region	Región	1:16	Ver diccionario códigos territoriales	1:16	Todos los registros	18.480.432
provincia	Provincia	11:163	Ver diccionario códigos territoriales	11:163	Todos los registros	18.480.432
comuna	Comuna	1101:16305	Ver diccionario códigos territoriales	1101:16305	Todos los registros	18.480.432
area	Área urbana o rural	1	Urbano	1:2	Todos los registros	16.018.474
		2	Rural			2.461.958
tipo_operativo	Tipo de operativo	1	Persona en situación de calle	1:3	Todos los registros	21.750
		2	Vivienda particular			18.339.655
		3	Vivienda colectiva			119.027
parentesco	19. ¿Qué relación de parentesco tiene con la jefa o jefe de hogar?	1	Jefe/a de hogar	1:17, 19	Todas las personas	6.596.527
		2	Esposo/a o cónyuge			2.090.815
		3	Conviviente por unión civil			61.672
		4	Conviviente de hecho o pareja			1.226.712
		5	Hijo/a			5.861.272
		6	Hijo/a del cónyuge, conviviente o pareja			200.617
		7	Hermano/a			281.643
		8	Padre/Madre			247.279
		9	Cuñado/a			59.280
		10	Suegro/a			73.702
		11	Yerno/Nuera			193.892
		12	Nieto/a			853.354
		13	Abuelo/a			10.030
		14	Otro Pariente			330.270
		15	No Pariente			236.858
		16	Servicio doméstico puertas adentro			15.732
		17	Persona en vivienda colectiva			119.027
		19	Persona en operativo calle			21.750
sexo	20. ¿Cuál es su sexo?	1	Hombre	1:2	Todas las personas	8.967.033
		2	Mujer			9.513.399

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
edad	21. ¿Cuántos años cumplidos tiene?	0:85	Válidos	0:85; -66 (anonimizado)	Todas las personas	18.329.233
		-66	Valor suprimido por anonimización			151.199
edad_quinquenal	Edad en quinquenios	0	0-4	0:85	Todas las personas	870.693
		5	5-9			1.147.515
		10	10-14			1.256.440
		15	15-19			1.219.347
		20	20-24			1.273.193
		25	25-29			1.383.669
		30	30-34			1.527.489
		35	35-39			1.408.198
		40	40-44			1.270.493
		45	45-49			1.151.776
		50	50-54			1.173.352
		55	55-59			1.133.239
		60	60-64			1.077.790
		65	65-69			870.801
		70	70-74			646.241
		75	75-79			477.186
		80	80-84			317.424
		85	85 o más			275.586
p23_est_civil	23. ¿Cuál es su estado conyugal o civil actual?	1	Casado/a	1:8; -99 (no respuesta); NA (no aplica)	Personas de 15 años o más	4.608.520
		2	Conviviente o pareja sin acuerdo de unión civil			2.638.561
		3	Conviviente civil (con acuerdo de unión civil)			133.624
		4	Anulado/a			16.450
		5	Separado/a			462.915
		6	Divorciado/a			482.427
		7	Viudo/a			740.311
		8	Soltero/a			6.068.849
		-99	No respuesta			54.127
		NA	No aplica			3.274.648
p24_lug_resid5	24. ¿En qué comuna o país vivía en abril de 2019?	1	Aún no nacía	1:4; -99 (no respuesta)	Todas las personas	870.693
		2	En esta comuna			14.426.941
		3	En otra comuna			2.442.116
		4	En otro país			604.170
		-99	No respuesta			136.512

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p24_lug_resid5_esp	Lugar específico de residencia hace 5 años	2:16305	Ver diccionario códigos territoriales específicos	2:16305; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	17.466.038
		-66	Valor suprimido por anonimización			2.459
		-99	No respuesta en país de residencia hace 5 años			141.242
		NA	No aplica			870.693
p25_lug_nacimiento	25. Cuándo nació, ¿en qué comuna o país vivía su madre?	1	En esta comuna	1:3; -99 (no respuesta)	Todas las personas	9.457.200
		2	En otra comuna			7.301.521
		3	En otro país			1.608.650
		-99	No respuesta			113.061
p25_lug_nacimiento_rec	Nació en Chile	1	Sí	1:2; -99 (no respuesta)	Todas las personas	16.758.721
		2	No			1.608.650
		-99	No respuesta			113.061
p25_lug_nacimiento_esp	Lugar específico de nacimiento	2:16305	Ver diccionario códigos territoriales específicos	2:16305; -66 (anonimizado); -99 (no respuesta)	Todas las personas	18.191.426
		-66	Valor suprimido por anonimización			151.199
		-99	No respuesta			137.807
p26_llegada_perodo	26. ¿En qué año llegó a vivir a Chile?	1	Entre 2023 y 2024	1:8; -99 (no respuesta); NA (no aplica)	Personas que nacieron en otro país	188.790
		2	Entre 2020 y 2022			366.878
		3	Entre 2017 y 2019			533.689
		4	Entre 2014 y 2016			190.158
		5	Entre 2010 y 2013			111.485
		6	Entre 2000 y 2009			121.553
		7	Entre 1990 y 1999			54.026
		8	Antes de 1990			36.811
		-99	No respuesta			5.260
		NA	No aplica			16.871.782
p27_nacionalidad	27. ¿Cuál es su nacionalidad?	1	Chilena (exclusivamente)	1:3; -66 (anonimizado); -99 (no respuesta)	Todas las personas	16.500.963
		2	Chilena y otra (doble nacionalidad)			242.909
		3	Otra nacionalidad			1.484.478
		-66	Valor suprimido por anonimización			151.199
		-99	No respuesta			100.883
p27_nacionalidad_esp	País de nacionalidad	2:862	Ver diccionario códigos territoriales específicos	2:862; -66 (anonimizado); -99 (no respuesta)	Todas las personas	18.226.433
		-66	Valor suprimido por anonimización			151.199
		-99	No respuesta			102.800

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p27_nacionalidad_rec	Nacionalidad chilena	1	Sí	1;2; -99 (no respuesta)	Todas las personas	16.889.540
		2	No			1.489.385
		-99	No respuesta			101.507
p28_autoid_pueblo	28. ¿Es o se considera perteneciente a algún pueblo indígena u originario?	1	Sí	1;2; -99 (no respuesta)	Todas las personas	2.105.863
		2	No			16.267.072
		-99	No respuesta			107.497
p28_pueblo_pert	28. ¿A cuál?	1	Mapuche	1;12; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Todas las personas que son o se consideran de algún pueblo indígena u originario	1.621.748
		2	Aymara			175.353
		3	Rapa Nui			4.383
		4	Atacameño o Lickanantay			33.686
		5	Quechua			43.779
		6	Colla			20.075
		7	Diaguita			150.692
		8	Kawésqar			868
		9	Yagán			397
		10	Chango			10.244
		11	Selk'nam			143
		12	Otro			20.630
		-66	Valor suprimido por anonimización			21.470
		-99	No respuesta			2.395
		NA	No aplica			16.374.569
p29_afrodescendencia	29. De acuerdo con sus antepasados, tradiciones y cultura, es o se considera:	1	Afrodescendiente	1;7; -66 (anonimizado); -99 (no respuesta)	Todas las personas	90.254
		2	Afrochileno/a			35.711
		3	Negro/a			11.771
		4	Del Pueblo Tribal Afrodescendiente Chileno			10.066
		5	Moreno/a de Azapa			6.162
		6	Negro/a de la Chimba			1.463
		7	Ninguna de las anteriores			18.195.851
		-66	Valor suprimido por anonimización			18.763
		-99	No respuesta			110.391
p29_afrodescendencia_rec	Afrodescendencia	1	Sí	1;2; -99 (no respuesta)	Todas las personas	174.190
		2	No			18.195.851
		-99	No respuesta			110.391

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p30_lengua_indigena	30. ¿Habla o entiende una de las siguientes lenguas indígenas u originarias?	1	Mapuzugún (lengua mapuche)	1:9; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	380.780
		2	Aymara			60.080
		3	Quechua			37.837
		4	Rapa Nui			6.376
		5	Ckunza			2.541
		6	Kawésqar			546
		7	Yagán			710
		8	Otra lengua indígena de Chile			17.604
		9	No habla ni entiende ninguna lengua indígena de Chile			16.996.225
		-66	Valor suprimido por anonimización			3.990
		-99	No respuesta			103.050
		NA	No aplica			870.693
p30_lengua_indigena_rec	Habla o entiende alguna lengua originaria	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	510.464
		2	No			16.996.225
		-99	No respuesta			103.050
		NA	No aplica			870.693
p31_religion	31. ¿Cuál es su religión o credo? ⁴	1	Católica	1:999; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 15 años o más	8.168.578
		2	Evangélica o Protestante			2.466.199
		5	Iglesia de Jesucristo de los Santos de los Últimos Días			91.796
		10	Testigo de Jehová			112.688
		12	Ninguna			3.903.308
		999	Otra religión o credo			363.305
		-66	Valor suprimido por anonimización			12.395
		-99	No respuesta			87.515
		NA	No aplica			3.274.648
p31_religion_rec	Tiene religión o credo	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Personas de 15 años o más	11.214.961
		2	No			3.903.308
		-99	No respuesta			87.515
		NA	No aplica			3.274.648

⁴ Los números de las categorías no son correlativos porque se mantiene los valores de la versión impresa del cuestionario de viviendas particulares.

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p32a_dificultad_ver	32. ¿Tiene dificultad para ver, aun usando anteojos o lentes?	1	No, sin dificultad	1:4; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	13.905.057
		2	Sí, algo de dificultad			2.758.800
		3	Sí, mucha dificultad			795.335
		4	No puede hacerlo			51.667
		-99	No respuesta			98.880
		NA	No aplica			870.693
p32b_dificultad_oir	32. ¿Tiene dificultad para oír, aun usando un dispositivo auditivo	1	No, sin dificultad	1:4; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	16.144.858
		2	Sí, algo de dificultad			964.940
		3	Sí, mucha dificultad			357.490
		4	No puede hacerlo			41.580
		-99	No respuesta			100.871
		NA	No aplica			870.693
p32c_dificultad_mover	32. ¿Tiene dificultad para caminar o subir escaleras?	1	No, sin dificultad	1:4; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	15.289.715
		2	Sí, algo de dificultad			1.307.950
		3	Sí, mucha dificultad			759.016
		4	No puede hacerlo			149.696
		-99	No respuesta			103.362
		NA	No aplica			870.693
p32d_dificultad_cogni	32. ¿Tiene dificultad para recordar o concentrarse?	1	No, sin dificultad	1:4; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	15.198.358
		2	Sí, algo de dificultad			1.808.201
		3	Sí, mucha dificultad			424.594
		4	No puede hacerlo			73.643
		-99	No respuesta			104.943
		NA	No aplica			870.693
p32e_dificultad_cuidado	32. ¿Tiene dificultad para realizar tareas de cuidado personal, como bañarse o vestirse?	1	No, sin dificultad	1:4; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	16.880.403
		2	Sí, algo de dificultad			324.960
		3	Sí, mucha dificultad			155.590
		4	No puede hacerlo			146.513
		-99	No respuesta			102.273
		NA	No aplica			870.693

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p32f_dificultad_comunic	32. ¿Tiene dificultad para hablar o comunicarse usando su idioma habitual, por ejemplo, entender o ser entendido por otras personas?	1	No, sin dificultad	1:4; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	17.002.154
		2	Sí, algo de dificultad			307.637
		3	Sí, mucha dificultad			127.428
		4	No puede hacerlo			72.283
		-99	No respuesta			100.237
		NA	No aplica			870.693
discapacidad	Discapacidad	1	Con discapacidad	1:2; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	1.950.388
		2	Sin discapacidad			15.561.356
		-99	No respuesta			97.995
		NA	No aplica			870.693
p33_edu_asiste	33. ¿Asiste actualmente a la educación formal? Incluye educación parvularia, especial, básica, media y superior	1	Sí	1:2; -99 (no respuesta)	Todas las personas	5.106.338
		2	No			13.274.483
		-99	No respuesta			99.611
asistencia_parv	Asistencia a nivel de educación parvularia o preescolar	1	Sí	1:2; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 0 a 5 años	555.041
		2	No			506.360
		-66	Valor suprimido por anonimización			8.430
		-99	No respuesta			9.544
		NA	No aplica			17.401.057
asistencia_basica	Asistencia a nivel de educación básica	1	Sí	1:2; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 6 a 13 años	1.822.358
		2	No			87.531
		-66	Valor suprimido por anonimización			15.076
		-99	No respuesta			12.922
		NA	No aplica			16.542.545
asistencia_media	Asistencia a nivel de educación media	1	Sí	1:2; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 14 a 17 años	858.688
		2	No			123.361
		-66	Valor suprimido por anonimización			7.975
		-99	No respuesta			6.645
		NA	No aplica			17.483.763

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
asistencia_superior	Asistencia a nivel de educación superior	1	Sí	1:2; -66 (anonimizado); -99 (no respuesta); NA (no aplica)	Personas de 18 a 24 años	808.427
		2	No			921.756
		-66	Valor suprimido por anonimización			10.849
		-99	No respuesta			12.225
		NA	No aplica			16.727.175
p37_alfabet	37. ¿Sabe leer y escribir?	1	Sí	1:2; -99 (no respuesta); NA (no aplica)	Personas de 5 años o más	16.769.140
		2	No			725.905
		-99	No respuesta			114.694
		NA	No aplica			870.693
escolaridad	Años de escolaridad	0:24	Válidos	0:24; -99 (no respuesta)	Todas las personas	18.348.374
		-99	No respuesta			132.058
cine11	Logro educativo de acuerdo con la Clasificación Internacional Normalizada de la Educación	1	01: Nunca cursó un programa educativo	1:12; -99 (no respuesta)	Todas las personas	552.698
		2	02: Educación de la primera infancia (incluye la forma parcial)			821.251
		3	03: Primaria en forma parcial (sin conclusión del nivel)			2.122.256
		4	10: Educación primaria (nivel 1)			1.192.382
		5	14: Educación primaria (nivel 2), con orientación general			3.362.035
		6	24: Educación secundaria, con orientación general			4.371.768
		7	25: Educación secundaria, con orientación vocacional			1.568.048
		8	35: Educación terciaria de ciclo corto, con orientación vocacional			1.412.112
		9	46: Grado de educación terciaria o nivel equivalente, con orientación no especificada			2.545.699
		10	56: Nivel de maestría, especialización o equivalente, con orientación no especificada			277.835
		11	64: Nivel de doctorado o equivalente, con orientación académica			31.694
		12	98: Educación especial o diferencial			100.691
		-99	No respuesta			121.963

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
sit_fuerza_trabajo	Situación en la fuerza de trabajo	1	Ocupado	1;3; -99 (no respuesta); NA (no aplica)	Personas de 15 años o más	8.277.501
		2	Desocupado			777.920
		3	Fuera de la fuerza de trabajo			6.060.837
		-99	No respuesta			89.526
		NA	No aplica			3.274.648
p40_cise_rec	40. ¿Trabajó cómo...?	1	Independiente	1;3; -99 (no respuesta); NA (no aplica)	Personas ocupadas	2.794.082
		2	Dependiente			5.352.532
		3	Trabajador/a familiar o personal no remunerado en un negocio de un integrante de su familia			127.429
		-99	No respuesta			3.458
		NA	No aplica			10.202.931
depend_econ_deficit_hab	Dependencia económica	1	Económicamente independiente	1;2; -99 (no respuesta)	Todas las personas	10.417.681
		2	Económicamente dependiente			7.970.061
		-99	No respuesta			92.690
cod_ciuo	CIUO-08.CL a 1 dígito	1	Directores, gerentes y administradores	0;9, 999; -99 (no respuesta); NA (no aplica)	Personas ocupadas	383.706
		2	Profesionales, científicos e intelectuales			1.490.307
		3	Técnicos y profesionales de nivel medio			884.233
		4	Personal de apoyo administrativo			518.314
		5	Trabajadores de los servicios y vendedores de comercios y mercados			1.622.622
		6	Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros			224.291
		7	Artesanos y operarios de oficios			1.038.385
		8	Operadores de instalaciones, máquinas y ensambladores			662.440
		9	Ocupaciones elementales			1.132.704
		0	Ocupaciones de las fuerza armadas			24.165
		999	Respuesta no codificable			275.787
		-99	No respuesta			20.547
		NA	No aplica			10.202.931

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
cod_caenes	CAENES a 1 dígito	A	Agricultura, ganadería, silvicultura y pesca	A:U, 999; -99 (no respuesta); NA (no aplica)	Personas ocupadas	484.597
		B	Explotación de minas y canteras			275.517
		C	Industrias manufactureras			637.429
		D	Suministro de electricidad, gas, vapor y aire acondicionado			85.063
		E	Suministro de agua; evacuación de aguas residuales, gestión de desechos y descontaminación			60.844
		F	Construcción			640.645
		G	Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas			1.365.418
		H	Transporte y almacenamiento			558.899
		I	Actividades de alojamiento y de servicio de comidas			390.006
		J	Información y comunicaciones			169.467
		K	Actividades financieras y de seguros			158.834
		L	Actividades inmobiliarias			74.536
		M	Actividades profesionales, científicas y técnicas			315.654
		N	Actividades de servicios administrativos y de apoyo			324.195
		O	Administración pública y defensa; planes de seguridad social de afiliación obligatoria			378.190
		P	Enseñanza			702.416
		Q	Actividades de atención de la salud humana y de asistencia social			671.212
		R	Actividades artísticas, de entretenimiento y recreativas			94.624
		S	Otras actividades de servicios			241.797
		T	Actividades de los hogares como empleadores; actividades no diferenciadas de los hogares como productores de bienes y servicios para uso propio			215.173
		U	Actividades de organizaciones y órganos extraterritoriales			2.917
		999	Respuesta no codificable			406.001
		-99	No respuesta			24.067
		NA	No aplica			10.202.931

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p44_lug_trab	44. ¿En qué comuna o país se ubica su trabajo?	1	En su comuna, en la vivienda donde reside	1:5; -99 (no respuesta); NA (no aplica)	Personas ocupadas	1.637.632
		2	En su comuna, fuera de la vivienda donde reside			3.311.492
		3	En otra comuna			2.856.674
		4	En otro país			6.751
		5	Varias comunas o países			460.346
		-99	No respuesta			4.606
		NA	No aplica			10.202.931
p44_lug_trab_esp	Lugar de trabajo	2:16305	Ver diccionario códigos territoriales específicos	2:16305; -99 (no respuesta); NA (no aplica)	Personas ocupadas	8.268.472
		-99	No respuesta			9.029
		NA	No aplica			10.202.931
p45_medio_ transporte	45. ¿Cuál es el medio de transporte principal que utiliza para dirigirse a su lugar de trabajo?	1	Auto particular	1:7; -99 (no respuesta); NA (no aplica)	Personas ocupadas que no trabajan en la vivienda donde residen	2.341.926
		2	Transporte público (bus, micro, metro, tren, taxi, colectivo)			2.793.070
		3	Caminando			677.171
		4	Bicicleta (incluye scooter)			254.215
		5	Motocicleta			95.687
		6	Caballo, lancha o bote			6.348
		7	Otro			465.580
		-99	No respuesta			1.266
		NA	No aplica			11.845.169
p46a_tot_hijs_nac	46. ¿Cuántas hijas e hijos nacidos vivos ha tenido en total?	0:23	Válidos	0:23; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más	7.862.575
		-99	No respuesta			44.706
		NA	No aplica			10.573.151
p46b_hijas_nac	46. Del total de hijas e hijos que se encuentran vivos, ¿cuántas son mujeres?	0:22	Válidos	0:22; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más	5.570.891
		-99	No respuesta			2.085
		NA	No aplica			12.907.456
p46c_hijos_nac	46. Del total de hijas e hijos que se encuentran vivos, ¿cuántos son hombres?	0:22	Válidos	0:22; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más	5.570.891
		-99	No respuesta			2.085
		NA	No aplica			12.907.456

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
p47a_tot_hijs_sobrev	47. ¿Cuántas hijas e hijos están vivos actualmente?	0:22	Válidos	0:22; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más que han tenido hijos nacidos vivos	5.569.843
		-99	No respuesta			3.133
		NA	No aplica			12.907.456
p47b_hijas_sobrev	47. Del total de hijas e hijos que se encuentran vivos, ¿cuántas son mujeres?	0:15	Válidos	0:15; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más que han tenido hijos nacidos vivos	5.475.748
		-99	No respuesta			58.447
		NA	No aplica			12.946.237
p47c_hijos_sobrev	47. Del total de hijas e hijos que se encuentran vivos, ¿cuántos son hombres?	0:19	Válidos	0:19; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más que han tenido hijos nacidos vivos	5.475.748
		-99	No respuesta			58.447
		NA	No aplica			12.946.237
p48_anio_nac_uh	48. ¿En qué mes y año nació su última hija o hijo nacido vivo? Año	1934:2024	Válidos	1934:2024; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más que han tenido hijos nacidos vivos	5.354.481
		-99	No responde			218.495
		NA	No aplica			12.907.456
p48_mes_nac_uh	48. ¿En qué mes y año nació su última hija o hijo nacido vivo? Mes	1	Enero	1:12; -99 (no respuesta); NA (no aplica)	Mujeres de 15 años o más que han tenido hijos nacidos vivos	459.083
		2	Febrero			410.413
		3	Marzo			455.275
		4	Abril			436.463
		5	Mayo			443.334
		6	Junio			441.553
		7	Julio			431.142
		8	Agosto			445.898
		9	Septiembre			457.217
		10	Octubre			452.482
		11	Noviembre			434.750
		12	Diciembre			444.343
		-99	No respuesta			261.023
		NA	No aplica			12.907.456

Nombre variable	Descripción de la variable	Valor	Etiqueta de categoría	Rango	Universo	Conteo
div_genero	Diversidades de género	1	Sí	1:3; -66 (anonimizado); NA (no aplica)	Personas de 18 años o más	64.101
		2	No			12.983.214
		3	No clasificable			143.512
		-66	Valor suprimido por anonimización			1.275.674
		NA	No aplica			4.013.931
comuna_bajo_umbral	Comuna con menos de 6.000 personas censadas	1	Sí	1:2	Todos los registros	151.199
		2	No			18.329.233

6.5 Diccionario de códigos territoriales

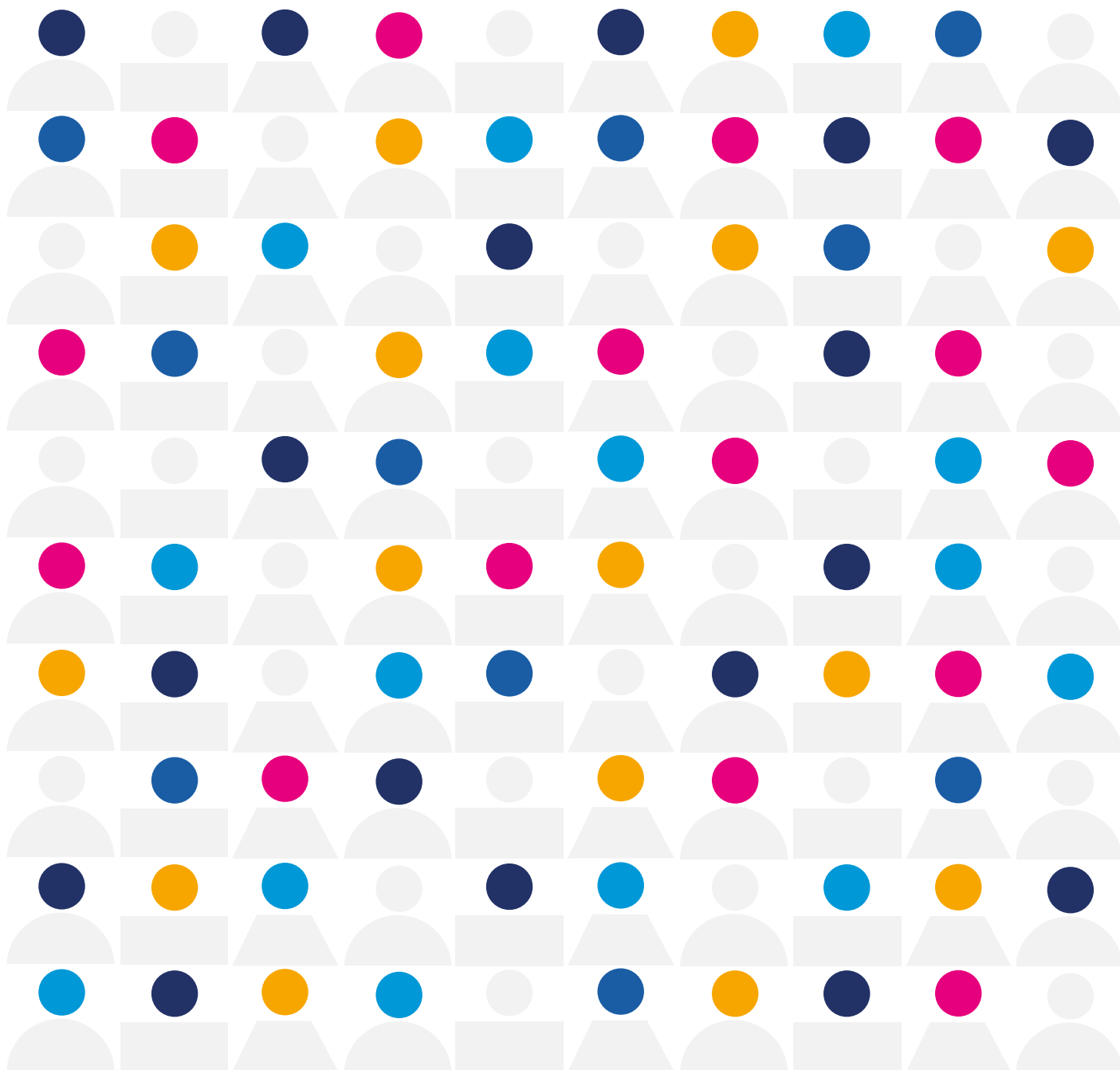
Este diccionario permite conocer las etiquetas asociadas a los códigos contenidos en las variables *región*, *provincia* y *comuna*. Se puede descargar en: <https://censo2024.ine.gob.cl/resultados/>

6.6 Diccionario de códigos territoriales específicos

Este diccionario permite conocer las etiquetas asociadas a los códigos contenidos en las variables *p24_lug_resid5_esp*, *p25_lug_nacimiento_esp*, *p27_nacionalidad_esp* y *p44_lug_trab_esp*. Se puede descargar en: <https://censo2024.ine.gob.cl/resultados/>

7.

Definición y construcción de indicadores



Con el objetivo de facilitar el análisis que las y los usuarios realicen con la base de microdatos del Censo a continuación se presenta una serie de indicadores y su forma de cálculo.

Cada indicador se presenta mediante una ficha técnica que incluye la siguiente información:

- Nombre del indicador
- Definición
- Forma de cálculo
- Universo
- Variables requeridas
- Programación en R

7.1. Aspectos técnicos generales

7.1.1. Redondeo de decimales

En las publicaciones del Censo 2024 los valores se reportan con un decimal. El redondeo se realizó mediante la función `round_half_up` del paquete `janitor`.

7.1.2. Manejo de la no respuesta al ítem

Para el cálculo de porcentajes, los valores de no respuesta (-99 o equivalentes) se excluyeron de los denominadores.

7.1.3. Cargar paquetes

Los indicadores que se presentan en este documento se construyen sobre la base de tres paquetes de R:

- `tidyverse`: transformación de datos y cálculo de indicadores.
- `arrow`: procesamiento eficiente de grandes volúmenes de datos.
- `janitor`: redondeo y conteo de totales.

En este sentido, antes de calcular los indicadores es necesario verificar que estos paquetes estén instalados, para luego cargarlos a la sesión actual a través del comando `library`.

```
library(tidyverse)
library(arrow)
library(janitor)
```

7.1.4. Tipos de variables y carga de datos

Todas las variables son de tipo entero, con excepción de la variable `caenes` que es de tipo cadena de caracteres.

7.1.4.1. Cargar datos

Se recomienda utilizar la función `open_dataset` del paquete `arrow` para la lectura de los datos. Para que la lectura sea más eficiente, se recomienda utilizar los datos en formato *parquet*.

```
personas_censo2024 <- open_dataset("ruta/archivo/personas_censo2024.parquet")
hogares_censo2024 <- open_dataset("ruta/archivo/hogares_censo2024.parquet")
viviendas_censo2024 <- open_dataset("ruta/archivo/viviendas_censo2024.parquet")
```

7.1.5. Desagregaciones temáticas y geográficas

Las desagregaciones de los indicadores se pueden realizar con las variables disponibles en las bases de datos de nivel persona, hogar y vivienda. Por su parte, las desagregaciones geográficas se pueden realizar a nivel regional, provincial, comunal y por área urbano o rural, según lo detallado en la sección de **Control de divulgación estadística**.

7.2. Ficha de indicadores

7.2.1. dem01: Estructura sexo-edad (para pirámide poblacional)

- **Definición:** Es la representación gráfica, mediante barras horizontales, de la distribución de la población por sexo y edad. En el eje horizontal se presentan los valores absolutos o las proporciones de la población, mientras que en el eje vertical se presentan las edades de la población, tradicionalmente por edad quinquenal o por edad simple. Por convención internacional, las barras de la izquierda de la pirámide corresponden a hombres y las de la derecha a mujeres.
- **Forma de cálculo:** Tabla de doble entrada entre edad quinquenal y sexo.
- **Universo:** Total de personas.
- **Variables requeridas:** edad_quinquenal y sexo.
- **Programación en R:**

```
personas_censo2024 %>%
  count(edad_quinquenal, sexo) %>%
  collect() %>%
  mutate(edad_quinquenal = factor(edad_quinquenal,
                                  levels = c(0, 5, 10, 15, 20, 25, 30, 35, 40, 45,
                                              50, 55, 60, 65, 70, 75, 80, 85),
                                  labels = c('0-4', '5-9', '10-14', '15-19',
                                              '20-24', '25-29', '30-34', '35-39',
                                              '40-44', '45-49', '50-54', '55-59',
                                              '60-64', '65-69', '70-74', '75-79',
                                              '80-84', '85 o más')),

          # Se etiqueta variable sexo
          sexo_rec = if_else(sexo == 1, 'Hombre', 'Mujer'),
          # Frec. de hombres en negativo para que queden a la izquierda del eje X
          n = if_else(sexo == 1, -n, n)) %>%
  ggplot(aes(edad_quinquenal, y = n, fill = sexo_rec)) +
  geom_col() +
  coord_flip() +
  scale_y_continuous(labels = abs) +
  labs(x = "Edad en tramos de edad (quinquenios)", y = "Población", fill = "Sexo")
```

7.2.2. dem02: Razón hombre-mujer (índice de masculinidad)

- **Definición:** Corresponde a la relación entre la población de hombres por cada 100 mujeres.
- **Forma de cálculo:**

$$\frac{\text{Total de hombres}}{\text{Total de mujeres}} \times 100$$

- **Universo:** Total de personas.
- **Variables requeridas:** sexo.
- **Programación en R:**

```
personas_censo2024 %>%
  summarise(
    n_hombre = sum(sexo == 1),
    n_mujer = sum(sexo == 2),
    razon_hombre_mujer = n_hombre / n_mujer * 100) %>%
  collect() %>%
  mutate(razon_hombre_mujer = round_half_up(razon_hombre_mujer, 1))
```

7.2.3. dem03: Índice de envejecimiento

- **Definición:** Corresponde a la relación entre la población de 65 años o más por cada 100 personas de 14 años o menos.
- **Forma de cálculo:**

$$\frac{\text{Población de 65 años o más}}{\text{Población entre 0-14 años}} \times 100$$

- **Universo:** Personas de 14 años o menos y de 65 años o más.
- **Variables requeridas:** edad_quinquenal.
- **Programación en R:**

```
personas_censo2024 %>%
  summarise(
    n_pob65 = sum(edad_quinquenal >= 65, na.rm = TRUE),
    n_pob14 = sum(edad_quinquenal < 15, na.rm = TRUE),
    ind_envejecimiento = n_pob65 / n_pob14 * 100
  ) %>%
  collect() %>%
  mutate(ind_envejecimiento = round_half_up(ind_envejecimiento, 1))
```

7.2.4. dem04: Relación de dependencia total

- **Definición:** Corresponde a la relación entre la población potencialmente dependiente (0-14 años y 65 o más) por cada 100 personas potencialmente activas (entre 15 y 64 años).
- **Forma de cálculo:**

$$\frac{\text{Población entre 0-14 años} + \text{Población de 65 años o más}}{\text{Población entre 15 y 64 años}} \times 100$$

- **Universo:** Total de personas.
- **Variables requeridas:** edad_quinquenal.

■ Programación en R:

```
personas_censo2024 %>%
  summarise(
    n_pob65 = sum(edad_quinquenal >= 65, na.rm = TRUE),
    n_pob15_64 = sum(edad_quinquenal >= 15 & edad_quinquenal < 65, na.rm = TRUE),
    n_pob14 = sum(edad_quinquenal < 15, na.rm = TRUE),
    ind_dep_total = (n_pob65 + n_pob14) / n_pob15_64 * 100
  ) %>%
  collect() %>%
  mutate(ind_dep_total = round_half_up(ind_dep_total, 1))
```

7.2.5. dem05: Relación de dependencia de niños/as y jóvenes

- **Definición:** Corresponde a la relación entre la población potencialmente dependiente de 14 años o menos por cada 100 personas potencialmente activas (entre 15 y 64 años).

- **Forma de cálculo:**

$$\frac{\text{Población entre 0-14 años}}{\text{Población entre 15 y 64 años}} \times 100$$

- **Universo:** Personas de 64 años o menos.
- **Variables requeridas:** edad_quinquenal.
- **Programación en R:**

```
personas_censo2024 %>%
  summarise(
    n_pob15_64 = sum(edad_quinquenal >= 15 & edad_quinquenal < 65, na.rm = TRUE),
    n_pob14 = sum(edad_quinquenal < 15, na.rm = TRUE),
    ind_dep_niños = n_pob14 / n_pob15_64 * 100
  ) %>%
  collect() %>%
  mutate(ind_dep_niños = round_half_up(ind_dep_niños, 1))
```

7.2.6. dem06: Relación de dependencia de personas mayores

- **Definición:** Corresponde a la relación entre la población potencialmente dependientes de 65 años o más por cada 100 personas potencialmente activas (entre 15 y 64 años).

- **Forma de cálculo:**

$$\frac{\text{Población de 65 años o más}}{\text{Población entre 15 y 64 años}} \times 100$$

- **Universo:** Personas de 15 años o más.
- **Variables requeridas:** edad_quinquenal.

■ Programación en R:

```
personas_censo2024 %>%
  summarise(
    n_pob65 = sum(edad_quinquenal >= 65, na.rm = TRUE),
    n_pob15_64 = sum(edad_quinquenal >= 15 & edad_quinquenal < 65, na.rm = TRUE),
    ind_dep_mayores = n_pob65 / n_pob15_64 * 100
  ) %>%
  collect() %>%
  mutate(ind_dep_mayores = round_half_up(ind_dep_mayores, 1))
```

7.2.7. dem07: Porcentaje de hogares de personas mayores

- **Definición:** Porcentaje de hogares que están compuestos exclusivamente por personas de 65 años o más.

- **Forma de cálculo:**

$$\frac{\text{Hogares conformados exclusivamente por personas de 65 años o más}}{\text{Total de hogares}} \times 100$$

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `id_vivienda`, `id_hogar`, `tipo_operativo`, `edad_quinquenal`.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  group_by(id_vivienda, id_hogar) %>%
  summarise(
    hog_mayores = all(edad_quinquenal >= 65), .groups = "drop") %>%
  summarise(
    total_hogares = n(),
    hog_mayores = sum(hog_mayores, na.rm = TRUE),
    pct_hogares_mayores = hog_mayores / total_hogares * 100) %>%
  collect() %>%
  mutate(pct_hogares_mayores = round_half_up(pct_hogares_mayores, 1))
```

7.2.8. dem08: Porcentaje de hogares con niños/as

- **Definición:** Porcentaje de hogares con personas de 14 años o menos.

- **Forma de cálculo:**

$$\frac{\text{Hogares con al menos una persona entre 0 a 14 años}}{\text{Total de hogares}} \times 100$$

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `id_vivienda`, `id_hogar`, `tipo_operativo`, `edad_quinquenal`.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  group_by(id_vivienda, id_hogar) %>%
  summarise(
    hog_con_niños = any(edad_quinquenal < 15), .groups = "drop") %>%
  collect() %>%
  summarise(
    total_hogares = n(),
    hog_con_niños = sum(hog_con_niños, na.rm = TRUE),
    pct_hog_con_niños = hog_con_niños / total_hogares * 100) %>%
  mutate(pct_hog_con_niños = round_half_up(pct_hog_con_niños, 1))
```

7.2.9. dem09: Porcentaje de hogares unipersonales

- **Definición:** Porcentaje de hogares conformados únicamente por una persona.

- **Forma de cálculo:**

$$\frac{\text{Hogares conformados por una persona}}{\text{Total de hogares}} \times 100$$

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `id_vivienda`, `id_hogar`, `tipo_operativo`.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  group_by(id_vivienda, id_hogar) %>%
  summarise(
    tamaño_hogar = n(), .groups = "drop") %>%
  collect() %>%
  summarise(
    total_hogares = n(),
    hog_unipersonal = sum(tamaño_hogar == 1, na.rm = TRUE),
    pct_hog_unipersonal = hog_unipersonal / total_hogares * 100) %>%
  mutate(pct_hog_unipersonal = round_half_up(pct_hog_unipersonal, 1))
```

7.2.10. dem10: Promedio de personas por hogar

- **Definición:** Promedio de la cantidad de personas que conforman un hogar.

- **Forma de cálculo:**

$$\frac{\text{Total de personas censadas en viviendas particulares}}{\text{Total de hogares}} \times 100$$

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `id_vivienda`, `id_hogar`, `tipo_operativo`.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  group_by(id_vivienda, id_hogar) %>%
  summarise(cant_per_hog = n(), .groups = "drop") %>%
  collect() %>%
  summarise(prom_personas_hogar = mean(cant_per_hog)) %>%
  mutate(prom_personas_hogar = round_half_up(prom_personas_hogar, 1))
```

7.2.11. dem11: Mediana de edad

- **Definición:** Mediana de edad de la población.
- **Forma de cálculo:** Se debe ubicar el valor de la mitad de la distribución al ordenar la edad de menor a mayor.
- **Universo:** Total de personas.
- **Variables requeridas:** `edad`.
- **Programación en R:**

Por motivos de control de divulgación, la variable `edad` no se encuentra disponible para todas las comunas, por lo que, el ejemplo de programación se realizará para una desagregación territorial que tenga toda su población con información en `edad`. En este caso, para la Región Metropolitana.

Al calcular la mediana, el paquete `arrow` utiliza un algoritmo que permite obtener una aproximación de la mediana de forma eficiente, en vez de calcular la mediana exacta. Por esto, se recomienda derivar el cálculo a `DuckDB`, mediante `arrow::to_duckdb()`, lo que permite calcular la mediana de forma exacta y eficiente.

```
personas_censo2024 %>%
  filter(region == 13) %>%
  arrow::to_duckdb() %>%
  summarise(mediana_edad = median(edad)) %>%
  collect()
```


7.2.12. dem12: Porcentaje de hogares con jefatura de hogar mujer

- **Definición:** Porcentaje de hogares en que su jefatura de hogar es de sexo mujer.
- **Forma de cálculo:**

$$\frac{\text{Hogares con jefatura de hogar mujer}}{\text{Total de hogares}} \times 100$$

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `parentesco`, `sexo` y `tipo_operativo`.
- **Programación en R:**

Considerando que cada hogar cuenta con una única jefatura de hogar, se debe filtrar por `parentesco == 1` para obtener el total de hogares. Luego, se debe calcular el total de mujeres (que sería el equivalente a hogares con jefatura de hogar mujer) y dividirlo por el total de hogares (que sería el equivalente a la suma entre hogares con jefatura de hogar mujer y hombre).

```
personas_censo2024 %>%
  filter(parentesco == 1 & tipo_operativo == 2) %>%
  summarise(n_jh_hombre = sum(sexo == 1),
            n_jh_mujer = sum(sexo == 2),
            porc_jh_mujer = n_jh_mujer / (n_jh_hombre + n_jh_mujer) * 100
  ) %>%
  collect() %>%
  mutate(porc_jh_mujer = round_half_up(porc_jh_mujer, 1))
```

7.2.13. mig01: Porcentaje de población inmigrante internacional

- **Definición:** Porcentaje de personas nacidas fuera del país.
- **Forma de cálculo:**

$$\frac{\text{Población nacida fuera del país}}{\text{Total de población (excluye no respuesta en p25_lug_nacimiento)}} \times 100$$

- **Universo:** Total de personas.
- **Variables requeridas:** `p25_lug_nacimiento`.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(p25_lug_nacimiento != -99) %>% # Excluye no respuesta
  summarise(pob_inmigrante_int = sum(p25_lug_nacimiento == 3, na.rm = TRUE),
            pct_inmigrante_int = pob_inmigrante_int / n() * 100
  ) %>%
  collect() %>%
  mutate(pct_inmigrante_int = round_half_up(pct_inmigrante_int, 1))
```

7.2.14. fec01: Razón niño-mujer

- **Definición:** Relación entre la cantidad de personas de 4 años o menos por cada cien mujeres entre 15-49 años.

- **Forma de cálculo:**

$$\frac{\text{Población de 4 años o menos}}{\text{Mujeres entre 15 y 49 años}} \times 100$$

- **Universo:** Población entre 0 y 4 años y mujeres entre 15 y 49 años.

- **Variables requeridas:** edad_quinquenal y sexo.

- **Programación en R:**

```
personas_censo2024 %>%
  summarise(pob04 = sum(edad_quinquenal < 5, na.rm = TRUE),
            pob_muj_1549 = sum(edad_quinquenal >= 15 & edad_quinquenal <= 49
                               & sexo == 2, na.rm = TRUE),
            raz_niño_mujer = pob04 / pob_muj_1549 * 100)
  ) %>%
  collect() %>%
  mutate(raz_niño_mujer = round_half_up(raz_niño_mujer, 1))
```

7.2.15. fec02: Promedio de hijos nacidos vivos de mujeres entre 15-49 años (paridez media)

- **Definición:** Indicador conocido bajo el nombre de paridez media. Corresponde al promedio de hija/os nacidos que han tenido las mujeres entre 15 y 49 años.

- **Forma de cálculo:**

$$\frac{\text{Total de hijos nacidos vivos de mujeres entre 15 y 49 años}}{\text{Mujeres entre 15 y 49 años (excluye no respuesta en p46a_tot_hijs_nac)}} \times 100$$

- **Universo:** Mujeres entre 15 y 49 años.

- **Variables requeridas:** p46a_tot_hijs_nac, sexo y edad_quinquenal.

- **Programación en R:**

```
personas_censo2024 %>%
  # edad_quinquenal 45 representa grupo 45-49 años
  filter(edad_quinquenal >= 15 & edad_quinquenal <= 45 & sexo == 2 &
         p46a_tot_hijs_nac != -99) %>%
  summarise(paridez_media = mean(p46a_tot_hijs_nac, na.rm = TRUE)) %>%
  collect() %>%
  mutate(paridez_media = round_half_up(paridez_media, 1))
```

7.2.16. fec03: Porcentaje de mujeres entre 15-49 años con hijos

- **Definición:** Porcentaje de mujeres entre 15 y 49 años que registran tener al menos un hijo/a nacido vivo.
- **Forma de cálculo:**

$$\frac{\text{Total de mujeres entre 15 a 49 años con 1 hijo o más}}{\text{Mujeres entre 15 y 49 años (excluye mujeres con no respuesta en p46a_tot_hijs_nac)}} \times 100$$

- **Universo:** Mujeres entre 15 y 49 años.
- **Variables requeridas:** p46a_tot_hijs_nac, sexo y edad_quinquenal.
- **Programación en R:**

```
personas_censo2024 %>%
  # edad_quinquenal 45 representa grupo 45-49 años
  filter(edad_quinquenal >= 15 & edad_quinquenal <= 45 & sexo == 2 &
    p46a_tot_hijs_nac != -99) %>% # Excluye no respuesta
  summarise(pob_muj_con_hijs = sum(edad_quinquenal >= 15 & edad_quinquenal <= 45 &
    sexo == 2 &
    p46a_tot_hijs_nac > 0, na.rm = TRUE),
    pob_muj_1549 = sum(edad_quinquenal >= 15 & edad_quinquenal <= 45 &
    sexo == 2,
    na.rm = TRUE),
    pct_mjr_con_hijs = pob_muj_con_hijs / pob_muj_1549 * 100
  ) %>%
  collect() %>%
  mutate(pct_mjr_con_hijs = round_half_up(pct_mjr_con_hijs, 1))
```

7.2.17. fec04: Porcentaje de mujeres entre 15-49 años sin hijos

- **Definición:** Porcentaje de mujeres entre 15 y 49 años que no registran hijos/as nacidos vivos.
- **Forma de cálculo:**

$$\frac{\text{Mujeres entre 15 y 49 años sin hijos/as nacidos vivos}}{\text{Mujeres entre 15 y 49 años (excluye mujeres con no respuesta en p46a_tot_hijs_nac)}} \times 100$$

- **Universo:** Mujeres entre 15 y 49 años.
- **Variables requeridas:** p46a_tot_hijs_nac, sexo y edad_quinquenal.

■ Programación en R:

```
personas_censo2024 %>%
  # edad_quinquenal 45 representa grupo 45-49 años
  filter(edad_quinquenal >= 15 & edad_quinquenal <= 45 & sexo == 2 &
    p46a_tot_hijs_nac != -99) %>% # Excluye no respuesta
  summarise(pob_muj_sin_hijs = sum(edad_quinquenal >= 15 &
    edad_quinquenal <= 45 &
    sexo == 2 &
    p46a_tot_hijs_nac == 0, na.rm = TRUE),
    pob_muj_1549 = sum(edad_quinquenal >= 15 &
    edad_quinquenal <= 45 &
    sexo == 2, na.rm = TRUE),
    pct_mjr_sin_hijs = pob_muj_sin_hijs / pob_muj_1549 * 100) %>%
  collect() %>%
  mutate(pct_mjr_sin_hijs = round_half_up(pct_mjr_sin_hijs, 1))
```

7.2.18. mig_int01: Matriz de migración interna (nivel regional)

- **Definición:** Tabla de doble entrada que compara la región de residencia habitual en dos períodos de tiempo. En el caso de la matriz de nivel regional, en las columnas se distribuye la población según la región en que residía en abril de 2019, y de manera análoga, en las filas se distribuye la misma población según la región de residencia al momento del Censo 2024. En la diagonal de la matriz se encuentra la población no migrante, es decir, aquellas personas que declararon residir en la misma región en ambas fechas. En los totales de columnas y filas se observa la población residente por región en 2019 y 2024, respectivamente.
- **Forma de cálculo:** Para la construcción de la matriz de migración interna se sugiere seguir los siguientes pasos:
 1. Crear una variable que tenga la región de residencia hace cinco años (`region_resid5`). La función de esta variable es transformar la información de comuna de `p24_lug_resid5_esp` a nivel regional y mantener la información de región en este nivel. Esto se puede realizar dividiendo por mil y truncando la información geográfica de aquellas personas con respuesta “En esta comuna” y “En otra comuna” de `p24_lug_resid5`¹.

Para las personas que aún no nacían hace cinco años, residían en otro país, no respondieron la pregunta 24 o respondieron que residían en otra comuna, pero no especificaron el nombre de la comuna o de la región, la variable `region_resid5` debe tomar valores NA en tanto estas personas no entran en el análisis de migración interna.

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(region_resid5 = case_when(
    p24_lug_resid5 == 1 ~ NA_integer_, # Personas que hace cinco años aún no nacían
    p24_lug_resid5 == 4 ~ NA_integer_, # Personas que residían en otro país
    p24_lug_resid5 == -99 ~ NA_integer_, # Personas sin respuesta en p24_lug_resid5
    # Personas sin respuesta en comuna o región
    p24_lug_resid5 == 3 & p24_lug_resid5_esp == -99 ~ NA_integer_,
    p24_lug_resid5 %in% 2:3 ~ trunc(p24_lug_resid5_esp/1000)
  )) %>%
  compute()
```

¹ Esto se puede hacer porque en el Código Único Territorial (CUT) los primeros 2 dígitos representan la región (incluyendo un cero en las comunas cuyo número de región es menor a 10) y los últimos 3 a la provincia y comuna. Al dividir por mil y truncar, se “eliminan” los últimos 3 dígitos.

2. Calcular la frecuencia del cruce entre `region` y `region_resid5` usando la función `count`. Se sugiere cambiar el formato de la variable a `factor` para que sea posible etiquetar la información de región y ordenar la tabla a nivel de filas y columnas mediante la función `arrange`.

3. Transformar la tabla resultante a formato ancho.

4. Contabilizar los totales por fila y por columna.

■ **Universo:** Personas de 5 años o más.

■ **Variables requeridas:** `p24_lug_resid5`, `p24_lug_resid5_esp`, `region`.

■ **Programación en R:**

```
personas_censo2024 %>%
  # Se considera sólo la población con información en región hace cinco años
  filter(region_resid5 %in% 1:16) %>%
  count(region_resid5, region) %>%
  collect() %>%
  mutate(region = factor(region,
    levels = c(1:16),
    labels = c("Tarapacá", "Antofagasta", "Atacama", "Coquimbo",
      "Valparaíso", "OHiggins", "Maule", "Biobío", "La Araucanía",
      "Los Lagos", "Aysén", "Magallanes", "Metropolitana",
      "Los Ríos", "Arica y Parinacota", "Ñuble")),
    region_resid5 = factor(region_resid5,
    levels = c(1:16),
    labels = c("Tarapacá", "Antofagasta", "Atacama", "Coquimbo",
      "Valparaíso", "OHiggins", "Maule", "Biobío", "La Araucanía",
      "Los Lagos", "Aysén", "Magallanes", "Metropolitana",
      "Los Ríos", "Arica y Parinacota", "Ñuble"))) %>%
  arrange(region_resid5, region) %>%
  pivot_wider(names_from = "region_resid5",
    values_from = "n") %>%
  adorn_totals(where = "col") %>%
  adorn_totals(where = "row")
```

7.2.19. mig_int02: Indicadores derivados de la matriz de migración interna

Definición: A partir de la matriz de migración interna se pueden calcular una serie de indicadores derivados. Estos permiten identificar las regiones que poseen tendencias migratorias “expulsoras” o “atractivas” de población. Al igual que la matriz de migración interna, se descartan del análisis las personas que aún no nacían, las personas que hace cinco años residían en “Otro país”, las personas sin respuesta en la pregunta de residencia hace cinco años y aquellas que, habiendo respondido que hace cinco años vivían en otra comuna, dejaron sin respuesta la pregunta de comuna o región.

Los indicadores que se pueden calcular son los siguientes:

■ **Residencia habitual actual (regional):** La residencia habitual actual corresponde a la región donde una persona ha vivido la mayor parte de los últimos 12 últimos meses previos al censo (es decir, al menos seis meses y un día), o donde habiendo vivido menos tiempo, tenía la intención de seguir residiendo allí. Refiere al lugar donde fue censada la persona.

- **Residencia habitual hace cinco años (regional):** Corresponde a la región donde la persona censada residía habitualmente hace cinco años antes del censo, es decir, la región registrada en la variable `region_resid5`.
- **Población media (regional):** Promedio entre la cantidad de residentes habituales en la actualidad y residentes habituales hace cinco años en la región. Corresponde al denominador para el cálculo de las tasas.
- **No migrantes (regional):** Corresponde a las personas residentes habituales de la región que hace cinco años residían en el mismo lugar que al momento del Censo 2024.
- **Inmigrantes (regional):** Corresponde a las personas que en 2019 residían habitualmente en una región distinta a donde lo hacen en 2024. Se deben contabilizar en función del lugar de residencia habitual actual (2024).
- **Emigrantes (regional):** Corresponde a las personas que en 2024 residen habitualmente en una región distinta a donde lo hacían en 2019. Se deben contabilizar en función del lugar de residencia anterior (2019).
- **Saldo de migración interna neta (regional):** Representa la diferencia entre el número absoluto de inmigrantes y el de emigrantes de una región. La migración interna neta puede ser positiva (número de inmigrantes mayor que el número de emigrantes), negativa (número de inmigrantes menor que el número de emigrantes), o cero (número de inmigrantes igual que el número de emigrantes).
- **Tasa de inmigración interna (regional):** Expresa el número de inmigrantes dividido por cinco (años que comprende el período 2019-2024), cuyo resultado es dividido por la población media de la región durante el período de referencia y multiplicado por 1000. La cifra final se expresa por cada mil habitantes.
- **Tasa de emigración interna (regional):** Expresa el número de emigrantes dividido por cinco, cuyo resultado es dividido por la población media de la región durante el período de referencia y multiplicado por mil. La cifra final se expresa por cada mil habitantes.
- **Tasa de migración interna neta (regional):** Expresa el saldo de migración interna neta dividido por cinco, cuyo resultado es dividido por la población media de la región durante el período de referencia y multiplicado por mil. La cifra final se expresa por cada mil habitantes y se interpreta como el aporte que la migración hace en un período determinado al cambio de población. Si es negativa, implica una sustracción neta de personas. Si es positiva, implica una adición neta de personas. Si es cero, implica un equilibrio entre llegada y salida de residentes habituales.

Forma de cálculo:

- **Residencia habitual actual (regional):** Se debe realizar un conteo de `region`. Para el cálculo se deben excluir a las personas que aún no nacían hace cinco años, a las personas que hace cinco años residían en otro país, ignorados de `p24_lug_resid5` (-99) e ignorados de `p24_lug_resid5_esp`.
- **Residencia habitual hace cinco años (regional):** Se debe realizar un conteo de `region_resid5`. Para el cálculo se deben excluir a las personas que aún no nacían hace cinco años, a las personas que hace cinco años residían en otro país, ignorados de `p24_lug_resid5` (-99) e ignorados de `p24_lug_resid5_esp`.

■ **Población media (regional):**

Residentes habituales actuales de la región + residentes habituales hace cinco años de la región

2

- **No migrantes (regional):** Se debe realizar un conteo de `region`, realizando previamente un filtro de aquellas personas que registran el mismo valor entre `region` y `region_resid5`.
- **Inmigrantes (regional):** Se debe realizar un conteo de `region`, realizando previamente un filtro de aquellas personas que registran un valor distinto entre `region` y `region_resid5`.
- **Emigrantes (regional):** Se debe realizar un conteo de `region_resid5`, realizando previamente un filtro de aquellas personas que registran un valor distinto entre `region` y `region_resid5`.
- **Saldo de migración neta (regional):** Una vez obtenidos los inmigrantes y emigrantes, se debe realizar una resta simple entre estos indicadores a nivel de región.

■ **Tasa de inmigración (regional):**

$$\frac{\text{Inmigrantes regionales} \div 5}{\text{Población media regional}} \times 1000$$

■ **Tasa de emigración (regional):**

$$\frac{\text{Emigrantes regionales} \div 5}{\text{Población media regional}} \times 1000$$

■ **Tasa de migración neta (regional):**

$$\frac{\text{Saldo de migración neta regional} \div 5}{\text{Población media regional}} \times 1000$$

- **Universo:** Personas de 5 años o más.

- **Variables requeridas:** `p24_lug_resid5` y `p24_lug_resid5_esp`.

- **Programación en R:** Para la programación de los indicadores se utiliza la variable creada para generar la matriz de migración interna `region_resid5`. Se sugiere utilizar el filtro `region_resid5 %in% 1:16` para conservar únicamente los registros con información válida en región de residencia hace cinco años.

Primero se calculan los indicadores que pueden ser calculados mediante un conteo.

```
# Residentes habituales 2024
pob2024 <- personas_censo2024 %>%
  filter(region_resid5 %in% 1:16) %>%
  count(region, name = "pob_2024") %>%
  collect()

# Residentes habituales 2019
pob2019 <- personas_censo2024 %>%
  filter(region_resid5 %in% 1:16) %>%
  count(region_resid5, name = "pob_2019") %>%
  # Se modifica el nombre para usar de llave de unión al usar función left_join
  rename(region = region_resid5) %>%
  collect()

# No migrantes internos
no_migrantes <- personas_censo2024 %>%
  filter(region_resid5 %in% 1:16) %>%
  group_by(region) %>%
  filter(region_resid5 == region) %>%
  count(name = "no_migrantes") %>%
  collect()

# Inmigrantes internos
inmigrantes <- personas_censo2024 %>%
  filter(region_resid5 %in% 1:16) %>%
  group_by(region) %>%
  filter(region_resid5 != region) %>%
  count(name = "inmigrantes") %>%
  collect()

# Emigrantes internos
emigrantes <- personas_censo2024 %>%
  filter(region_resid5 %in% 1:16) %>%
  group_by(region_resid5) %>%
  filter(region_resid5 != region) %>%
  count(name = "emigrantes") %>%
  # Se modifica el nombre para usar de llave de unión al usar función left_join
  rename(region = region_resid5) %>%
  collect()
```

Luego, se realiza un `left_join` de los indicadores a través de la columna `region` y se calculan los indicadores restantes, que son operaciones que se realizan a partir de los indicadores calculados anteriormente.


```
# Tabla de indicadores derivados
pob2024 %>%
  left_join(pob2019, by = c("region")) %>%
  left_join(no_migrantes, by = c("region")) %>%
  left_join(inmigrantes, by = c("region")) %>%
  left_join(emigrantes, by = c("region")) %>%
  arrange(region) %>%
  mutate(region = factor(region,
                        levels = c(1:16),
                        labels = c("Tarapacá", "Antofagasta", "Atacama", "Coquimbo",
                                   "Valparaíso", "OHiggins", "Maule", "Biobío", "La Araucanía",
                                   "Los Lagos", "Aysén", "Magallanes", "Metropolitana",
                                   "Los Ríos", "Arica y Parinacota", "Ñuble"))) %>%
  mutate(saldo = inmigrantes - emigrantes, # Saldo migratorio
         pob_media = (pob_2024 + pob_2019)/2) %>% # Población media
  # Se calculan los totales nacionales para los indicadores expresados en valores
  # absolutos
  adorn_totals(where = "row") %>%
  mutate(
    # Tasa de inmigración
    tasa_inmigracion = round_half_up((inmigrantes/5)/pob_media*1000, 1),
    # Tasa de emigración
    tasa_emigracion = round_half_up((emigrantes/5)/pob_media*1000, 1),
    # Tasa de migración neta
    tasa_migracion_neta = round_half_up((saldo/5)/pob_media*1000, 1))
```

7.2.20. mig_int03: Condición migratoria (nivel regional)

- **Definición:** Variable que permite clasificar a la población en las categorías de no migrantes y migrantes. A su vez, las personas migrantes se pueden dividir en dos categorías: migrantes inter-regionales o intra-regionales, sin embargo, en este documento sólo se abordarán los movimientos entre regiones².
 - **No migrantes (regional):** Corresponde a las personas residentes habituales de la región que hace cinco años residían en la misma región que al momento del Censo 2024.
 - **Migrantes inter-regionales:** Corresponde a las personas residentes habituales de la región que hace cinco años residían una región distinta a la que residían al momento del Censo 2024.
- **Forma de cálculo:** Se utiliza la variable construida en el indicador mig_int01 (`region_resid5`) y se compara con la región de residencia habitual (`region`). Las personas que tengan igualdad entre ambas variables se consideran no migrantes; las personas que tengan desigualdad entre ambas variables, se consideran migrantes inter-regionales.
- **Universo:** Personas de 5 años o más.
- **Variables requeridas:** `region_resid5`.

² Para cuantificar a los migrantes intra-regionales, es decir, a aquellas personas que cambiaron su lugar de residencia dentro de la misma región, se debe analizar la matriz de migración al interior de la región usando una matriz a nivel de una división administrativa menor, como comuna.

- **Programación en R:** Para la programación de los indicadores se utiliza la variable creada para generar la matriz de migración interna (region_resid5).

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(cond_migratoria = case_when(
    region == region_resid5 ~ "No migrante",
    region != region_resid5 ~ "Migrante inter-regional",
    TRUE ~ NA_character_
  )) %>%
  compute()

personas_censo2024 %>%
  filter(!is.na(cond_migratoria)) %>%
  summarise(pct_mig_ir = sum(cond_migratoria == "Migrante inter-regional",
    na.rm = TRUE) / n() * 100,
    pct_no_mig = sum(cond_migratoria == "No migrante",
    na.rm = TRUE) / n() * 100) %>%
  collect() %>%
  mutate(pct_mig_ir = round_half_up(pct_mig_ir, 1),
    pct_no_mig = round_half_up(pct_no_mig, 1))
```

7.2.21. educ01: Promedio de años de escolaridad

- **Definición:** Promedio de años de educación formalmente completados por la población residente habitual del país.

- **Forma de cálculo:**

$$\frac{\text{Sumatoria de años de escolaridad de la población}}{\text{Población total (excluye no respuesta en escolaridad)}} \times 100$$

- **Universo:** Toda la población.
- **Variables requeridas:** escolaridad.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(escolaridad != -99) %>% # Excluye no respuesta
  summarise(prom_escolaridad = mean(escolaridad)) %>%
  collect() %>%
  mutate(prom_escolaridad = round_half_up(prom_escolaridad, 1))
```

7.2.22. educ02: Tasa de alfabetización

- **Definición:** Proporción de personas de 15 años o más que sabe leer y escribir.
- **Forma de cálculo:**

$$\frac{\text{Población que sabe leer y escribir}}{\text{Población de 15 años o más (excluye no respuesta p37_alfabet)}} \times 100$$

- **Universo:** Población de 15 años o más.
- **Variables requeridas:** `p37_alfabet` y `edad_quinquenal`.
- **Programación en R:**

```
# Cálculo de tasa de alfabetización
personas_censo2024 %>%
  filter(edad_quinquenal >= 15 &
         p37_alfabet != -99) %>% # Excluye no respuesta
  summarise(sabe_leer_escribir = sum(p37_alfabet == 1),
            tasa_alfabetizacion = sabe_leer_escribir / n() * 100
            ) %>%
  collect() %>%
  mutate(tasa_alfabetizacion = round_half_up(tasa_alfabetizacion, 1))
```

7.2.23. educ03: Tasa de asistencia neta a la educación parvularia

- **Definición:** Cociente de personas entre 0 y 5 años que asisten a educación parvularia respecto a la población de 0 a 5 años con información válida en la variable de asistencia a la educación formal y de nivel educativo, multiplicado por 100.

- **Forma de cálculo:**

Población entre 0 y 5 años que asisten a educación parvularia

X 100

Población entre 0 y 5 años (excluye no respuesta en asistencia a la educación formal y nivel educativo)

- **Universo:** Población de 0 a 5 años.
- **Variables requeridas:** `asistencia_parv`.
- **Programación en R:** Por motivos de control de divulgación, la variable `asistencia_parv` no se encuentra disponible para todas las comunas, por lo que, el ejemplo de programación se realizará para una desagregación territorial que tenga toda su población con información en `asistencia_parv`. En este caso, para la Región Metropolitana.

```
personas_censo2024 %>%
  filter(region == 13 &
         asistencia_parv != -99) %>% # Se excluye no respuesta
  mutate(denominador = ifelse(asistencia_parv %in% 1:2, 1, 0)) %>%
  summarise(tasa_neta_parv = sum(asistencia_parv == 1, na.rm = T) /
            sum(denominador == 1, na.rm = T) * 100) %>%
  collect() %>%
  mutate(tasa_neta_parv = round_half_up(tasa_neta_parv, 1))
```

7.2.24. educ04: Tasa de asistencia neta a la educación básica

- **Definición:** Cociente de personas entre 6 y 13 años que asisten a educación básica respecto a la población de 6 y 13 años con información válida en la variable de asistencia a la educación formal, de nivel educativo y de término de nivel, multiplicado por 100.

- **Forma de cálculo:**

Población entre 6 y 13 años que asisten a educación básica

X 100

Población entre 6 y 13 años
(excluye no respuesta en asistencia a la educación formal, nivel educativo y término de nivel)

- **Universo:** Población de 6 a 13 años.
- **Variables requeridas:** `asistencia_basica`.
- **Programación en R:** Por motivos de control de divulgación, la variable `asistencia_basica` no se encuentra disponible para todas las comunas, por lo que, el ejemplo de programación se realizará para una desagregación territorial que tenga toda su población con información en `asistencia_basica`. En este caso, para la Región Metropolitana.

```
personas_censo2024 %>%
  filter(region == 13 &
         asistencia_basica != -99) %>% # Se excluye no respuesta
  mutate(denominador = ifelse(asistencia_basica %in% 1:2, 1, 0)) %>%
  summarise(tasa_neta_basica = sum(asistencia_basica == 1, na.rm = T) /
            sum(denominador == 1, na.rm = T) * 100) %>%
  collect() %>%
  mutate(tasa_neta_basica = round_half_up(tasa_neta_basica, 1))
```

7.2.25. educ05: Tasa de asistencia neta a la educación media

- **Definición:** Cociente de personas entre 14 y 17 años que asisten a educación media respecto a la población de 14 y 17 años con información válida en la variable de asistencia a la educación formal, de nivel educativo y de término de nivel, multiplicado por 100.

- **Forma de cálculo:**

Población entre 14 y 17 años que asisten a educación media

X 100

Población entre 14 y 17 años
(excluye no respuesta en asistencia a la educación formal, nivel educativo y término de nivel)

- **Universo:** Población de 14 a 17 años.
- **Variables requeridas:** `asistencia_media`.

■ Programación en R:

Por motivos de control de divulgación, la variable `asistencia_media` no se encuentra disponible para todas las comunas, por lo que, el ejemplo de programación se realizará para una desagregación territorial que tenga toda su población con información en `asistencia_media`. En este caso, para la Región Metropolitana.

```
personas_censo2024 %>%
  filter(region == 13 &
         asistencia_media != -99) %>% # Se excluye no respuesta
  mutate(denominador = ifelse(asistencia_media %in% 1:2, 1, 0)) %>%
  summarise(tasa_neta_media = sum(asistencia_media == 1, na.rm = T) /
            sum(denominador == 1, na.rm = T) * 100) %>%
  collect() %>%
  mutate(tasa_neta_media = round_half_up(tasa_neta_media, 1))
```

7.2.26. educ06: Tasa de asistencia neta a la educación superior

- **Definición:** Cociente de personas entre 18 a 24 años que asisten a educación superior respecto a la población de 18 a 24 años con información válida en la variable de asistencia a la educación formal, de nivel educativo y de término de nivel, multiplicado por 100.

- **Forma de cálculo:**

Población entre 18 y 24 años que asisten a educación superior

X 100

Población entre 18 y 24 años

(excluye no respuesta en asistencia a la educación formal, nivel educativo y término de nivel)

- **Universo:** Población de 18 a 24 años.
- **Variables requeridas:** `asistencia_superior`.
- **Programación en R:** Por motivos de control de divulgación, la variable `asistencia_superior` no se encuentra disponible para todas las comunas, por lo que, el ejemplo de programación se realizará para una desagregación territorial que tenga toda su población con información en `asistencia_superior`. En este caso, para la Región Metropolitana.

```
personas_censo2024 %>%
  filter(region == 13 &
         asistencia_superior != -99) %>% # Se excluye no respuesta
  mutate(denominador = ifelse(asistencia_superior %in% 1:2, 1, 0)) %>%
  summarise(tasa_neta_superior = sum(asistencia_superior == 1, na.rm = T) /
            sum(denominador == 1, na.rm = T) * 100) %>%
  collect() %>%
  mutate(tasa_neta_superior = round_half_up(tasa_neta_superior, 1))
```

7.2.27. disc01: Discapacidad

- **Definición:** Clasifica a las personas de 5 años o más que, debido a dificultades para realizar ciertas actividades básicas y universales, se encuentran en mayor riesgo de experimentar restricciones en su participación social en comparación a personas sin estas dificultades. El punto de corte recomendado por el Grupo de Washington sobre Estadísticas de Discapacidad (WG) para el indicador de discapacidad considera clasificar como personas con discapacidad a aquellas que respondieron tener mucha dificultad o no poder realizar al menos una de las seis actividades consultadas. Así, se construye un indicador que clasifica a las personas con y sin discapacidad.

- **Forma de cálculo:**

Para construir la variable `disc` se deben evaluar las siguientes condiciones de manera secuencial:

- Si la persona tiene 0-4 años, se clasifica como no aplica (NA).
- Si en alguna variable de dificultad se registra los valores 3 o 4, se clasifica con discapacidad (1).
- Si en alguna variable de dificultad se registra 1 o 2, se clasifica sin discapacidad (2).
- Las personas restantes se clasifican con la categoría "No responde" (-99).

- **Universo:** Población de 5 años o más.

- **Variables requeridas:** `p32a_dificultad_ver`, `p32b_dificultad_oir`, `p32c_dificultad_mover`, `p32d_dificultad_cogni`, `p32e_dificultad_cuidado`, `p32f_dificultad_comunic`, `edad_quinquenal`.

- **Programación en R:**

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(
    disc = case_when(
      # edad_quinquenal 0 representa grupo 0-4 años
      edad_quinquenal == 0 ~ NA_integer_,

      (p32a_dificultad_ver %in% 3:4) | (p32b_dificultad_oir %in% 3:4) |
      (p32c_dificultad_mover %in% 3:4) | (p32d_dificultad_cogni %in% 3:4) |
      (p32e_dificultad_cuidado %in% 3:4) | (p32f_dificultad_comunic %in% 3:4) ~ 1,

      (p32a_dificultad_ver %in% 1:2) | (p32b_dificultad_oir %in% 1:2) |
      (p32c_dificultad_mover %in% 1:2) | (p32d_dificultad_cogni %in% 1:2) |
      (p32e_dificultad_cuidado %in% 1:2) | (p32f_dificultad_comunic %in% 1:2) ~ 2,

      TRUE ~ -99)) %>%
  compute()
```

7.2.28. pueblos01: Porcentaje de la población que es o se considera perteneciente a un pueblo indígena u originario

- **Definición:** Cantidad de personas que son o se identifican con un pueblo indígena u originario sobre el total de la población.

- **Forma de cálculo:**

$$\frac{\text{Población que es o se considera perteneciente a un pueblo indígena}}{\text{Población total (excluye no respuesta en p28_autoid_pueblo)}} \times 100$$

- **Universo:** Total de personas.
- **Variables requeridas:** p28_autoid_pueblo.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(p28_autoid_pueblo != -99) %>% # Excluye no respuesta
  summarise(pob_pueblos = sum(p28_autoid_pueblo == 1),
            pct_pueblos = pob_pueblos / n() * 100) %>%
  collect() %>%
  mutate(pct_pueblos = round_half_up(pct_pueblos, 1))
```

7.2.29. pueblos02: Porcentaje de la población afrodescendiente

- **Definición:** Cantidad de personas que son o se identifican como afrodescendientes sobre el total de la población.

- **Forma de cálculo:**

$$\frac{\text{Población que es o se considera afrodescendiente}}{\text{Población total (excluye no respuesta en p29_afrodescendencia_rec)}} \times 100$$

- **Universo:** Total de personas.
- **Variables requeridas:** p29_afrodescendencia.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(p29_afrodescendencia_rec != -99) %>% # Excluye no respuesta
  summarise(pob_afro = sum(p29_afrodescendencia_rec == 1),
            pct_afro = pob_afro / n() * 100) %>%
  collect() %>%
  mutate(pct_afro = round_half_up(pct_afro, 1))
```

7.2.30. conm01: Matriz de conmutación laboral

- **Definición:** Tabla de doble entrada que compara la región de residencia habitual y la región de trabajo de la población ocupada. En el caso de la matriz de nivel regional, en las columnas se distribuye la población ocupada según región de trabajo, y de manera análoga, en las filas se distribuye la región de residencia habitual. En la diagonal de la matriz se encuentra la población que reside y trabaja en la misma región.
- **Forma de cálculo:** Para la construcción de la matriz de conmutación laboral se sugiere seguir los siguientes pasos:
 1. Crear una variable que tenga la región del lugar de trabajo (`region_lugtrab`). La función de esta variable es transformar la información de comuna de `p44_lug_trab_esp` a nivel regional y mantener la información de región en este nivel. Esto se puede realizar truncando y dividiendo por mil la información geográfica de aquellas personas con respuesta “En su comuna, en la vivienda donde reside,” “En su comuna, fuera de la vivienda donde reside” y “En otra comuna” (`p44_lug_trab %in% c(1:3)`).

Se dejan como NA en la variable `region_lugtrab`: las personas de 15 años o menos, las personas que su situación laboral es desocupada o fuera de la fuerza de trabajo, que trabajan en otro país, que trabajan en varias comunas o países, las personas que no respondieron la pregunta `p44_lug_trab` y las personas que habiendo respondido `p44_lug_trab` no registraron información de región. Ya sea por su respuesta o por la falta de información estas personas no se incluyen en la matriz de conmutación laboral.

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(region_trabajo = case_when(
    p44_lug_trab %in% 1:3 ~ as.integer(trunc(p44_lug_trab_esp/1000)),
    TRUE ~ NA_integer_
  )) %>%
  compute()
```

2. Calcular la frecuencia del cruce entre `region` y `region_lugtrab` mediante un `count`. Se recomienda cambiar el formato de la variable a `factor` para etiquetar la información de región y ordenar la tabla a nivel de filas y columnas mediante la función `arrange`.
 3. Transformar la tabla resultante a formato ancho.
 4. Contabilizar los totales por fila y columna.
- **Universo:** Población ocupada de 15 años o más.
 - **Variables requeridas:** `p44_lug_trab_esp`, `region_lugtrab`, `region`.

■ Programación en R:

```
personas_censo2024 %>%
  # Se considera sólo la población ocupada con información en lugar de trabajo
  filter(region_trabajo %in% 1:16) %>%
  count(region_trabajo, region) %>%
  collect() %>%
  mutate(region = factor(region,
    levels = c(1:16),
    labels = c("Tarapacá", "Antofagasta", "Atacama", "Coquimbo",
      "Valparaíso", "OHiggins", "Maule", "Biobío", "La Araucanía",
      "Los Lagos", "Aysén", "Magallanes", "Metropolitana",
      "Los Ríos", "Arica y Parinacota", "Ñuble")),
    region_trabajo = factor(region_trabajo,
      levels = c(1:16),
      labels = c("Tarapacá", "Antofagasta", "Atacama", "Coquimbo",
        "Valparaíso", "OHiggins", "Maule", "Biobío", "La Araucanía",
        "Los Lagos", "Aysén", "Magallanes", "Metropolitana",
        "Los Ríos", "Arica y Parinacota", "Ñuble"))) %>%
  arrange(region_trabajo, region) %>%
  pivot_wider(names_from = "region_trabajo",
    values_from = "n") %>%
  adorn_totals(where = "col") %>%
  adorn_totals(where = "row")
```

7.2.31. conm02: Condición de conmutación laboral (nivel regional)

- **Definición:** Variable que permite clasificar a la población en las categorías de no conmutantes y conmutantes. A su vez, las personas que conmutan se pueden dividir en dos categorías: conmutantes inter-regionales o intra-regionales, sin embargo, en este documento sólo se abordarán los movimientos entre regiones³.
 - **No conmutantes regionales:** Corresponde a las personas residentes habituales que tienen su ocupación en su región de residencia.
 - **Conmutantes inter-regionales:** Corresponde a las personas residentes habituales de una región que tienen su ocupación en otra región.
- **Forma de cálculo:** Se utiliza la variable construida en el indicador conm01 (`region_trabajo`) y se compara con la región de residencia habitual (`region`). Las personas que tengan igualdad entre ambas variables se consideran no conmutantes; las personas que tengan desigualdad entre ambas variables, se consideran conmutantes inter-regionales.
- **Universo:** Población de 15 años o más.
- **Variables requeridas:** `region_trabajo`, `region`.

³ Para cuantificar a los conmutantes intra-regionales, es decir, a aquellas personas que se trasladan dentro de su región de residencia para ir a su trabajo, se debe analizar la matriz de conmutación laboral al interior de la región usando una matriz a un nivel administrativo menor, como comuna.

■ Programación en R:

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(cond_commutacion = case_when(
    region == region_trabajo ~ "No conmuta",
    region != region_trabajo ~ "Conmutante inter-regional",
    TRUE ~ NA_character_
  )) %>%
  compute()

personas_censo2024 %>%
  filter(!is.na(cond_commutacion)) %>%
  summarise(pct_conm_ir = sum(cond_commutacion == "Conmutante inter-regional",
    na.rm = TRUE) / n() * 100,
    pct_no_conm = sum(cond_commutacion == "No conmuta",
    na.rm = TRUE) / n() * 100) %>%
  collect() %>%
  mutate(pct_conm_ir = round_half_up(pct_conm_ir, 1),
    pct_no_conm = round_half_up(pct_no_conm, 1))
```

7.2.32. hog01: Porcentaje de hogares con acceso a internet

- **Definición:** Este indicador mide el porcentaje de hogares que disponen de acceso a servicios de internet, independiente del tipo de conexión.

- **Forma de cálculo:**

$$\frac{\text{Hogares con acceso a internet}}{\text{Total de hogares}} \times 100$$

(excluye hogares con no respuesta en p15d_serv_internet_fija, p15e_serv_internet_movil y p15f_serv_internet_satelital)

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `tipo_operativo`, `p15d_serv_internet_fija`, `p15e_serv_internet_movil` y `p15f_serv_internet_satelital`.

- **Programación en R:**

```
hogares_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  # Se excluye no respuesta de tres preguntas de servicios de internet
  filter(p15d_serv_internet_fija != -99 | p15e_serv_internet_movil != -99 |
    p15f_serv_internet_satelital != -99) %>%
  summarise(acceso_internet = sum(p15d_serv_internet_fija == 1 |
    p15e_serv_internet_movil == 1 |
    p15f_serv_internet_satelital == 1),
    pct_hog_int = acceso_internet / n() * 100) %>%
  collect() %>%
  mutate(pct_hog_int = round_half_up(pct_hog_int, 1))
```

7.2.33. hog02: Porcentaje de hogares con tenencia insegura de la vivienda

- **Definición:** Porcentaje de hogares que no tienen certeza jurídica sobre la permanencia en la vivienda en la que residen. Los hogares con tenencia insegura son aquellos que tienen viviendas arrendadas sin contrato, ocupadas de hecho o propiedades en sucesión o litigio.

- **Forma de cálculo:**

$$\frac{\text{Hogares con tenencia insegura de la vivienda}}{\text{Total de hogares (excluye hogares con no respuesta en p12_tenencia_viv)}} \times 100$$

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `tipo_operativo`, `p12_tenencia_viv`.
- **Programación en R:**

```
hogares_censo2024 <- hogares_censo2024 %>%
  mutate(
    inseguridad_tenencia = case_when(
      p12_tenencia_viv %in% c(4,8,9) ~ 1, # Tenencia insegura
      p12_tenencia_viv %in% c(1:3,5:7) ~ 2, # Tenencia segura
      TRUE ~ p12_tenencia_viv) %>%
  compute()

hogares_censo2024 %>%
  filter(tipo_operativo %in% 2 & inseguridad_tenencia != -99) %>% # Excluye no
                                                                    # respuesta
  summarise(inseguridad_tenencia = sum(inseguridad_tenencia == 1),
            pct_inseguridad_tenencia = inseguridad_tenencia / n() * 100) %>%
  collect() %>%
  mutate(pct_inseguridad_tenencia = round_half_up(pct_inseguridad_tenencia, 1))
```

7.2.34. hog03: Tipología de hogar

- **Definición:** Este indicador permite clasificar la composición de los hogares en función de las relaciones de parentesco que se encuentren en ellos.

- **Forma de cálculo:**

Para calcular esta variable se deben identificar las siguientes condiciones:

- **Hogar unipersonal:** Hogar conformado por una persona.
- **Hogar nuclear monoparental:** hogar integrado por la jefatura de hogar y sus hijas o hijos (esto incluye hijas o hijos del cónyuge, conviviente o pareja).
- **Hogar nuclear biparental sin hijos:** hogar integrado por la jefatura de hogar y su pareja⁴.

⁴ Considera como pareja de la jefatura de hogar a las siguientes categorías de parentesco: esposo/a o cónyuge, conviviente por unión civil, conviviente de hecho o pareja.

- **Hogar nuclear biparental con hijos:** hogar compuesto por la jefatura de hogar, su pareja y sus hijas o hijos. Estos últimos pueden ser de ambos, solo de la jefatura de hogar o solo de la pareja.
 - **Hogar compuesto:** hogar compuesto por hogares nucleares (es decir, por cualquiera de las combinatoria de hogares nucleares anteriormente descrita) que además incluyen a *no parientes* de la jefatura de hogar. Estos pueden integrar a *otros parientes* de la jefatura de hogar.
 - **Hogar extendido:** hogar compuesto por hogares nucleares (es decir, por cualquiera de las combinatoria de hogares nucleares anteriormente descrita), incluyendo a otros parientes de la jefatura de hogar. Estos hogares **NO** incluyen a algún *no pariente* de la jefatura de hogar.
 - **Hogar sin núcleo:** hogar integrado por la jefatura de hogar y algún *otro pariente*, *no pariente* y/o *servicio doméstico* *puertas adentro* de la jefatura de hogar.
- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
 - **Variables requeridas:** `parentesco`, `id_vivienda` y `id_hogar`.
 - **Programación en R:**

```
# Variables auxiliares por tipo de parentesco
aux_parentesco <- personas_censo2024 %>%
  filter(tipo_operativo == 2) %>% group_by(id_vivienda, id_hogar) %>%
  summarise(JH = any(parentesco %in% 1, na.rm = TRUE), # jefatura de hogar
            CC = any(parentesco %in% 2:4, na.rm = TRUE), # pareja jh
            HH = any(parentesco %in% 5:6, na.rm = TRUE), # hijo/a
            OP = any(parentesco %in% 7:14, na.rm = TRUE), # otro pariente
            NP = any(parentesco %in% 15, na.rm = TRUE), # no pariente
            SD = any(parentesco %in% 16, na.rm = TRUE)) %>%
  ungroup() %>%
  compute()

# Como es una variable a nivel hogar, se agrega a esta tabla
hogares_censo2024 <- hogares_censo2024 %>%
  left_join(aux_parentesco, by = c("id_vivienda", "id_hogar")) %>%
  compute()

# Construir tipología de hogar según composición de hogares por parentesco
hogares_censo2024 <- hogares_censo2024 %>%
  mutate(t_hogar = case_when(
    JH & !CC & !HH & !OP & !NP & !SD ~ 1, # Unipersonal
    JH & !CC & HH & !OP & !NP & !SD ~ 2, # Nuclear monoparental
    JH & CC & !HH & !OP & !NP & !SD ~ 3, # Nuclear pareja sin hijos
    JH & CC & HH & !OP & !NP & !SD ~ 4, # Nuclear pareja con hijos
    JH & (CC | HH) & OP & !(NP | SD) ~ 5, # Extenso
    JH & (CC | HH) & (NP | SD) ~ 6, # Compuesto
    JH & !(CC | HH) & (OP | NP | SD) ~ 7) %>% # Sin núcleo
  compute()
```

7.2.35. viv01: Promedio de hogares en la vivienda

- **Definición:** Cantidad de hogares promedio por vivienda.
- **Forma de cálculo:**

Total de hogares

Total de viviendas particulares ocupadas con moradores presentes

- **Universo:** Viviendas particulares ocupadas con moradores presentes.
- **Variables requeridas:** `tipo_operativo` y `id_vivienda`.
- **Programación en R:**

```
hogares_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  group_by(id_vivienda) %>%
  summarise(n_hogares = n(), .groups = "drop") %>%
  collect() %>%
  summarise(prom_hogares_viv = mean(n_hogares)) %>%
  mutate(prom_hogares_viv = round_half_up(prom_hogares_viv, 1))
```

7.2.36. viv02: Promedio de personas por vivienda

- **Definición:** Promedio de la cantidad de personas por vivienda.
- **Forma de cálculo:**

Total de personas

Total de viviendas particulares ocupadas con moradores presentes

- **Universo:** Viviendas particulares ocupadas con moradores presentes.
- **Variables requeridas:** `tipo_operativo`.
- **Programación en R:**

```
personas_censo2024 %>%
  filter(tipo_operativo == 2) %>%
  group_by(id_vivienda) %>%
  summarise(n_personas = n(), .groups = "drop") %>%
  collect() %>%
  summarise(prom_personas_viv = mean(n_personas)) %>%
  mutate(prom_personas_viv = round_half_up(prom_personas_viv, 1))
```

7.2.37. viv03: Índice de hacinamiento

- **Definición:** El hacinamiento se mide a partir de la relación entre el total de personas que reside en una vivienda y el total de dormitorios que hay en esta y que son de uso exclusivo para dormir. Se define que una vivienda se encuentra hacinada cuando:

- Hay 2,5 o más personas por dormitorio.
- No hay dormitorios en la vivienda y residen dos o más personas⁵.

Las viviendas pueden ser clasificadas en viviendas sin hacinamiento, con hacinamiento medio y hacinamiento crítico. Se considera que una vivienda tiene hacinamiento medio cuando la razón entre personas y dormitorios se encuentra entre 2,5 y menos de 5 personas por dormitorio. Por otro lado, el hacinamiento se considera crítico si el número de personas por dormitorio es 5 o más o cuando no hay dormitorios en la vivienda y residen dos o más personas en esta.

- **Forma de cálculo:**

Total de personas

Total de dormitorios de uso exclusivo para dormir en la vivienda
(excluye viviendas con no respuesta en p5_num_dormitorios)

- **Universo:** Total de viviendas particulares ocupadas con moradores presentes.

- **Variables requeridas:** tipo_operativo, cant_per y p5_num_dormitorios.

- **Programación en R:**

```
viviendas_censo2024 <- viviendas_censo2024 %>%
  mutate(hacinamiento = round(cant_per / p5_num_dormitorios, 2),
         i_hacinamiento = case_when(
           cant_per == 0 | tipo_operativo != 2 ~ NA_character_,
           p5_num_dormitorios == -99 ~ "No respuesta",
           p5_num_dormitorios == 0 & cant_per == 1 ~ "Sin hacinamiento",
           p5_num_dormitorios == 0 ~ "Hacinamiento crítico",
           hacinamiento < 2.5 & p5_num_dormitorios > 0 ~ "Sin hacinamiento",
           hacinamiento < 5 & p5_num_dormitorios > 0 ~ "Hacinamiento medio",
           hacinamiento >= 5 & p5_num_dormitorios > 0 ~ "Hacinamiento crítico")) %>%
  compute()

viviendas_censo2024 %>%
  filter(tipo_operativo == 2 & cant_per > 0) %>%
  count(i_hacinamiento) %>%
  collect()
```

⁵ La definición de hacinamiento fue actualizada por el Ministerio de Vivienda y Urbanismo para el cálculo del déficit habitacional. Esta definición modifica la definición utilizada en la publicación de junio 2025.

7.2.38. viv04: Índice de materialidad

- **Definición:** Este índice representa la calidad de vivienda en su conjunto, es decir, evalúa la calidad de pisos, paredes y techos.
- **Forma de cálculo:**

Los materiales que se consideran *aceptables* son:

- **Techo:** tejas o tejuelas de arcilla, metálicas, de cemento, de madera, asfálticas o plásticas, losa de hormigón, planchas metálicas de zinc o cobre, planchas de fibrocemento tipo pizarreño.
- **Piso:** parquet, piso flotante, cerámico, madera, alfombra, flexit, cubrepiso u otro similar; sobre radier o vigas de madera.
- **Paredes:** hormigón armado, albañilería: bloque de cemento, ladrillo o piedra, tabique forrado por ambas caras (madera o acero).

Los materiales *recuperables* son:

- **Techo:** fonolita o plancha de fieltro embreado, paja, coirón, totora o caña.
- **Piso:** radier sin revestimiento, baldosa de cemento, capa de cemento sobre tierra.
- **Paredes:** tabique sin forro interior (madera u otro), adobe, barro, pirca, quincha u otro material artesanal.

Los materiales *irrecuperables* son:

- **Techo:** materiales precarios o de desecho o sin cubierta sólida de techo.
- **Piso:** tierra.
- **Paredes:** materiales precarios o de desecho.

A partir de lo anterior, las viviendas se clasifican de acuerdo con las siguientes categorías que componen el índice de materialidad:

- **Materialidad aceptable:** todos sus materiales son aceptables o las paredes se clasifican como recuperables y el resto de materiales como aceptables.
- **Materialidad irrecuperable:** todos sus materiales se clasifican como irrecuperables.
- **Materialidad recuperable:** el resto de las combinaciones de materiales se identifican como viviendas con materialidad recuperable.
- **Universo:** Total de viviendas particulares ocupadas con moradores presentes.
- **Variables requeridas:** p4a_mat_paredes, p4b_mat_techo y p4c_mat_piso.

■ Programación en R:

```
viviendas_censo2024 <- viviendas_censo2024 %>%
  mutate(
    aux_mat_muro = case_when(
      p4a_mat_paredes %in% 1:3 ~ 1L, # Aceptable
      p4a_mat_paredes %in% 4:5 ~ 2L, # Recuperable
      p4a_mat_paredes == 6 ~ 3L, # Irrecuperable
      p4a_mat_paredes == -99 ~ -99, # Ignorada
      is.na(p4a_mat_paredes) ~ NA_integer_), # No aplica
    aux_mat_techo = case_when(
      p4b_mat_techo %in% 1:4 ~ 1L, # Aceptable
      p4b_mat_techo %in% 5:6 ~ 2L, # Recuperable
      p4b_mat_techo %in% 7:8 ~ 3L, # Irrecuperable
      p4b_mat_techo == -99 ~ -99, # Ignorada
      is.na(p4b_mat_techo) ~ NA_integer_), # No aplica
    aux_mat_piso = case_when(
      p4c_mat_piso == 1 ~ 1L, # Aceptable
      p4c_mat_piso %in% 2:4 ~ 2L, # Recuperable
      p4c_mat_piso == 5 ~ 3L, # Irrecuperable
      p4c_mat_piso == -99 ~ -99, # Ignorada
      is.na(p4c_mat_piso) ~ NA_integer_), # No aplica
    # Índice de materialidad
    i_materialidad = case_when(
      is.na(aux_mat_muro) | is.na(aux_mat_piso) |
        is.na(aux_mat_techo) ~ NA_character_,
      aux_mat_muro == -99 | aux_mat_piso == -99 |
        aux_mat_techo == -99 ~ "No responde",
      aux_mat_muro == 1 & aux_mat_piso == 1 &
        aux_mat_techo == 1 ~ "Materialidad aceptable",
      aux_mat_muro == 3 | aux_mat_piso == 3 |
        aux_mat_techo == 3 ~ "Materialidad irrecuperable",
      TRUE ~ "Materialidad recuperable")) %>%
  compute()

viviendas_censo2024 %>%
  filter(tipo_operativo == 2 & cant_per > 0) %>%
  count(i_materialidad) %>%
  collect()
```

7.2.39.viv05: Índice de tipo de vivienda

- **Definición:** Este indicador clasifica a las viviendas particulares de acuerdo con su tipo. Las categorías en que se clasifican las viviendas son: viviendas de tipo aceptable y viviendas de tipo irrecuperable.
- **Forma de cálculo:**

Las viviendas de *tipo aceptable* son:

- Casa (con acceso directo o en condominio cerrado)
- Departamento (en edificio con o sin ascensor)
- Vivienda tradicional indígena (ruka u otras)
- Pieza en casa antigua o conventillo

Las viviendas de *tipo irrecuperable* son:

- Mediagua, mejora, vivienda de emergencia, rancho o choza
 - Móvil (carpa, casa rodante o similar)
 - Otro tipo de vivienda particular
- **Universo:** Viviendas particulares ocupadas con moradores presentes.
 - **Variables requeridas:** `p2_tipo_vivienda`.
 - **Programación en R:**

```
viviendas_censo2024 <- viviendas_censo2024 %>%
  mutate(ind_tipo = case_when(
    # cant_per > 0 se asegura que el análisis sea sobre las viv particulares con moradores
    # presentes que respondieron el censo
    p2_tipo_vivienda %in% c(1,3,5,6) & (cant_per > 0 & tipo_operativo == 2) ~ "Aceptable",
    p2_tipo_vivienda %in% 7:9 & (cant_per > 0 & tipo_operativo == 2) ~ "Irrecuperable",
    TRUE ~ NA_character_)) %>%
  compute()

viviendas_censo2024 %>%
  filter(tipo_operativo == 2 & cant_per > 0) %>%
  count(ind_tipo) %>%
  collect()
```

7.2.40. viv06: Viviendas irrecuperables

- **Definición:** Las viviendas irrecuperables son aquellas que debido a su tipo o a la materialidad de sus paredes, techo y piso, deben ser reemplazadas. De acuerdo con el Ministerio de Vivienda y Urbanismo (Minvu), el indicador de viviendas irrecuperables es uno de los cuatro componentes del déficit habitacional cuantitativo.
- **Forma de cálculo:**

Las siguientes combinaciones de tipo de vivienda y materialidad clasifican a una vivienda como irrecuperable:

 - **Tipo de vivienda:** mediagua, mejora, vivienda de emergencia, rancho o choza, móvil (carpa, casa rodante o similar) y otros tipos de vivienda particular.
 - **Materialidad:** materiales precarios o de desecho en paredes o cubierta de techo; sin cubierta en el techo o con piso de tierra.
- **Universo:** Viviendas particulares ocupadas con moradores presentes.
- **Variables requeridas:** `i_materialidad` y `ind_tipo`.

■ Programación en R:

```
viviendas_censo2024 <- viviendas_censo2024 %>%
  mutate(viv_irr = case_when(
    is.na(ind_tipo) | is.na(i_materialidad) ~ NA_integer_,
    ind_tipo == "No responde" | i_materialidad == "No responde" ~ -99,
    ind_tipo == "Irrecuperable" |
      i_materialidad == "Materialidad irrecuperable" ~ 1L,
    TRUE ~ 0)) %>%
  compute()

viviendas_censo2024 %>%
  summarise(viv_irrecuperables = sum(viv_irr==1, na.rm = T)) %>%
  collect()
```

7.2.41. viv07: Hogares allegados

- **Definición:** Indicador que identifica cuando dos o más grupos de personas que no comparten presupuesto de alimentación, comparten una misma vivienda. También se puede definir como el excedente de hogares en relación al total de viviendas particulares ocupadas con moradores presentes que contestan. Este componente también se conoce como allegamiento externo.

De acuerdo con el Ministerio de Vivienda y Urbanismo (Minvu), el indicador que identifica a los hogares allegados es uno de los cuatro componentes del déficit habitacional cuantitativo.

- **Forma de cálculo:**

Total de hogares en la vivienda - 1

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`), ya que sólo en este operativo se identifican hogares.
- **Variables requeridas:** `tipo_operativo` y `cant_hog`.

- **Programación en R:**

```
viviendas_censo2024 <- viviendas_censo2024 %>%
  mutate(hog_allegados = if_else(tipo_operativo == 2 & cant_hog > 0,
                                cant_hog - 1, 0, NA)) %>%
  compute()

viviendas_censo2024 %>%
  summarise(hogares_allegados = sum(hog_allegados, na.rm = T)) %>%
  collect()
```

7.2.42. viv08: Núcleos allegados hacinados

- **Definición:** Los núcleos allegados, también conocido como allegamiento interno, corresponden a núcleos familiares adicionales al núcleo principal del hogar⁶. Los núcleos se identifican a partir de las relaciones de parentesco al interior del hogar. En este sentido, al interior de un hogar pueden encontrarse uno o más núcleos familiares.

Se considera que existen núcleos allegados hacinados cuando estos residen en viviendas con 2,5 o más personas por dormitorio, o cuando residen 2 o más personas en una vivienda que no tiene dormitorios de uso exclusivo para dormir.

⁶ El núcleo principal es el núcleo familiar de la jefatura de hogar.

De acuerdo con el Ministerio de Vivienda y Urbanismo (Minvu), el indicador que identifica a los núcleos allegados hacinados es uno de los cuatro componentes del déficit habitacional cuantitativo.

- **Forma de cálculo:** Para la construcción de esta variable se sugiere seguir los siguientes pasos:

1. Identificar las relaciones de parentesco para la construcción de los núcleos (*recparen*).

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(
    recparen = case_when(
      parentesco == 1 ~ 1, # jefatura de hogar
      parentesco %in% c(2:4) ~ 2, # pareja o cónyuge
      parentesco %in% c(5:6) ~ 3, # hijos/as
      parentesco == 11 ~ 4, # yerno/nuera
      parentesco == 12 ~ 5, # nieto/a
      parentesco %in% c(7, 9) ~ 6, # hermano/cuñado
      parentesco %in% c(8, 10) ~ 7, # padres/suegros
      parentesco %in% c(13, 14) ~ 8, # otro pariente
      parentesco == 15 ~ 9, # no pariente
      parentesco == 16 ~ 10) %>% # sdpa
  compute()
```

2. Identificar el estado civil o conyugal actual de los hijos de la jefatura de hogar (*hijoecivil*).

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(
    hijoecivil = case_when(
      recparen == 3 & p23_est_civil == 8 ~ 1,
      recparen == 3 & p23_est_civil == -99 & edad_quinquenal >= 15 ~ NA_integer_,
      # hijos/as solteros/as
      recparen == 3 & p23_est_civil != 8 & edad_quinquenal >= 15 ~ 2,
      recparen == 3 & edad_quinquenal < 15 ~ NA_integer_,
      TRUE ~ NA_integer_) %>%
  compute()
```

3. Identificar el total de núcleos a nivel hogar a partir de las relaciones de parentesco (*total_nucleos_hogar*).

Los tipos de núcleos que se consideran son:

- **Núcleos de yernos o nueras:** se calcula contando la cantidad de yernos o nueras de cada hogar. Cada uno de ellos corresponde a un núcleo allegado.
- **Núcleos de hijos/as no solteros según la presencia de yernos o nueras:** se calcula contando la cantidad de hijos/as no solteros (ya sea de la jefatura de hogar o de la pareja o cónyuge) si es que en el hogar no hay presencia de núcleos de yernos o nueras.
- **Núcleo de hijas madres solteras:** se calcula contando la cantidad de hijas con estado civil o conyugal actual solteras, que tienen hijos nacidos sobrevivientes. Esto solo se realiza en hogares en donde hay presencia del parentesco nieto/a.
- **Núcleo de padres o suegros:** contabiliza un núcleo allegado por cada dos padres y/o suegros, considerando un máximo de dos núcleos de este tipo por hogar.

- **Núcleos de hermanos/as o cuñados/as:** contabiliza un núcleo allegado por cada dos hermanos y/o cuñados, considerando un máximo de dos núcleos de este tipo por hogar.
- **Núcleos de otros parientes:** define como otros parientes a nietos, abuelos y la categoría otro pariente. Contabiliza un núcleo allegado por cada dos otros parientes, considerando un máximo de dos núcleos de este tipo por hogar.
- **Núcleos de no parientes:** contabiliza un núcleo allegado por cada dos no parientes, considerando un máximo de dos núcleos de este tipo por hogar.

Para los núcleos de hermanos/as o cuñados/as, núcleos de otros parientes y núcleos de no parientes el valor impar⁷ se nivela hacia abajo, con un tope máximo de dos núcleos.

```
personas_censo2024 <- personas_censo2024 %>%
  mutate(
    # yernos y nueras
    nyern = case_when(
      reccaren == 4 ~ 1L,
      TRUE ~ 0),
    # hijos no solteros
    hijonosol = case_when(
      hijoecivil == 2 ~ 1L,
      TRUE ~ 0),
    # nietos
    connieto = case_when(
      reccaren == 5 ~ 1L,
      TRUE ~ 0),
    # madres con estado civil/conyugal soltero
    masol = case_when(
      hijoecivil == 1 & p47a_tot_hijs_sobrev > 0 ~ 1L,
      TRUE ~ 0),
    # padres y suegros
    padsuegr = case_when(
      reccaren == 7 ~ 1L,
      TRUE ~ 0),
    # hermanos y cuñados
    hermcun = case_when(
      reccaren == 6 ~ 1L,
      TRUE ~ 0),
    # otros parientes
    otropar = case_when(
      reccaren == 8 ~ 1L,
      TRUE ~ 0),
    # identificar no parientes
    nopar = case_when(
      reccaren == 9 ~ 1L,
      TRUE ~ 0)) %>%
  to_duckdb() %>%
  group_by(id_vivienda, id_hogar) %>%
  mutate(
    # contar yernos y nueras en el hogar
    nuc_yer_nue = sum(nyern, na.rm = TRUE),
    # contar hijos no solteros en el hogar
    tot_hijonosol = sum(hijonosol, na.rm = TRUE),
    # contar nietos residentes en el hogar
    tot_connieto = sum(connieto, na.rm = TRUE),
    # contar madres solteras en el hogar
    tot_masol = sum(masol, na.rm = TRUE),
```

⁷ Dado que el máximo de núcleos es dos, esto solo aplica para cuando se detectan tres núcleos de estos tipos.

```

# contar padres y suegros en el hogar
tot_pad_suegr = sum(padsuegr, na.rm = TRUE),
# contar hermanos y cuñados en el hogar
tot_hermcun = sum(hermcun, na.rm = TRUE),
# contar otros parientes en el hogar
tot_otropar = sum(otropar, na.rm = TRUE),
# contar no parientes en el hogar
tot_nopar = sum(nopar, na.rm = TRUE)) %>%
ungroup() %>%
to_arrow() %>%
mutate(
  # núcleos de hijos no solteros sin yernos
  nuc_hijns = case_when(
    nuc_yer_nue == 0 ~ tot_hijonosol,
    TRUE ~ 0),
  # núcleos de madres solteras sin yernos y con nietos
  nuc_masol = case_when(
    nuc_yer_nue == 0 & tot_connieto > 0 ~ tot_masol,
    TRUE ~ 0),
  # núcleos de padres y suegros
  nuc_pad_suegr = case_when(
    tot_pad_suegr %in% 0:1 ~ 0L,
    tot_pad_suegr == 2 ~ 1,
    tot_pad_suegr >= 3 ~ 2),
  # núcleos de hermanos y cuñados
  nuc_hnocun = case_when(
    tot_hermcun %in% 0:1 ~ 0L,
    tot_hermcun %in% 2:3 ~ 1L,
    tot_hermcun > 3 ~ 2L),
  # núcleos de otros parientes
  nuc_otropar = case_when(
    tot_otropar %in% 0:1 ~ 0L,
    tot_otropar %in% 2:3 ~ 1L,
    tot_otropar > 3 ~ 2L),
  # núcleos de no parientes
  nucnopar = case_when(
    tot_nopar %in% 0:1 ~ 0L,
    tot_nopar %in% 2:3 ~ 1L,
    tot_nopar > 3 ~ 2L)) %>%
compute()

# total de núcleos secundarios por hogar
nucleos_secundarios <- personas_censo2024 %>%
  select(id_vivienda, id_hogar,
         parentesco, nuc_yer_nue, nuc_pad_suegr,
         nuc_hnocun, nuc_otropar, nucnopar, nuc_masol, nuc_hijns) %>%
  filter(parentesco == 1) %>%
  compute()

nucleos_secundarios <- nucleos_secundarios %>% mutate(
  total_nucleos_hogar = coalesce(nuc_yer_nue, 0) +
    coalesce(nuc_hijns, 0) +
    coalesce(nuc_masol, 0) +
    coalesce(nuc_pad_suegr, 0) +
    coalesce(nuc_hnocun, 0) +
    coalesce(nuc_otropar, 0) +
    coalesce(nucnopar, 0)) %>%
  compute()

```

4. Calcular el total de núcleos allegados hacinados.

- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`). Solo en este operativo se identifican hogares.
- **Variables requeridas:** `tipo_operativo`, `i_hacinamiento`, `total_nucleos_hogar`.
- **Programación en R:**

Para usar esta programación, se asume que el o la usuaria ha construido el indicador v03.

```
hogares_censo2024 <- hogares_censo2024 %>%
  left_join(nucleos_secundarios, by = c("id_vivienda", "id_hogar")) %>%
  left_join(viviendas_censo2024, by = "id_vivienda") %>%
  compute()

rm(nucleos_secundarios)

hogares_censo2024 <- hogares_censo2024 %>%
  mutate(
    nuc_hac = case_when(
      i_hacinamiento %in% c("Hacinamiento medio", "Hacinamiento crítico")
      ~ total_nucleos_hogar,
      TRUE ~ NA_integer_) %>%
  compute()

hogares_censo2024 %>%
  summarise(nucleos_hacinados_allegados =
    sum(nuc_hac, na.rm = T)) %>%
  collect()
```

7.2.43. viv09: Hogares que residen en viviendas hacinadas no ampliables

- **Definición:** Corresponde a aquellos hogares que no pueden ampliar la vivienda en la que residen, ya sea porque viven en departamentos, piezas o porque son arrendatarios.

De acuerdo con el Ministerio de Vivienda y Urbanismo (Minvu), el indicador que identifica a los hogares que residen en viviendas hacinadas no ampliables es uno de los cuatro componentes del déficit habitacional cuantitativo.

- **Forma de cálculo:**

Para identificar los requerimientos a partir de este componente se requiere:

- Identificar a viviendas hacinadas que no han sido identificadas en componentes del déficit anteriores, es decir, no presentan hogares o núcleos allegados, ni tampoco se identifican como viviendas irrecuperables. A partir de esto se identifican las viviendas con requerimientos de ampliación.

- Identificadas las viviendas con requerimientos de ampliación es necesario identificar a aquellas que no son ampliables. Estas viviendas son:
 - Viviendas que son departamentos o piezas en casas antiguas o conventillos, o,
 - Hogares cuya tenencia de la vivienda es arrendada (con o sin contrato).
- **Universo:** Total de hogares. Para contabilizar correctamente a los hogares se debe filtrar por el operativo de viviendas particulares (`tipo_operativo == 2`). Solo en este operativo se identifican hogares.
- **Variables requeridas:** `p2_tipo_vivienda`, `p12_tenencia_viv`, `i_hacinamiento`, `cant_hog`, `total_nucleos_hogar` y `viv_irr`.
- **Programación en R:**

Para usar esta programación, se asume que el o la usuaria ha construido los indicadores `v03` y `v06`.

```
# anteriormente se debe agregar la variable de viviendas irrecuperables (viv_irr)
# y de hacinamiento (i_hacinamiento) desde la tabla de viviendas a la tabla
# de hogares

hogares_censo2024 <- hogares_censo2024 %>%
  mutate(
    # requerimientos de ampliación, viviendas hacinadas
    reqampli = case_when(
      id_hogar == 1 &
        i_hacinamiento %in% c("Hacinamiento medio", "Hacinamiento crítico") &
        total_nucleos_hogar == 0 & cant_hog == 1 & viv_irr != 1 ~ 1,
      TRUE ~ 0),
    # requerimientos de amplicacion, pero no ampliable
    reqhna = case_when(
      reqampli == 1 & p2_tipo_vivienda %in% c(3, 6) ~ 1, # dpto con y sin ascensor +
                                                         # pieza
      reqampli == 1 & p12_tenencia_viv %in% 3:4 ~ 1, # arrendada
      TRUE ~ 0)) %>%
  compute()

hogares_censo2024 %>%
  summarise(hog_viv_hac_no_ampliable = sum(reqhna, na.rm = T)) %>%
  collect()
```

7.2.44. viv10: Déficit habitacional cuantitativo

- **Definición:** El déficit habitacional cuantitativo mide la cantidad de viviendas que se requieren para atender la necesidad de hogares que no cuentan con una vivienda exclusiva o de hogares que residen en viviendas que no cumplen los estándares mínimos de habitabilidad.
- **Forma de cálculo:**

De acuerdo con el Minvu, la metodología oficial para el cálculo del déficit habitacional cuantitativo está basada en la suma de cuatro componentes:

- **Viviendas irrecuperables:** corresponden a aquellas que por su tipo o materialidad deben ser reemplazadas.

- **Hogares allegados:** ocurre cuando dos o más grupos de personas que no comparten presupuesto de alimentación comparten una misma vivienda.
 - **Núcleos allegados hacinados:** se genera cuando existen núcleos adicionales al núcleo principal del hogar y residen en viviendas hacinadas.
 - **Viviendas con hacinamiento no ampliable:** corresponde a hogares que no pueden ampliar la vivienda en la que residen, ya sea por el tipo de vivienda o la tenencia de esta.
- **Universo:** Viviendas particulares ocupadas con moradores presentes.
 - **Variables requeridas:** tipo_operativo, cant_hog, i_hacinamiento, total_nucleos_hogar, viv_irr, y reqhna.
 - **Programación en R:**

```
hogares_censo2024 <- hogares_censo2024 %>%
  mutate(viv_irrecuperables = if_else(id_hogar == 1 & viv_irr == 1, 1, 0),
         hog_allegados = if_else(id_hogar == 1, hog_allegados, 0)) %>%
  compute()

deficit_cuantitativo <- hogares_censo2024 %>%
  summarise(tot_viv_irrecuperables = sum(viv_irrecuperables, na.rm = T),
            tot_hog_allegados = sum(hog_allegados, na.rm = T),
            tot_nuc_allegados_hac = sum(nuc_hac, na.rm = T),
            tot_viv_hac_no_ampliable = sum(reqhna, na.rm = T)) %>%
  collect()

deficit_cuantitativo <- deficit_cuantitativo %>%
  pivot_longer(cols = matches("tot_"), names_to = "Componente",
               values_to = "n") %>%
  janitor::adorn_totals() %>%
  collect()
```

7.2.45. nr01: Tasa de no respuesta al ítem

- **Definición:** Porcentaje de viviendas, hogares o personas que no responden una pregunta, habiendo cumplido la condición de habilitación de la misma.
- **Forma de cálculo:**

$$\frac{\text{Unidad (vivienda, hogar, persona) con no respuesta en una pregunta}}{\text{Unidad (vivienda, hogar, persona) a la que se le habilitó la pregunta}} \times 100$$
- **Universo:** Unidades a las que se les habilitó la pregunta. Se utilizará de ejemplo la pregunta p25_lug_nacimiento. En este caso, la pregunta debió habilitarse a toda la población.
- **Variables requeridas:** p25_lug_nacimiento.
- **Programación en R:**

```
personas_censo2024 %>%
  summarise(tasa_nr_p25_lug_nacimiento = sum(p25_lug_nacimiento == -99,
                                             na.rm = TRUE) / n() * 100) %>%
  collect() %>%
  mutate(tasa_nr_p25_lug_nacimiento = round_half_up(tasa_nr_p25_lug_nacimiento, 1))
```


7.2.46. na01: Frecuencia de casos en los que no aplica respuesta

- **Definición:** Cantidad de viviendas, hogares, personas para los que no aplica responder una pregunta (NA) al no cumplir la condición de habilitación de esta dentro del diseño del cuestionario.
- **Forma de cálculo:** Conteo de casos que no cumplen con la condición de habilitación de alguna pregunta.
- **Universo:** Conjunto de casos en que la pregunta no se habilitó para la unidad. Se utiliza de ejemplo a las personas que se le habilitó la pregunta p37_alfabet por filtro de edad.
- **Variables requeridas:** p37_alfabet.
- **Programación en R:**

```
personas_censo2024 %>%
  summarise(na_p37_alfabet = sum(is.na(p37_alfabet))) %>%
  collect()
```

7.2.47. tip_op01: Frecuencia de casos por tipo de operativo

- **Definición:** Cantidad de personas según tipo de operativo.

Los tipos de operativo son los siguientes:

1. **Personas en situación calle:** operativo destinado a censar a aquellas personas que carecen de residencia fija y que pernoctan en lugares públicos o privados, que no tienen las características básicas de una vivienda, aunque cumplan esa función.
2. **Viviendas particulares:** operativo destinado a censar toda construcción estructuralmente separada e independiente ubicada dentro de un sitio, la cual se encuentra destinada, total o parcialmente, para el alojamiento de personas y hogares.
3. **Viviendas colectivas:** operativo destinado a censar viviendas que han sido construidas y destinadas al alojamiento de personas, los que por diversas razones hacen vida en común, ya sea por trabajo, estudios u otro.

- **Forma de cálculo:** Conteo de personas por tipo de operativo.
- **Universo:** Total de personas.
- **Variables requeridas:** tipo_operativo.
- **Programación en R:**

```
personas_censo2024 %>%
  count(tipo_operativo) %>%
  collect()
```

