



MANUAL DE USUARIO 10° ENCUESTA NACIONAL DE JUVENTUDES 2022

Área de Estudios
Departamento de Planificación y Estudios
Instituto Nacional de la Juventud

Noviembre 2022

Contenido

1. Introducción.....	3
2. Ficha metodológica.....	4
2.1. Ficha metodológica encuesta a jóvenes.....	4
2.2. Ficha metodológica encuesta a adultos.....	4
3. Diseño de los cuestionarios.....	5
3.1. Estructura cuestionario población joven.....	8
3.2. Estructura cuestionario población adulta.....	9
4. Encuesta a población joven.....	9
4.1. Diseño muestral población joven.....	10
4.2. Universo y tamaño muestral encuesta a población joven.....	12
5. Encuesta a población adulta.....	20
5.1. Diseño muestral población adulta.....	21
5.2. Universo y tamaño muestral encuesta a población adulta.....	21
6. Marco muestral, unidades muestrales y capturas de datos en terreno.....	28
6.1 Fase 2: Mix-mode, Encuesta telefónica y Encuesta online a personas sorteadas en hogares como técnica de levantamiento.....	37
6.2 Fase 3: Recorrido de casos pendientes de las fases 1 y 2.....	40
7. Validación, cálculo de expansores y ponderación de las bases de datos.....	40
8. Cálculo y clasificación de niveles socioeconómicos.....	47
9. Cuadro comparativo Encuesta Nacional de Juventud (1994-2022).....	48

1. Introducción

Desde 1994 que el Instituto Nacional de la Juventud (INJUV) desarrolla la Encuesta Nacional de Juventudes, en adelante ENJUV, la cual es la herramienta cuantitativa sobre jóvenes con mayor consistencia y alcance técnico de Iberoamérica. El principal foco de estudio son las personas entre los 15 y 29 años, teniendo como objetivo principal la **obtención de información pertinente y oportuna para el diagnóstico de las realidades juveniles en Chile**. Los datos que recaba la encuesta dotan de fuentes empíricas a académicos, organizaciones de la sociedad civil y servicios públicos, entre otros, para el diseño y elaboración de políticas e iniciativas públicas orientadas a este relevante segmento de la población.

Para cada nuevo proceso la ENJUV se somete a mecanismos de revisión y validación, los cuales han complejizado y optimizado su representatividad estadística, precisión y utilidad para quienes trabajan con este instrumento. Un ejemplo de lo anterior es la evolución del cuestionario de esta encuesta, el cual edición tras edición es actualizado y perfeccionado.

La primera Encuesta Nacional de Juventudes, desarrollada en 1994, solo se aplicó en los principales centros urbanos del país; ya la segunda versión, en 1997, contó con representatividad a nivel nacional en población urbana. Para la tercera versión se incluye la representatividad en población rural, con mejoras en sus protocolos de trabajo y cuestionarios gracias a la colaboración de expertos y académicos. A partir de 2003, en lo que fue la cuarta versión de la encuesta, es posible realizar inferencia estadística a nivel regional, y se instaura una política sistemática de difusión nacional como regional, incorporando también, en 2006 y 2009, la participación de una red de expertos en el análisis de los resultados.

Desde 2012, en la séptima versión de la encuesta, se generan mejoras e innovaciones en el cuestionario, con el fin de mejorar no sólo la precisión y confiabilidad estadística, sino que también resguardar la información de la población en estudio. Así, desde aquella versión se utiliza un cuestionario autoaplicado en aquellos módulos de la encuesta que abordan temas sensibles -como consumo de sustancias, vida sexual, salud mental y violencia-, que tiene como objetivo preservar el anonimato de las respuestas y disminuir la declaración de datos inexactos. Posteriormente, desde la octava y novena versión de 2015 y 2018, respectivamente, se aplica la encuesta a jóvenes, y, además, un cuestionario acotado a una muestra de población adulta -entre 30 y 59 años- representativa a nivel nacional, para cotejar si aquellos resultados asociados a jóvenes son cuestiones específicas a esta población -la juvenil-, o responden a estructuras sociales más amplias.

La presente versión, la Décima Encuesta Nacional de Juventudes 2022, fue desarrollada bajo contextos sociales complejos. Las protestas de octubre de 2019 –donde la población joven tuvo un papel relevante–, el confinamiento producto de la crisis sanitaria por COVID-19 (el cual también golpeó de manera diferencial a las juventudes), y el proceso constituyente en marcha desde 2020, supusieron no sólo generar condiciones para flexibilizar y facilitar la aplicación de la encuesta, sino que también se posicionó el foco sobre los ejes institucionales del INJUV en pos de la generación de iniciativas públicas. En este sentido, además de mantener la continuidad de la mayoría de los indicadores en relación a la novena versión de 2018, una de las medidas adoptadas para la presente edición fue la aplicación del instrumento en una modalidad mixta-secuencial, con la posibilidad de que los datos sean recolectados mediante un llamado telefónico, previa selección probabilística del respondiente mediante los sorteos protocolares establecidos en la metodología.

Para efectos de esta versión, se continúa con la búsqueda de comprender las principales problemáticas de la población joven, junto con mantener lo que le otorga un valor agregado a esta herramienta, a decir, los indicadores asociados a percepciones y evaluaciones de diversos aspectos

sociales. Finalmente, la **10ma. Encuesta Nacional de Juventudes 2022** es una Encuesta que, si bien es más acotada que la versión de 2018, fue diseñada para ser funcional a agendas de investigación y organismos públicos que requieren de la continuidad característica de la serie de Encuestas Nacionales de Juventud de Chile.

2. Ficha metodológica

A continuación, se presentan las fichas metodológicas resumidas tanto de la encuesta a jóvenes como adultos para el levantamiento de la 10ma. Encuesta Nacional de Juventudes 2022:

2.1. Ficha metodológica encuesta a jóvenes

Universo	Jóvenes, hombres y mujeres, de entre 15 y 29 años, residentes en todas las regiones del país, en zonas urbanas y rurales.
Áreas de representación	Nacional, regional, urbano-rural.
Diseño muestral	Muestreo probabilístico polietápico. El tamaño de la muestra es de 9.700 casos. El margen de error muestral asciende a +/- 1,00%, a nivel nacional (y de un máximo de 4,9% a nivel regional), con un nivel de confianza del 95%, bajo supuesto de varianza máxima.
Cobertura	141 comunas, correspondientes a un 88,2%, en función del tamaño del universo contenido en el marco muestral efectivo en relación al marco muestral teórico.
Técnica	Entrevista presencial en hogares, en tablet. Formato autoaplicado para todas las preguntas referidas a vida sexual, consumo de sustancias, salud mental y violencia. Posibilidad de levantamiento remoto mediante metodología mixta secuencial (636 casos, 6,6% del total de la muestra).
Periodo de recolección de datos	Diciembre 2021 a mayo de 2022.
Instrumento	Cuestionario con 112 preguntas: 69 cara a cara y 43 de manera autoaplicada.
Ejecutor	IPSOS bajo la dirección del INJUV.

2.2. Ficha metodológica encuesta a adultos

Universo	Adultos, hombres y mujeres, de entre 30 y 59 años, residentes en todas las regiones del país, en zonas urbanas y rurales.
Áreas de representación	Nacional.
Diseño muestral	Muestreo probabilístico polietápico. El tamaño de la muestra es de 1.000 casos. El margen de error muestral asciende a +/- 2,55% a nivel nacional con un nivel de confianza del 95%, bajo supuesto de varianza máxima.
Cobertura	141 comunas, correspondientes a un 86,5%, en función del tamaño del universo contenido en el marco muestral efectivo en relación al marco muestral teórico.
Técnica	Entrevista presencial en hogares, en tablet. Formato autoaplicado para todas las preguntas referidas a vida sexual, consumo de sustancias, salud mental y violencia. Posibilidad de levantamiento remoto mediante metodología mixta secuencial (47 casos, 4,7% del total de la muestra).
Periodo de recolección de datos	Diciembre 2021 a mayo de 2022.
Instrumento	Cuestionario con 99 preguntas: 50 cara a cara y 27 de manera autoaplicada.
Ejecutor	IPSOS bajo la dirección del INJUV.

3. Diseño de los cuestionarios

Los cuestionarios, tanto de la población adulta como de la población joven, fueron diseñados por el equipo del Área de Estudios del Instituto Nacional de la Juventud, el cual procuró mantener parte del instrumento con el fin de permitir la continuidad de la aplicación de la encuesta. Para esta edición, el cuestionario de la 10ma. ENJUV contó con 13 módulos temáticos, a saber: Participación social; Representaciones valóricas; Educación; Política; Confianza y discriminación; Trabajo; Pareja y familia; Prácticas económicas; Caracterización general; Consumo de sustancias; Vida sexual; Salud mental; y Violencia.

Sin perjuicio de lo anterior, producto de las condiciones sanitarias que impactan a nivel global debido a la pandemia por COVID-19, se establecieron variaciones a la 10ma. edición de la Encuesta Nacional de Juventudes, otorgándole flexibilidad en las diversas etapas del diseño y levantamiento de información, tal como lo indican organismos internacionales. Esta flexibilidad dice relación con integrar mecanismos de levantamiento remoto, vía web o telefónico, para dar seguridad tanto a encuestadores como a encuestados. Esto supone pasar desde la modalidad presencial, a una mixta secuencial, tal como lo realizan otras encuestas nacionales de alto alcance metodológico, como la de Caracterización Socioeconómica Nacional (CASEN) en su versión 2020¹.

El principal cambio que genera esta metodología, además del tipo de levantamiento, es la necesidad de reducir de manera drástica el cuestionario, entendiendo que en la literatura el promedio de duración máxima de la entrevista debe ser de 30 minutos. Esto resulta particularmente problemático entendiendo que el cuestionario cara a cara de la 9° ENJUV tiene una extensión de 111 preguntas, lo que redundaba en un tiempo aproximado de aplicación de 50 minutos en términos presenciales; cuestión inviable de reproducir en términos telefónicos.

En este sentido, y como forma de generar los menores efectos de este cambio en la comunidad investigativa y los propios objetivos del INJUV, es que el Área de Estudios promovió mecanismos de consulta pública para conocer cuáles son los principales indicadores y preguntas utilizadas por la comunidad investigativa y nuestra red de expertos y expertas en materia juvenil, como también consultas internas con los diversos equipos de programas del INJUV para determinar cuáles son los indicadores imprescindibles para su trabajo en el desarrollo de las diversas iniciativas y programas (formulación y generación de fichas ex ante). Esto generó una reducción significativa del cuestionario de la 9° a la 10° ENJUV, pasando desde las 111 preguntas del módulo cara a cara en la primera, a 69 para esta décima versión.

Así, el presente documento muestra los principales cambios de preguntas en los módulos de la Encuesta establecidos para la 10° ENJUV, con base en el cuestionario de la 9° versión. Además de la utilización de estos indicadores, los principales criterios para incluir o no una pregunta fueron: 1) la

¹ Para mayor información sobre este método de consulta, revisar: http://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/casen/2020/Nota_tecnica1_Modalidad_de_aplicacion_Casen_en_pandemia_2020.pdf.

historicidad de la pregunta, 2) si se afecta el indicador y su calidad cambiando el modo de consulta a vía telefónica, y 3) la funcionalidad para el trabajo interno y externo de INJUV.

En los siguientes apartados se presenta la estructura de cada cuestionario, señalando luego los nuevos indicadores agregados a la encuesta.

Principales cambios por módulo

Módulo 1: Participación sociopolítica

Se mantienen indicadores sobre participación en organizaciones, como dirigentes, y también el de participación en actividades de voluntariado.

Se eliminan preguntas sobre razones de participación y de no participación en organizaciones, por temas de cambio de metodología (ya que es una lista a respuesta espontánea, lo que impide comparabilidad). También se elimina la pregunta sobre frecuencia de actividades de tiempo libre, debido a lo complicado de agregar esa forma de parrilla en una aplicación telefónica y por la validez de la pregunta en tiempos de pandemia.

Se agrega indicador comprometido con CELADE en reunión latinoamericana sobre el Consenso de Montevideo y Desarrollo, sobre la participación juvenil en instancias gubernamentales de decisión.

Módulo 2: Representaciones

Si bien se mantiene el grueso del módulo, se eliminan las preguntas de segunda mención por necesidades de extensión e inutilización (específicamente sobre las condiciones para tener éxito y ser feliz). Además, debido a la necesidad de mantener indicadores históricos, se eliminan las preguntas sobre percepciones de desarrollo intergeneracional.

En términos de las parrillas de consulta, se decide eliminar al menos la mitad de los indicadores, promoviendo aquellos que sean más funcionales al estudio juvenil y programas (como en el de satisfacción con aspectos de la vida; de roles de género y acuerdo con diversas iniciativas sociales).

Módulo 3: Educación

Este módulo posee variados cambios: se elimina la pregunta sobre matrícula, ya que desde la 9° ENJUV se decide preguntar por asistencia (para 2022, virtual o presencial). También se deciden eliminar por no uso las preguntas sobre el último curso aprobado, el tipo de universidad donde cursó estudios superiores, la segunda principal razón para no estudiar y la que indica lo que se puede lograr con la educación. Sumado a ello, dado el cambio de modalidad a aplicación telefónica, se eliminan indicadores sobre el principal factor que define una educación de calidad (se necesita mostrar tarjeta o *tablet*), y, además, por la incapacidad de comparar, se decide también eliminar el indicador sobre evaluación de diversos aspectos del establecimiento educacional. Por asuntos de longitud del cuestionario, también se elimina el indicador sobre nivel de estudios que le gustaría alcanzar.

Módulo 4: Política

Es, debido a su uso público e interno, el módulo que sufre menores modificaciones. Solo se modifican por temas de longitud la segunda acción que realizarían frente a una ley que consideran injusta, como la pregunta sobre justificación de acciones de violencia (sobre esta última, la decisión pasó por la dificultad de replicar la escala de respuesta vía telefónica -parrilla de 1 a 10 en torno a nivel de

justificación-). Finalmente, por temas de comparabilidad y restricciones de movilidad por pandemia, se modifica pregunta por participación no convencional, aumentando el periodo a 2 años (desde las movilizaciones de octubre de 2019).

Módulo 5: Confianza y discriminación

Las principales modificaciones del módulo se encuentran en la eliminación de los indicadores de confianza en instituciones y personas. Esto se decide debido a su extensión (ambas parrillas de 14 categorías), como también a la incapacidad de poder comparar debido a su escala de respuesta (de 1 a 10). Por último, se elimina el segundo grupo más discriminado, recabando solo la primera mención.

Módulo 6: Trabajo

Este módulo presenta pocas modificaciones. Se eliminan los indicadores de horas de trabajo en la actividad principal (solo manteniendo pregunta histórica sobre horas en una semana normal), la segunda razón de reinserción laboral, y la pregunta de tiempo máximo en un trabajo, ya que se privilegió tener el indicador de tiempo buscando empleo.

Además, por temas de poco uso, se eliminan indicadores de rama económica de empleo, ocupación, y flexibilidad salarial (como recibo y cómo le gustaría recibir su salario).

Módulo 7: Tecnologías

Este módulo se elimina debido a la incapacidad de comparabilidad dado los cambios acontecidos producto de la pandemia (niveles de respuesta y escalas hacen imposible de establecer comparaciones con versiones previas de la encuesta). Además, las parrillas eran muy extensas para las necesidades del cuestionario.

Módulo 8: Pareja y familia

Se mantienen las preguntas sobre caracterización de familia (pareja, hijos e hijas, expectativas de tenencia). Se eliminan indicadores sobre autonomía residencial, edad que dejó el hogar de origen, y percepciones de independencia residencial; principalmente por su no uso. Se mantienen las de caracterización: dónde y con quién viven; por factores de uso e historicidad.

Módulo 9: Prácticas económicas

Debido a su uso, se mantiene gran parte del módulo. Solo se modifica parrilla sobre deudas, agregando la categoría “Con personas conocidas, amistades o familiares”, debido a la idea de ser un aspecto relevante con relación al estudio de las prácticas económicas de las juventudes; cuestión también destacada en la consulta pública.

Módulo 10: Otros

Debido a no uso, se elimina la pregunta sobre las razones para migrar a Chile de migrantes internacionales. Además, si bien se elimina la pregunta respecto de la religión con la cual se identifica, se mantiene el indicador si se identifica o no con alguna religión por su relevancia en el cuestionario a nivel histórico.

Módulo 1 Autoaplicado: Consumo de sustancias

Es el módulo que más se modifica en términos del cuestionario autoaplicado. Por longitud, se eliminan las preguntas de cantidad de cigarrillos y de tragos de alcohol. Además, se acuerda eliminar la parrilla

de preguntas sobre exposición a situaciones de riesgo en el último año, debido a su incapacidad de comparar con respecto a la aplicación previa, y además que esas actividades también se redujeron por las prohibiciones de movilidad (producto de la pandemia).

Módulo 2 Autoaplicado: Vida sexual

Se incorporó la pregunta sobre haber tenido relaciones sexuales sin penetración, para complementar indicador histórico sobre iniciación sexual penetrativa. Además, se cambia ligeramente el fraseo en las preguntas sobre uso de mecanismos de protección con la pareja sexual (antes era solo para parejas).

Módulo 3 Autoaplicado: Salud mental

Este módulo sufre la única variación al cambiar la parrilla de sentimientos de salud mental por el indicador PQH-9, validado a nivel internacional y promovido por redes de expertos y expertas en investigación juvenil. Esto afectará comparar con indicadores de salud mental de aplicaciones anteriores.

Módulo 4 Autoaplicado: Violencia

En este módulo se eliminan las preguntas sobre vivencia de situaciones de violencia física y psicológica el último año, por no uso. Se agrega indicador sobre cyberbullying (“tomar pantallazos sin consentimiento”). Todo el resto se mantiene.

3.1. Estructura cuestionario población joven

Modalidad	N°	Módulos temáticos	Preguntas
Cara a cara	A	Selección y sorteo entrevista	A01 – A08
	1	Participación sociopolítica	P1 – P4
	2	Representaciones y orientaciones valóricas	P5 – P12
	3	Educación	P13 – P21
	4	Política y procesos de integración al sistema democrático	P22 – P30
	5	Confianza y discriminación	P31 – P34
	6	Trabajo	P35 – P52
	7	Pareja y familia	P53 – P58
	8	Prácticas económicas	P59 – P62
	9	Otros	P63 – P69
Autoaplicado	B	Datos de clasificación social	P70-P75
	1	Consumo de sustancias	P76 – P77
	2	Vida sexual	P78 – P101
	3	Salud mental	P102 – P105
	4	Violencia	P106 – P112

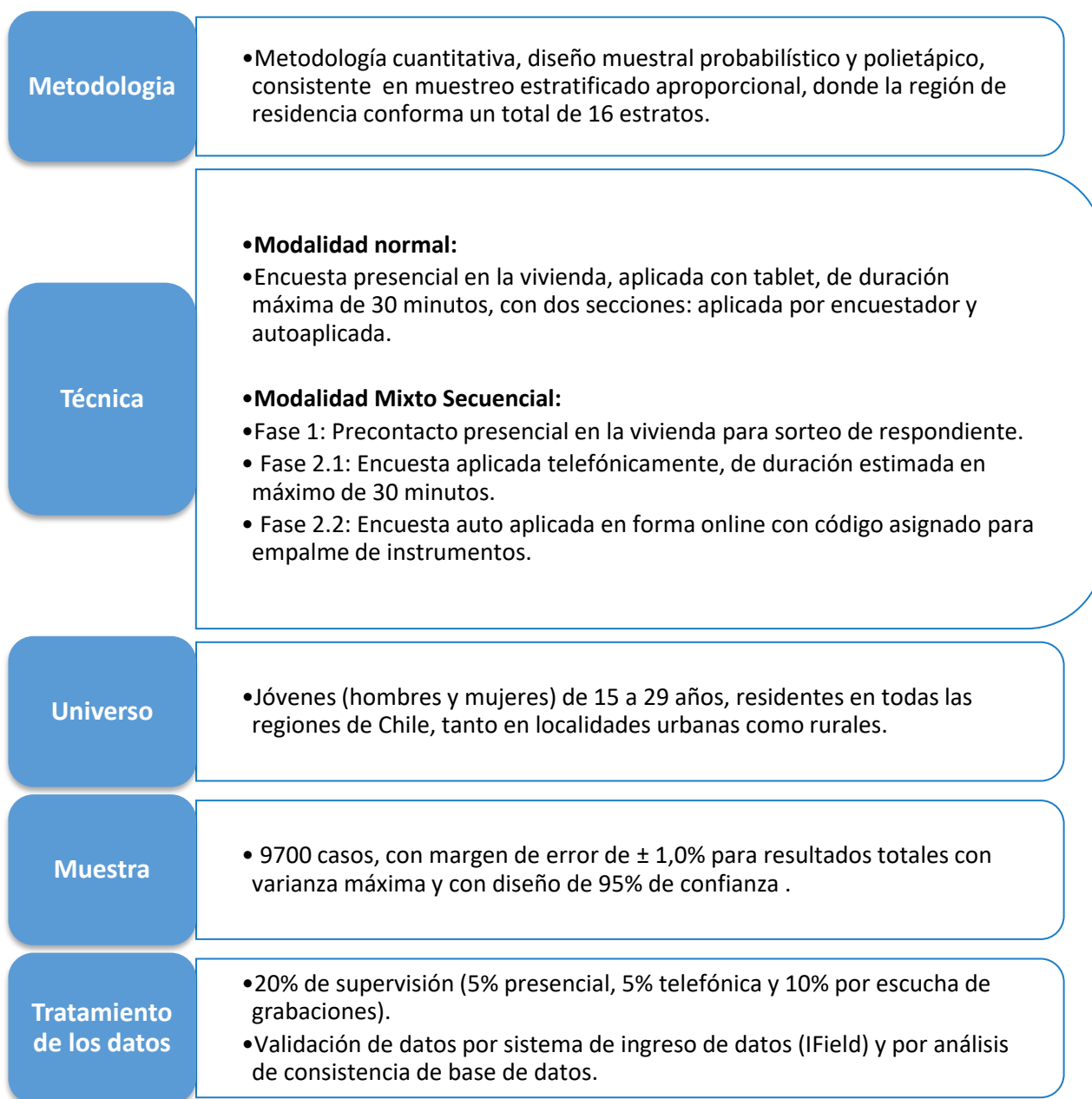
3.2. Estructura cuestionario población adulta

Modalidad	N°	Módulos temáticos	Preguntas
Cara a cara	A	Selección y sorteo entrevista	A01 – A08
	1	Participación sociopolítica	P1 – P3
	2	Representaciones y orientaciones valóricas	P4 – P11
	3	Educación	P12 – P15
	4	Política y procesos de integración al sistema democrático	P16 – P22
	5	Confianza y discriminación	P23 – P24
	6	Trabajo	P25 – P37
	7	Pareja y familia	P38 – P42
	8	Prácticas económicas	P43 – P44
	9	Otros	P45 – P50
	B	Datos de clasificación social	P51 – P56
Autoaplicado	1	Consumo de sustancias	P57 – P58
	2	Vida sexual	P59 – P76
	3	Salud mental	P77 – P79
	4	Violencia	P80 – P84

4. Encuesta a población joven

Tal como se mencionó anteriormente, debido a que las condiciones sanitarias impedían el normal desarrollo del levantamiento del trabajo de campo, la 10ma. Encuesta Nacional de Juventudes 2022 contó con la posibilidad de realizar un levantamiento vía remoto, con una metodología mixta secuencial. Así, tal como otras encuestas de hogares como la CASEN o la Encuesta Nacional de Empleo, se permitió que la persona que encuestaba contactara a la persona seleccionada mediante llamada telefónica, luego de los procesos de sorteos probabilísticos respectivos. De todas formas, debido a que las condiciones sanitarias mejoraron en el transcurso del trabajo de campo, esta opción solo fue utilizada por una fracción de la muestra de jóvenes (636 casos, correspondientes al 6,6% del total de la muestra en jóvenes). De esta forma, y a modo de resumen de la sección, se presenta en el siguiente cuadro las características metodológicas de la muestra de la población joven:

Figura 1. Resumen metodológico de la 10ma. Encuesta de Juventudes.



Fuente: Décima Encuesta Nacional de Juventudes 2022.

4.1. Diseño muestral población joven

El diseño muestral corresponde a un muestreo estratificado aporportional. Esta definición incluye las siguientes características:

1. **Diseño muestral probabilístico:** Para el objetivo del estudio fue obligatorio contar con un diseño muestral de orden probabilístico que permita condiciones de muestreo habilitantes para la estimación de parámetros (resultados atribuibles al universo). Este tipo de diseño implicó la selección aleatoria en todas las etapas del muestreo, descartando muestreos por cuotas o similares.

2. Diseño muestral polietápico: Se trabajó con tres unidades de muestreo, manzanas, viviendas y respondientes, por tanto, se aplicaron tres sorteos aleatorios por cada encuesta aplicada.
3. Muestreo estratificado: El muestreo estratificado implica que, para facilitar el abordaje del universo, se clasificó a la población en grupos homogéneos en función de variables clasificatorias. Estos grupos son denominados estratos. Para efectos de este estudio, se trabajó con 16 estratos o grupos homogéneos que se obtienen por la región de residencia. Sin embargo, en cada estrato se consideraron dos zonas o sectores, urbano y rural, para una mejor administración de la muestra. Los estratos a medir fueron:

Cuadro 1. Definición de estratos Encuesta Nacional de juventudes

		Estrato	Zona Urbana	Zona Rural
Región	Arica y Parinacota	E1	E1.1	E1.2
	Tarapacá	E2	E2.1	E2.2
	Antofagasta	E3	E3.1	E3.2
	Atacama	E4	E4.1	E4.2
	Coquimbo	E5	E5.1	E5.2
	Valparaíso	E6	E6.1	E6.2
	O'Higgins	E7	E7.1	E7.2
	Maule	E8	E8.1	E8.2
	Ñuble	E9	E9.1	E9.2
	Biobío	E10	E10.1	E10.2
	Araucanía	E11	E11.1	E11.2
	Los Ríos	E12	E12.1	E12.2
	Los Lagos	E13	E13.1	E13.2
	Aysén	E14	E14.1	E14.2
	Magallanes	E15	E15.1	E15.2
	RM	E16	E16.1	E16.2
TOTAL DE ESTRATOS		16	16	16

Fuente: Décima Encuesta Nacional de la Juventudes, 2022

4. Opción aporporcional: Para aplicar muestreo estratificado es posible optar entre la opción proporcional y la aporporcional. La primera opción implica que cada uno de los estratos tiene una muestra proporcional a su peso en el universo, no necesitando ponderación posterior de la muestra (al menos por la variable estrato). Si bien en principio se aplicó el ejercicio de cálculo de muestra por región -como se observa en el siguiente cuadro-, el resultado indicaba que los estratos de muy bajo peso quedarían con una muestra con alto margen de error y que no habilitaba para análisis a nivel de esa unidad. Por ejemplo, si la muestra de 9.700 casos propuesta por Ipsos se repartiera proporcionalmente, la Región Metropolitana contaría con 4.231 casos, mientras que la muestra de la Región de Aysén debería participar con solo 46 casos. Además, no se cumpliría con el requisito solicitado por bases técnicas de margen de error muestral máximo de 5% para cada región.

Cuadro 2. Ejercicio de cálculo de muestra proporcional según estrato región

REGION	Total universo	% Universo	Muestra proporcional	Margen de error
15 Región de Arica	58.997	1,38%	133	8,49%
01 Región de Tarapacá	94.386	2,20%	213	6,70%
02 Región de Antofagasta	172.545	4,02%	390	4,96%
03 Región de Atacama	66.373	1,55%	150	7,99%
04 Región de Coquimbo	179.874	4,19%	407	4,86%
05 Región de Valparaíso	426.515	9,94%	965	3,15%
06 Región del Lib. Gral Bdo. O'Higgins	190.959	4,45%	432	4,72%
07 Región del Maule	229.389	5,35%	519	4,30%
16 Región del Ñuble	100.080	2,33%	226	6,52%
08 Región del Biobío	359.251	8,38%	813	3,44%
09 Región de la Araucanía	212.523	4,96%	481	4,47%
14 Región de Los Ríos	86.952	2,03%	197	6,97%
10 Región de Los Lagos	181.882	4,24%	411	4,83%
11 Región de Aysén	20.218	0,47%	46	14,43%
12 Región de Magallanes	37.987	0,89%	86	10,56%
13 Región Metropolitana de Santiago	1.870.874	43,62%	4.231	1,51%
TOTAL	4.288.805	100%	9.700	1,00%

Fuente: Décima Encuesta Nacional de Juventudes 2022.

Por otra parte, la opción aporportional considera muestras prefijadas bajo algún criterio previo (por ejemplo, en este caso, tamaño del estrato y margen de error), pero implican obligatoriamente la ponderación de cada segmento en el universo cuando se realice el análisis de la información. No obstante, aunque el diseño muestral sea proporcional, la mayoría de las veces se termina ponderando igualmente para corregir distorsiones producto de muestreo en variable como sexo, edad, grupo socioeconómico, entre otras.

Un diseño de muestreo estratificado aporportional implica que las selecciones son aleatorias dentro de cada estrato. Por tanto, se asegurará sorteo de manzanas, hogares y respondientes de forma aleatoria.

4.2. Universo y tamaño muestral encuesta a población joven

El universo se ha definido como jóvenes (hombres y mujeres) de 15 a 29 años, de todos los grupos socioeconómicos, residentes en todas las regiones de Chile, tanto en localidades urbanas como rurales. Para efectos de dimensionar el tamaño del universo, se consideraron los datos del Censo 2017 proyectados para al año 2022, correspondiente a la versión alojada en la página web del Instituto Nacional de Estadísticas.

El tamaño del universo alcanza a 4.288.805 jóvenes, distribuyéndose por región de la siguiente forma:

Cuadro 3. Distribución del universo según región y localidad (urbano – rural)

REGION	Universo urbano	Universo rural	Total universo	% Urbano	% Rural
15 Región de Arica	53.610	5.387	58.997	90,87%	9,13%
01 Región de Tarapacá	89.943	4.443	94.386	95,29%	4,71%
02 Región de Antofagasta	167.354	5.191	172.545	96,99%	3,01%
03 Región de Atacama	61.872	4.501	66.373	93,22%	6,78%
04 Región de Coquimbo	153.216	26.658	179.874	85,18%	14,82%
05 Región de Valparaíso	396.707	29.808	426.515	93,01%	6,99%
06 Región del Lib. Gral. Bdo. O'Higgins	147.861	43.098	190.959	77,43%	22,57%
07 Región del Maule	179.392	49.997	229.389	78,20%	21,80%
16 Región del Ñuble	75.556	24.524	100.080	75,50%	24,50%
08 Región del Biobío	327.522	31.729	359.251	91,17%	8,83%
09 Región de la Araucanía	164.065	48.458	212.523	77,20%	22,80%
14 Región de Los Ríos	68.449	18.503	86.952	78,72%	21,28%
10 Región de Los Lagos	145.834	36.048	181.882	80,18%	19,82%
11 Región de Aysén	17.390	2.828	20.218	86,01%	13,99%
12 Región de Magallanes	34.800	3.187	37.987	91,61%	8,39%
13 Región Metropolitana de Santiago	1.809.403	61.471	1.870.874	96,71%	3,29%
TOTAL GENERAL	3.892.974	395.831	4.288.805	90,77%	9,23%

Fuente: Décima Encuesta Nacional de Juventudes 2022.

Para representar al universo, la muestra propuesta alcanza a 9.700 casos, lo que implica margen de error máximo de $\pm 1,00\%$ para resultados totales estimados a nivel de confianza de 95%.

El cálculo de la muestra fue obtenido mediante el siguiente procedimiento:

1. Determinación de tamaño muestral óptimo para una muestra nacional en relación al tipo de estudio y su objetivo.
2. Cálculo del margen de error asociado a la muestra propuesta.
3. Organización de las regiones en segmentos según margen de error deseado:
 - Margen de error inferior a 2% para regiones que representen sobre el 30% del universo.
 - Margen de error inferior a 3% para regiones que representen sobre el 8% y 29% del universo.
 - Margen de error inferior a 5% para regiones que representen sobre el 1% y el 7% del universo.
 - Asignación de la muestra a cada segmento de regiones.

De esta forma, se replica el procedimiento y los márgenes de error logrados tanto en la Octava como en la Novena Encuesta Nacional de Juventudes, así como el crecimiento de la muestra nacional por agregación de la nueva Región del Ñuble de la 9na. ENJUV. Siguiendo el procedimiento antes descrito, la muestra calculada se distribuyó de la siguiente forma:

Cuadro 4. Muestra propuesta según región y localidad (urbano – rural)

REGION	Total universo	% Urbano	% Rural	Muestra urbana	Muestra rural	Muestra total	Margen de error
15 Región de Arica	58.997	90,87%	9,13%	363	37	400	4,88%
01 Región de Tarapacá	94.386	95,29%	4,71%	381	19	400	4,89%
02 Región de Antofagasta	172.545	96,99%	3,01%	388	12	400	4,90%
03 Región de Atacama	66.373	93,22%	6,78%	373	27	400	4,89%
04 Región de Coquimbo	179.874	85,18%	14,82%	341	59	400	4,90%
05 Región de Valparaíso	426.515	93,01%	6,99%	930	70	1.000	3,10%
06 Región del Lib. Gral. Bdo. O'Higgins	190.959	77,43%	22,57%	310	90	400	4,90%
07 Región del Maule	229.389	78,20%	21,80%	313	87	400	4,90%
16 Región del Ñuble	100.080	75,50%	24,50%	302	98	400	4,90%
08 Región del Biobío	359.251	91,17%	8,83%	912	88	1.000	3,10%
09 Región de la Araucanía	212.523	77,20%	22,80%	309	91	400	4,90%
14 Región de Los Ríos	86.952	78,72%	21,28%	315	85	400	4,89%
10 Región de Los Lagos	181.882	80,18%	19,82%	321	79	400	4,90%
11 Región de Aysén	20.218	86,01%	13,99%	344	56	400	4,85%
12 Región de Magallanes	37.987	91,61%	8,39%	366	34	400	4,87%
13 Región Metropolitana de Santiago	1.870.874	96,71%	3,29%	2418	82	2.500	1,96%
TOTAL GENERAL	4.288.805	90,77%	9,23%	8.686	1.014	9.700	1,00%
Margen de error	--	--	--	1,05%	3,08%	1,00%	

Fuente: Décima Encuesta Nacional de Juventudes 2022.

Por otra parte, para el cálculo del margen de error se ocupó la fórmula correspondiente a universos infinitos (tamaño de 100 mil casos o más), siendo la siguiente:

Figura 1. Fórmula de cálculo de tamaño muestral en universos finitos

$$E = \sqrt{\frac{Z^2 * P * Q}{n}}$$

Donde:

- E = Margen de error deseado para la extrapolación de resultados.
- n = Muestra estimada para supuestos del cálculo.
- Z² = Probabilidad de error en el cálculo del parámetro en base al estadígrafo. Se adoptó el alpha de 0,05 o nivel de confianza de 95% en cada cálculo.
- PQ ó p(1-p) = Varianza estimada de resultados. Se adoptó escenario de varianza máxima (V = 0,25).

Mientras que, en el caso de los universos finitos, es decir, los que se encuentran bajo los cien mil casos, se ocupó la siguiente fórmula.

Figura 2. Fórmula de cálculo de tamaño muestral en universos infinitos

$$E = z \sqrt{\frac{p(1-p)}{n} \left(1 - \frac{n}{N}\right)}$$

Donde:

- E = Margen de error deseado para la extrapolación de resultados.
- n = Muestra estimada en formula de universo infinito.
- p(1-p) = Varianza estimada de resultados. Se adoptó escenario de varianza máxima (V = 0,25).
- N = Tamaño del universo.

El siguiente paso fue la determinación de las comunas incluidas en el levantamiento de la muestra. Para tal efecto se procedió a revisar la información del INE en relación a los tamaños de cada comuna y su distribución urbana y rural. A continuación, se seleccionaron las principales comunas (con mayor concentración de población) y luego se analizó si esas comunas permitían el abordaje de la muestra rural o era necesario incorporar comunas con mayor presencia de ruralidad. Además, se consideró como criterio incluir la totalidad de las comunas en las que se aplicaron encuestas en tanto en la Octava como en la *Novena Encuesta Nacional de Juventud*. La muestra de comunas seleccionadas fue la siguiente:

Cuadro 5. Muestra de comunas seleccionadas y cobertura del marco muestral

Región	Comunas muestra urbana	Comunas muestra rural	Población incluida	Población regional	Cobertura
15 Región de Arica	Arica	Arica	57.773	58.997	97,93%
01 Región de Tarapacá	Iquique - Alto Hospicio	Pozo Almonte - Huara	92.598	94.386	98,11%
02 Región de Antofagasta	Antofagasta - Calama - Tocopilla	Calama - Tocopilla	161.393	172.545	93,54%
03 Región de Atacama	Copiapó - Caldera - Tierra Amarilla - Vallenar	Copiapó - Tierra Amarilla - Vallenar - Freirina	58.222	66.373	87,72%
04 Región de Coquimbo	Coquimbo - La Serena - Ovalle	Coquimbo - La Serena - Ovalle - Vicuña - Illapel - Los Vilos	158.221	179.874	87,96%
05 Región de Valparaíso	Valparaíso - Viña del Mar - San Antonio - Quilpué - Villa Alemana - Concón - La Calera	Casablanca - Quintero - Los Andes - La Ligua - Quillota - San Felipe - Limache	353.835	426.515	82,96%
06 Región del Lib. Gral. Bernardo O'Higgins	Rancagua - Doñihue - Machalí - Rengo - Graneros - Santa Cruz - Mostazal - San Fernando	Las Cabras - Pichidegua - Rengo - Requínoa - San Vicente - Chimbarongo - San Fernando - Santa Cruz	149.487	190.959	78,28%

07 Región del Maule	Talca - Linares - Constitución - Maule - Curicó	Talca - Linares - San Clemente - Teno - Curicó - Maule - Constitución - Molina - Parral - San Javier - Cauquenes	182.780	229.389	79,68%
16 Región del Ñuble	Chillán - San Carlos - Bulnes - Chillán Viejo	Chillán - San Carlos - Bulnes - Chillán Viejo	64.561	100.080	64,51%
08 Región del Biobío	Concepción - Talcahuano - Los Angeles - San Pedro de la Paz - Chillán - Hualpén - Chiguayante - Coronel - Lota - Penco - Tomé	Los Angeles - San Carlos - Tomé - Arauco - Mulchén - Nacimiento	294.544	359.251	81,99%
09 Región de la Araucanía	Temuco - Padre Las Casas - Angol - Victoria - Villarrica - Lautaro	Temuco - Padre Las Casas - Angol - Victoria - Lautaro - Villarrica - Lautaro - Nueva Imperial - Pucón - Vilcún	150.094	212.523	70,62%
14 Región de Los Ríos	Valdivia - Panguipulli - La Unión	Panguipulli - Río Bueno - Los Lagos - La Unión	67.716	86.952	77,88%
10 Región de Los Lagos	Puerto Montt - Osorno - Castro - Ancud - Llanquihue - Puerto Varas	Puerto Montt - Osorno - Castro - Ancud - Calbuco - Puerto Varas - Río Negro	137.764	181.882	75,74%
11 Región de Aysén	Coyhaique - Aysén	Coyhaique - Aysén	16.607	20.218	82,14%
12 Región de Magallanes	Punta Arenas - Puerto Natales	Punta Arenas - Puerto Natales	35.023	37.987	92,20%
13 Región Metropolitana de Santiago	Gran Santiago	Melipilla - Colina - Lampa - Buin - Paine - Talagante - Peñaflor	1.799.805	1.870.874	96,20%
Total			3.780.423	4.288.805	88,15%

Fuente: Décima Encuesta Nacional de Juventudes 2022.

La muestra de comunas propuesta alcanza a un nivel de cobertura del 88,15%, en función del tamaño del universo contenido en el marco muestral efectivo en relación al marco muestral teórico.

Posteriormente a la selección de comunas, se procedió a distribuir la muestra urbana y rural según los criterios antes presentados: división proporcional de la muestra regional en urbana y rural, y, a su vez, en cada segmento se distribuyó proporcionalmente la muestra según población de la zona de interés (urbana/rural) entre las comunas incluidas en la muestra. Por lo tanto, la muestra por comuna fue la siguiente:

Cuadro 6. Muestra urbana y rural de personas jóvenes por comuna

Región	Comuna	Jóvenes		
		Muestra urbana	Muestra rural	TOTAL
XV Arica y Parinacota	Arica	363	37	400
I Tarapacá	Iquique	233	6	239
	Alto Hospicio	148	0	148
	Pozo Almonte	0	9	9
	Huara	0	4	4
II Antofagasta	Antofagasta	264	0	264
	Calama	109	7	116
	Tocopilla	15	5	20
III Atacama	Copiapó	262	5	267
	Caldera	25	0	25
	Tierra Amarilla	17	8	25
	Vallenar	68	10	78
	Freirina	0	5	5
IV Coquimbo	La Serena	142	15	157
	Coquimbo	143	10	153
	Vicuña	0	8	8
	Illapel	0	6	6
	Los Vilos	0	4	4
	Ovalle	56	16	72
V Valparaíso	Valparaíso	268	1	269
	Casablanca	0	10	10
	Concón	32	4	36
	Quintero	0	5	5
	Viña del Mar	306	0	306
	Los Andes	0	7	7
	La Ligua	0	10	10
	Quillota	0	16	16
	La Calera	39	0	39
	San Antonio	65	0	65
	San Felipe	0	8	8
	Quilpué	120	0	120
	Limache	0	8	8
	Villa Alemana	100	0	100
VI O'Higgins	Rancagua	157	0	157
	Doñihue	11	0	11
	Graneros	20	0	20
	Las Cabras	0	12	12
	Machalí	30	0	30
	Mostazal	14	0	14

	Pichidegua	6	8	14
	Rengo	31	11	42
	Requínoa	0	10	10
	San Vicente	0	16	16
	San Fernando	41	8	49
	Chimbarongo	0	13	13
	Santa Cruz	0	12	12
VII Maule	Talca	144	6	150
	Constitución	22	6	28
	Maule	24	7	31
	San Clemente	0	14	14
	Cauquenes	0	3	3
	Curicó	80	10	90
	Molina	0	5	5
	Teno	0	13	13
	Linares	44	10	54
	Parral	0	6	6
	San Javier	0	8	8
Ñuble	Chillán	212	34	246
	San Carlos	40	40	80
	Bulnes	15	16	31
	Chillán Viejo	35	7	42
VIII Biobío	Concepción	208	0	208
	Coronel	86	0	86
	Chiguayante	68	0	68
	Lota	32	0	32
	Penco	35	0	35
	San Pedro de la Paz	98	0	98
	Talcahuano	115	0	115
	Tomé	35	88	123
	Hualpén	69	0	69
	Arauco	20	0	20
	Los Ángeles	113	0	113
	Mulchén	18	0	18
	Nacimiento	15	0	15
IX Araucanía	Temuco	192	15	207
	Lautaro	18	7	25
	Nueva Imperial	0	8	8
	Padre Las Casas	31	25	56
	Pucón	0	6	6
	Vilcún	0	9	9
	Villarrica	24	12	36
	Angol	29	3	32
	Victoria	15	6	21

XIV Los Ríos	Valdivia	256	0	256
	Los Lagos	0	16	16
	Panguipulli	22	32	54
	La Unión	37	17	54
	Río Bueno	0	20	20
X Los Lagos	Puerto Montt	158	23	181
	Calbuco	0	15	15
	Llanquihue	9	0	9
	Puerto Varas	20	10	30
	Castro	21	7	28
	Ancud	18	7	25
	Osorno	95	11	106
	Río Negro	0	6	6
XI Aysén	Coyhaique	245	38	283
	Aysén	99	18	117
XII Magallanes y la Antártica	Punta Arenas	320	25	345
	Natales	46	9	55
Región Metropolitana	Santiago	205	0	205
	Cerrillos	31	0	31
	Cerro Navia	52	0	52
	Conchalí	48	0	48
	El Bosque	63	0	63
	Estación Central	63	0	63
	Huechuraba	36	0	36
	Independencia	44	0	44
	La Cisterna	34	0	34
	La Florida	137	0	137
	La Granja	45	0	45
	La Pintana	71	0	71
	La Reina	33	0	33
	Las Condes	111	0	111
	Lo Barnechea	41	0	41
	Lo Espejo	40	0	40
	Lo Prado	37	0	37
	Macul	43	0	43
	Maipú	207	0	207
	Ñuñoa	69	0	69
	Pedro Aguirre Cerda	38	0	38
	Peñalolén	95	0	95
	Providencia	50	0	50
	Pudahuel	90	0	90
	Quilicura	88	0	88
	Quinta Normal	43	0	43
Recoleta	62	0	62	

Renca	57	0	57
San Joaquín	36	0	36
San Miguel	40	0	40
San Ramón	32	0	32
Vitacura	27	0	27
Puente Alto	231	0	231
Colina	0	15	15
Lampa	0	12	12
San Bernardo	119	0	119
Buín	0	8	8
Paine	0	14	14
Melipilla	0	21	21
Talagante	0	8	8
Peñaflor	0	4	4
TOTAL	8.686	1.014	9.700

Fuente: Décima Encuesta Nacional de Juventudes 2022.

5. Encuesta a población adulta

Para la realización de la Décima Encuesta Nacional de Juventudes se ha planeado por tercera vez la aplicación de una muestra de control de adultos de 30 a 59 años. Esta muestra es equivalente en cuanto a distribución y representatividad a la muestra de jóvenes, pero con tamaño pertinente a una muestra de control. Al igual que la población joven, la encuesta a la población adulta se aplicó con una metodología normal, cara a cara, y con una metodología mixto-secuencial. De todas formas, debido a que las condiciones sanitarias mejoraron en el transcurso del trabajo de campo (diciembre de 2021 a mayo de 2022), esta última opción solo fue utilizada por una pequeña porción de la muestra en adultos (4,7% del total de la muestra en adultos).

La muestra de adultos alcanza a 1000 casos, correspondiendo al 10,31% de la muestra de jóvenes. Las principales características asociadas a la muestra fueron las siguientes:

Figura 1. Resumen metodológico de la encuesta de adultos

Metodología	<ul style="list-style-type: none"> • Metodología cuantitativa, diseño muestral probabilístico y polietápico consistente en muestreo estratificado aporportional, donde la región conforma 16 estratos.
Técnica	<ul style="list-style-type: none"> • Modalidad normal: • Encuesta presencial en la vivienda, aplicada con tablet, de duración máxima de 30 minutos, con dos secciones: aplicada por el encuestador y autoaplicada. • Modalidad Mixtp Secuencial: • Fase 1: Precontacto presencial en la vivienda para sorteo de respondiente. • Fase 2.1: Encuesta aplicada telefónicamente, de duración estimada en máximo de 30 minutos. • Fase 2.2: Encuesta auto aplicada en forma online con código asignado para empalme de instrumentos.
Universo	<ul style="list-style-type: none"> • Adultos (hombres y mujeres) de 30 a 59 años, residentes en todas las regiones de Chile, tanto en localidades urbanas como rurales.
Muestra	<ul style="list-style-type: none"> • 1.000 casos, con margen de error de $\pm 3,10\%$ para resultados totales con varianza máxima y con diseño de 95% de confianza .
Tratamiento de los datos	<ul style="list-style-type: none"> • 20% de supervisión (5% presencial, 5% telefónica y 10% por escucha de grabaciones). • Validación de datos por sistema de ingreso de datos (IField) y por análisis de consistencia de base de datos.

Fuente: Décima Encuesta Nacional de la Juventudes, 2022

5.1. Diseño muestral población adulta

Para operar efectivamente como una muestra de control respecto de la Encuesta de Juventudes, el diseño muestral debió ser equivalente en todas sus fases a la cadena de decisiones muestrales adoptadas en la primera encuesta o grupo de estudio principal (jóvenes). Estas decisiones ya han sido explicadas en las secciones previas.

5.2. Universo y tamaño muestral encuesta a población adulta

El universo se ha definido como adultos (hombres y mujeres) de 30 a 59 años, residentes en todas las regiones de Chile, tanto en localidades urbanas como rurales. Al igual que en el caso de la encuesta de juventudes, el cálculo del tamaño del universo se realizó con los datos del Censo 2017 proyectados al año 2022, alcanzado a un total de 8.171.650 adultos, distribuyéndose por región de la siguiente forma:

Cuadro 7. Distribución del universo de adultos según región y localidad (urbano – rural)

REGION	Universo urbano	Universo rural	Total universo	% Urbano	% Rural
15 Región de Arica	92.941	8.716	101.657	91,43%	8,57%
01 Región de Tarapacá	149.757	9.258	159.015	94,18%	5,82%
02 Región de Antofagasta	284.332	13.291	297.623	95,53%	4,47%
03 Región de Atacama	116.868	11.706	128.574	90,90%	9,10%
04 Región de Coquimbo	277.616	63.357	340.973	81,42%	18,58%
05 Región de Valparaíso	713.164	68.724	781.888	91,21%	8,79%
06 Región del Lib. Gral. Bdo. O'Higgins	315.554	105.608	421.162	74,92%	25,08%
07 Región del Maule	344.811	122.045	466.856	73,86%	26,14%
16 Región del Ñuble	147.074	61.747	208.821	70,43%	29,57%
08 Región del Biobío	605.939	75.815	681.754	88,88%	11,12%
09 Región de la Araucanía	292.695	114.602	407.297	71,86%	28,14%
14 Región de Los Ríos	117.751	46.925	164.676	71,50%	28,50%
10 Región de Los Lagos	283.800	99.683	383.483	74,01%	25,99%
11 Región de Aysén	38.116	9.091	47.207	80,74%	19,26%
12 Región de Magallanes	69.672	6.705	76.377	91,22%	8,78%
13 Región Metropolitana de Santiago	3.374.021	130.266	3.504.287	96,28%	3,72%
TOTAL GENERAL	7.224.111	947.539	8.171.650	88,40%	11,60%

Fuente: Décima Encuesta Nacional de Juventudes 2022.

La muestra de control propuesta fue de 1.000 casos, lo que implica un margen de error máximo de $\pm 3,10\%$ para resultados totales estimados a nivel de confianza de 95%. La muestra fue obtenida considerando el procedimiento aplicado para la muestra de la encuesta a jóvenes que se resume en:

1. Determinación de tamaño muestral óptimo para una muestra nacional de control.
2. Cálculo del margen de error asociado a la muestra propuesta.
3. Organización de las regiones en segmentos según margen de error.
4. Asignación de la muestra a cada segmento de regiones.
5. Distribución de muestra por estrato urbano y rural. En este punto diferimos de la muestra de la encuesta a jóvenes, la que distribuyó la muestra en cada región según la distribución proporcional del universo por localidad. La cantidad de encuestas de cada región en la muestra de control es de menor tamaño que la muestra de jóvenes, lo que implica que al distribuirla proporcionalmente se obtienen cuotas rurales de muy bajo tamaño (en 9 regiones serían cuotas inferiores a 10 casos). Por lo que se optó por acercarse a la proporcionalidad, pero con una muestra rural que no fuese inferior a 15 casos.

El cuadro de conformación de la muestra para la encuesta a adultos es el siguiente:

Cuadro 8. Muestra propuesta según región y localidad (urbano – rural)

REGION	Total universo	% Urbano	% Rural	Muestra urbana	Muestra rural	Muestra total	Margen de error
15 Región de Arica	101.657	91,43%	8,57%	35	15	50	13,86%
01 Región de Tarapacá	159.015	94,18%	5,82%	35	15	50	13,86%
02 Región de Antofagasta	297.623	95,53%	4,47%	35	15	50	13,86%
03 Región de Atacama	128.574	90,90%	9,10%	35	15	50	13,86%
04 Región de Coquimbo	340.973	81,42%	18,58%	35	15	50	13,86%
05 Región de Valparaíso	781.888	91,21%	8,79%	60	15	75	11,32%
06 Región del Lib. Gral. Bdo. O'Higgins	421.162	74,92%	25,08%	35	15	50	13,86%
07 Región del Maule	466.856	73,86%	26,14%	35	15	50	13,86%
16 Región del Ñuble	208.821	70,43%	29,57%	35	15	50	13,86%
08 Región del Biobío	681.754	88,88%	11,12%	60	15	75	11,32%
09 Región de la Araucanía	407.297	71,86%	28,14%	35	15	50	13,86%
14 Región de Los Ríos	164.676	71,50%	28,50%	35	15	50	13,86%
10 Región de Los Lagos	383.483	74,01%	25,99%	35	15	50	13,86%
11 Región de Aysén	47.207	80,74%	19,26%	35	15	50	13,85%
12 Región de Magallanes	76.377	91,22%	8,78%	35	15	50	13,86%
13 Región Metropolitana de Santiago	3.504.287	96,28%	3,72%	185	15	200	6,93%
TOTAL GENERAL	8.171.650	88,40%	11,60%	760	240	1.000	3,10%
Margen de error	--	--	--	3,55%	6,33%	3,10%	

Fuente: Décima Encuesta Nacional de Juventudes 2022.

El cálculo del margen de error fue realizado según las fórmulas para universos infinitos y finitos. Por otra parte, respecto de las comunas seleccionadas para la toma de muestras urbanas y rurales, para efectos de comparabilidad con la encuesta de jóvenes, se deberán aplicar las encuestas en las mismas comunas en que se aplique la muestra principal, siendo estas las siguientes:

Cuadro 9. Muestra de comunas seleccionadas y cobertura del marco muestral

Región	Comunas muestra urbana	Comunas muestra rural	Población incluida	Población regional	Cobertura
15 Región de Arica	Arica	Arica	99.869	101.657	98,24%
01 Región de Tarapacá	Iquique - Alto Hospicio	Pozo Almonte - Huara	155.612	159.015	97,86%
02 Región de Antofagasta	Antofagasta - Calama - Tocopilla	Calama - Tocopilla	276.361	297.623	92,86%
03 Región de Atacama	Copiapó - Caldera - Tierra Amarilla - Vallenar	Copiapó - Tierra Amarilla - Vallenar - Freirina	110.230	128.574	85,73%
04 Región de Coquimbo	Coquimbo - La Serena - Ovalle	Coquimbo - La Serena - Ovalle - Vicuña - Illapel - Los Vilos	291.114	340.973	85,38%

05 Región de Valparaíso	Valparaíso - Viña del Mar - San Antonio - Quilpué - Villa Alemana - Concón - La Calera	Casablanca - Quintero - Los Andes - La Ligua - Quillota - San Felipe - Limache	627.149	781.888	80,21%
06 Región del Lib. Gral. Bernardo O'Higgins	Rancagua - Doñihue - Machalí - Rengo - Graneros - Santa Cruz - Mostazal - San Fernando	Las Cabras - Pichidegua - Rengo - Requinoa - San Vicente - Chimbarongo - San Fernando - Santa Cruz	323.569	421.162	76,83%
07 Región del Maule	Talca - Linares - Constitución - Maule - Curicó	Talca - Linares - San Clemente - Teno - Curicó - Maule - Constitución - Molina - Parral - San Javier - Cauquenes	360.880	466.856	77,30%
16 Región del Ñuble	Chillán - San Carlos - Bulnes - Chillán Viejo	Chillán - San Carlos - Bulnes - Chillán Viejo	127.428	208.821	61,02%
08 Región del Biobío	Concepción - Talcahuano - Los Angeles - San Pedro de la Paz - Chillán - Hualpén - Chiguayante - Coronel - Lota - Penco - Tomé	Los Angeles - San Carlos - Tomé - Arauco - Mulchén - Nacimiento	544.602	681.754	79,88%
09 Región de la Araucanía	Temuco - Padre Las Casas - Angol - Victoria - Villarrica - Lautaro	Temuco - Padre Las Casas - Angol - Victoria - Lautaro - Villarrica - Lautaro - Nueva Imperial - Pucón - Vilcún	271.639	407.297	66,69%
14 Región de Los Ríos	Valdivia - Panguipulli - La Unión	Panguipulli - Río Bueno - Los Lagos - La Unión	124.446	164.676	75,57%
10 Región de Los Lagos	Puerto Montt - Osorno - Castro - Ancud - Llanquihue - Puerto Varas	Puerto Montt - Osorno - Castro - Ancud - Calbuco - Puerto Varas - Río Negro	281.105	383.483	73,30%
11 Región de Aysén	Coyhaique - Aysén	Coyhaique - Aysén	37.644	47.207	79,74%
12 Región de Magallanes	Punta Arenas - Puerto Natales	Punta Arenas - Puerto Natales	70.728	76.377	92,60%
13 Región Metropolitana de Santiago	Gran Santiago	Melipilla - Colina - Lampa - Buin - Paine - Talagante - Peñaflor	3.365.879	3.504.287	96,05%
Total			7.068.255	8.171.650	86,50%

Fuente: Décima Encuesta Nacional de Juventudes 2022.

Finalmente, la muestra fue distribuida proporcionalmente por comuna y, posteriormente, por zona urbana y rural, conformando la siguiente distribución de la muestra:

Cuadro 10. Muestra urbana y rural de adultos por comuna

Región	Comuna	Adultos		
		Muestra urbana	Muestra rural	TOTAL
XV Arica y Parinacota	Arica	35	15	50
I Tarapacá	Iquique	23	3	26
	Alto Hospicio	12	0	12
	Pozo Almonte	0	10	10
	Huara	0	2	2
II Antofagasta	Antofagasta	22	0	22
	Calama	11	13	24
	Tocopilla	2	2	4
III Atacama	Copiapó	24	3	27
	Caldera	2	0	2
	Tierra Amarilla	2	5	7
	Vallenar	7	5	12
	Freirina	0	2	2
IV Coquimbo	La Serena	14	4	18
	Coquimbo	15	2	17
	Vicuña	0	2	2
	Illapel	0	2	2
	Los Vilos	0	1	1
	Ovalle	6	4	10
V Valparaíso	Valparaíso	16	0	16
	Casablanca	0	2	2
	Concón	3	1	4
	Quintero	0	1	1
	Viña del Mar	18	0	18
	Los Andes	0	2	2
	La Ligua	0	2	2
	Quillota	0	3	3
	La Calera	3	0	3
	San Antonio	5	0	5
	San Felipe	0	2	2
	Quilpué	8	0	8
	Limache	0	2	2
	Villa Alemana	7	0	7
VI O'Higgins	Rancagua	17	0	17
	Doñihue	1	0	1
	Graneros	2	0	2
	Las Cabras	0	2	2
	Machalí	4	0	4

	Mostazal	2	0	2
	Pichidegua	1	1	2
	Rengo	3	2	5
	Requínoa	0	2	2
	San Vicente	0	3	3
	San Fernando	5	1	6
	Chimbarongo	0	2	2
	Santa Cruz	0	2	2
VII Maule	Talca	13	1	14
	Constitución	3	1	4
	Maule	3	1	4
	San Clemente	0	2	2
	Cauquenes	0	1	1
	Curicó	10	2	12
	Molina	0	1	1
	Teno	0	2	2
	Linares	6	2	8
	Parral	0	1	1
	San Javier	0	1	1
Ñuble	Chillán	24	5	29
	San Carlos	5	6	11
	Bulnes	2	3	5
	Chillán Viejo	4	1	5
VIII Biobío	Concepción	10	0	10
	Coronel	6	0	6
	Chiguayante	4	0	4
	Lota	2	0	2
	Penco	2	0	2
	San Pedro de la Paz	7	0	7
	Talcahuano	8	0	8
	Tomé	3	15	18
	Hualpén	5	0	5
	Arauco	1	0	1
	Los Ángeles	8	0	8
	Mulchén	2	0	2
	Nacimiento	2	0	2
IX Araucanía	Temuco	20	2	22
	Lautaro	2	1	3
	Nueva Imperial	0	2	2
	Padre Las Casas	4	3	7
	Pucón	0	1	1
	Vilcún	0	2	2
	Villarrica	3	2	5
	Angol	4	1	5

	Victoria	2	1	3
XIV Los Ríos	Valdivia	25	0	25
	Los Lagos	0	3	3
	Panguipulli	3	5	8
	La Unión	4	3	7
	Río Bueno	3	4	7
X Los Lagos	Puerto Montt	16	3	19
	Calbuco	1	3	4
	Llanquihue	1	1	2
	Puerto Varas	2	2	4
	Castro	3	1	4
	Ancud	2	2	4
	Osorno	10	2	12
	Río Negro	0	1	1
XI Aysén	Coyhaique	25	10	35
	Aysén	10	5	15
XII Magallanes y la Antártica	Punta Arenas	30	10	40
	Natales	5	5	10
Región Metropolitana	Santiago	14	0	14
	Cerrillos	2	0	2
	Cerro Navia	4	0	4
	Conchalí	4	0	4
	El Bosque	5	0	5
	Estación Central	4	0	4
	Huechuraba	3	0	3
	Independencia	3	0	3
	La Cisterna	3	0	3
	La Florida	11	0	11
	La Granja	3	0	3
	La Pintana	5	0	5
	La Reina	3	0	3
	Las Condes	9	0	9
	Lo Barnechea	3	0	3
	Lo Espejo	3	0	3
	Lo Prado	3	0	3
	Macul	4	0	4
	Maipú	16	0	16
	Ñuñoa	7	0	7
	Pedro Aguirre Cerda	3	0	3
	Peñalolén	7	0	7
	Providencia	5	0	5
	Pudahuel	7	0	7
Quilicura	7	0	7	
Quinta Normal	3	0	3	

Recoleta	5	0	5
Renca	4	0	4
San Joaquín	3	0	3
San Miguel	3	0	3
San Ramón	2	0	2
Vitacura	2	0	2
Puente Alto	16	0	16
Colina	0	3	3
Lampa	0	2	2
San Bernardo	9	0	9
Buín	0	1	1
Paine	0	3	3
Melipilla	0	3	3
Talagante	0	2	2
Peñaflor	0	1	1
TOTAL	760	240	1.000

Fuente: Décima Encuesta Nacional de Juventudes 2022.

6. Marco muestral, unidades muestrales y capturas de datos en terreno

La captura de los datos en terreno operó exactamente igual para la encuesta a juventudes y para la encuesta a adultos (grupo de control), siendo un requisito necesario para la comparación de la información.

El diseño muestral propuesto implica operar con tres unidades de muestreo:

- Unidad primaria: manzanas.
- Unidad secundaria: hogares.
- Unidad terciaria: personas.

Cada unidad muestral fue abordada aleatoriamente con objeto de cautelar la condición probabilística de la muestra, como se ha indicado previamente. La forma de aplicar el sorteo se definió en función de las características de la unidad de muestreo. A continuación, se desarrollan los criterios de cada caso.

Es importante establecer que en el trabajo de muestreo hubo cuatro tipos de viviendas:

- Vivienda en zona urbana perteneciente a comuna urbana
- Vivienda en zona rural perteneciente a comuna urbana
- Vivienda en zona urbana perteneciente a comuna rural
- Vivienda en zona rural perteneciente a comuna rural

Para efectos de la muestra se consideraron como muestra de vivienda rural sólo aquellas encuestas aplicadas en viviendas pertenecientes a zonas rurales (fuera de las zonas urbanas de las comunas) independientemente que la comuna esté clasificada como rural. De esta forma se evitó que viviendas de comunas rurales pero que se encontraban en la zona urbana, y que por tanto es esperable que tenga comportamiento urbano, sean consideradas muestras rurales.

El marco muestral correspondió al conjunto de manzanas ordenadas en distritos censales en el marco muestral maestro con el que trabaja el Instituto Nacional de Estadísticas.

Tanto en el caso de las muestras urbanas como rurales se usó el marco muestral del Censo 2017, tal como se hizo en el año 2018 en la Novena Encuesta de Juventudes. Además, el marco muestral ocupado por Ipsos tiene la fortaleza de ser actualizado cada 6 meses mediante solicitud al Servicio de Impuestos Internos SII por vía de petición de transparencia de todas las direcciones de la base de propiedades con su rol activo y por destino habitacional. Esto permitió que el marco muestral se actualice oportunamente, en forma sistemática y con fuente conocida y pública de los datos. Con este marco muestral se disminuyó el error asociado a la antigüedad de los mapas del Censo 2017.

- **Primer sorteo aleatorio: sorteo de manzanas**

El primer sorteo permitió alcanzar la muestra de viviendas. Pero, en este punto fue necesario diferenciar el muestreo en función de la muestra de hogares urbanos y hogares rurales, por cuanto no es posible recorrer de igual forma ambas zonas.

Para el primer sorteo dentro de las comunas, las manzanas fueron organizadas en bases de datos administradas en SPSS en el caso de la Región Metropolitana, y en planos en el resto de las regiones, lo que permitió aplicar muestreo aleatorio en la selección. De esta forma se seleccionó una cantidad de manzanas proporcionales a la muestra a recolectar en cada región. Las muestras de manzanas fueron independientes por región y/o comuna, es decir, la elección de una manzana en una unidad geográfica no condicionó de ninguna forma la elección de una manzana en otra unidad de análisis. A medida que se necesitaron manzanas de reemplazo para alcanzar las cuotas, se realizaron nuevos sorteos en la base de datos o en los mapas correspondientes.

El muestreo se realizó en base a la probabilidad de selección según número de viviendas por manzana, de igual forma como se solicitó y ejecutó en el año 2015 y 2018 para asegurar la comparabilidad de la información.

Ahora bien, el sorteo generó dos elementos que permitió a las personas encuestadoras identificar la manzana seleccionada:

1. Mapa de terreno. El mapa que identificó claramente la manzana seleccionada y su entorno inmediato. El mapa también incluyó información necesaria para su vinculación con respuestas en la base de datos y para una mejor orientación de la persona encuestada.
2. Punto GPS. Cada mapa vino asociado a una coordenada GPS cargable por la persona encuestadora en su celular para tener el mapa online, comprobando en terreno la correcta ubicación de la manzana al comparar su posición al momento de hacer la consulta y la posición de la manzana seleccionada según coordina GPS.

Figura 3. Vista de mapa de manzana sorteada



Fuente: Décima Encuesta Nacional de Juventudes 2022.

La probabilidad de inclusión de una manzana i en la muestra de una comuna h se calculó de la siguiente forma:

Figura 4. Fórmula de probabilidad de selección de manzanas

$$f_1 = n_h * \frac{M_{hi}}{M_h}$$

Fuente: Ipsos Chile

Donde:

- f_1 = Probabilidad de selección de la manzana.
- n_h = Numero de manzanas a seleccionar en la comuna h .
- M_{hi} = Viviendas de la manzana i de la comuna h .
- M_h = Viviendas en la comuna h .

Criterio de reemplazo de manzanas urbanas:

Las manzanas fueron reemplazadas por una nueva selección en la base de datos sólo cuando se trató de una unidad que tenía un parque o una zona industrial. De todas formas, se sorteó un 20% adicional de manzanas por región y comuna por si hubiese sido necesario reemplazar un punto muestral completo o que éste no permitiera la obtención de las 8 encuestas a lograr.

Sorteo en zonas residenciales rurales. Considerando la cuota de encuestas rurales en cada región, ésta se repartió entre las zonas rurales de las comunas seleccionadas para el estudio. Del marco muestral teórico a utilizar en las localidades rurales definidas por el Instituto Nacional de Estadística, se construyó un segundo marco muestral efectivo generado en base a las localidades rurales de las comunas seleccionadas para la muestra rural. Sobre el marco muestral efectivo se procedió al sorteo de localidades en cada comuna. Una vez sorteados los puntos muestrales, le fue asignada una cuota de encuestas proporcional al tamaño.

Criterio de reemplazo de zonas residenciales rurales. Una localidad rural fue reemplazada sólo cuando una vez en terreno se observó que el sector se encontraba deshabitado o sin población suficiente para lograr la muestra necesaria.

- Segundo sorteo aleatorio: sorteo de viviendas

Sorteo de viviendas en zonas urbanas. A modo de valor agregado para el estudio, se implementó la selección de viviendas en base a manzanas empadronadas de la base de datos de Ipsos, lo que permitió avanzar con mayor rapidez y menor margen de error en terreno.

Para todas las comunas urbanas, para todas las manzanas del marco muestral del Censo 2017 y para todas las manzanas habitadas nuevas generadas en forma posterior al Censo 2017, Ipsos ya contaba con la información de todas las direcciones de cada manzana. Y esta información fue proporcionada por una fuente oficial como es el Servicio de Impuesto Internos. De esta forma, Ipsos estaba en condiciones de sortear las viviendas en la oficina y sus reemplazos, entregando esta información ya lista a cada persona encuestadora para concurrir a aplicar directamente el sorteo de respondiente en la fase de precontacto.

Este sorteo en oficinas en base al empadronamiento se realizó en función de:

- Un empadronamiento nacional urbano que abarca a la totalidad de las comunas seleccionadas para este estudio, y no para zonas parciales o sólo para algunas comunas.
- Ofrecer las actualizaciones necesarias de nuevos poblamientos urbanos requeridos
- Contar con fuente conocida y verificable.

Cada mapa entregado a una persona encuestadora contenía la información de las viviendas a visitar, las que estaban linkeadas digitalmente con las encuestas a aplicar. La información visible para el encuestador/a sobre el sorteo, además de la imagen y delimitación de la manzana sorteada, era la siguiente:

Figura 5. Ejemplo de información en mapa de viviendas sorteadas

DIRECCIONES SELECCIONADAS		
Estudio	258.111.00318.1 Novena Encuesta de Juventud - INJUV	
ID Manzana (ANOTE)		
Manzana (SELECCIONE)	1) Original 2) Reemplazo 1 3) Reemplazo 2 4) Reemplazo 3 5) Reemplazo 4	
ANOTE LAS DIRECCIONES SELECCIONADAS EN LA HOJA DE RUTA DIGITAL		
PRIMERA SELECCIÓN DE VIVIENDAS	PRIMERA OPCIÓN DE REEMPLAZO	SEGUNDA OPCIÓN DE REEMPLAZO
Dirección (nombre y número)	Dirección (nombre y número)	Dirección (nombre y número)
>	9	> 17
>	10	> 18
>	11	> 19
>	12	> 20
>	13	> 21
>	14	> 22
>	15	> 23
>	16	> 24
ANOTE SUS OBSERVACIONES:		

Fuente: Décima Encuesta Nacional de Juventudes 2022.

En otras palabras, la hoja de ruta de cada persona encuestadora estaba previamente confeccionada a la salida a terreno y la base de datos de direcciones sorteadas estaba cargada para identificar y facilitar el seguimiento en línea de cada encuesta.

En cuanto a la probabilidad de inclusión de una manzana i en la muestra de una comuna h , fue calculada de la siguiente forma:

Figura 6. Fórmula de probabilidad de selección de viviendas

$$f_2 = \frac{m_{hi}}{M_h}$$

Fuente: Ipsos Chile

Donde:

f_2 = Probabilidad de selección de vivienda

m_{hi} = Viviendas seleccionadas en la manzana i de la comuna h ($M_h = 8$ viviendas por manzana)

M_{hi} = Viviendas de la manzana i de la comuna h .

Reemplazo de viviendas en zonas urbanas. Una vivienda fue reemplazada si luego de una tercera visita la persona seleccionada para responder la encuesta no fue localizada o si existió rechazo inicial al momento de contactar por parte de encuestador. Por otra parte, en aquellas manzanas donde es posible que no hubiesen viviendas suficientes para la cantidad de encuestas a alcanzar el reemplazo operó de la siguiente forma:

- Primera alternativa de reemplazo: las viviendas de reemplazo sorteadas aleatoriamente en la misma manzana.

Junto con la selección de las 8 viviendas a visitar, se generaron 2 reemplazos aleatorios para cada una de esas selecciones. El mapa a recepcionar por la persona encuestadora contenía la información ya diseñada de los reemplazos, por lo que no se realizó ninguna acción adicional en terreno más que aplicar correctamente las instrucciones, lo que evitó cualquier intervención humana en la selección. A continuación, se muestra una imagen con la información de reemplazo del mapa:

Figura 7. Ejemplo de viviendas de reemplazo en ingreso a la hoja de ruta del mapa

DIRECCIONES SELECCIONADAS	
Estudio	258.111.00318.1 Novena Encuesta de Juventud - INJUV
ID Manzana (ANOTE)	
Manzana (SELECCIONE)	1) Original 2) Reemplazo 1 3) Reemplazo 2 4) Reemplazo 3 5) Reemplazo 4

ANOTE LAS DIRECCIONES SELECCIONADAS EN LA HOJA DE RUTA DIGITAL

PRIMERA SELECCIÓN DE VIVIENDAS	PRIMERA OPCIÓN DE REEMPLAZO	SEGUNDA OPCIÓN DE REEMPLAZO
Dirección (nombre y número)	Dirección (nombre y número)	Dirección (nombre y número)
1	> 9	> 17
2	> 10	> 18
3	> 11	> 19
4	> 12	> 20
5	> 13	> 21
6	> 14	> 22
7	> 15	> 23
8	> 16	> 24

ANOTE SUS OBSERVACIONES:

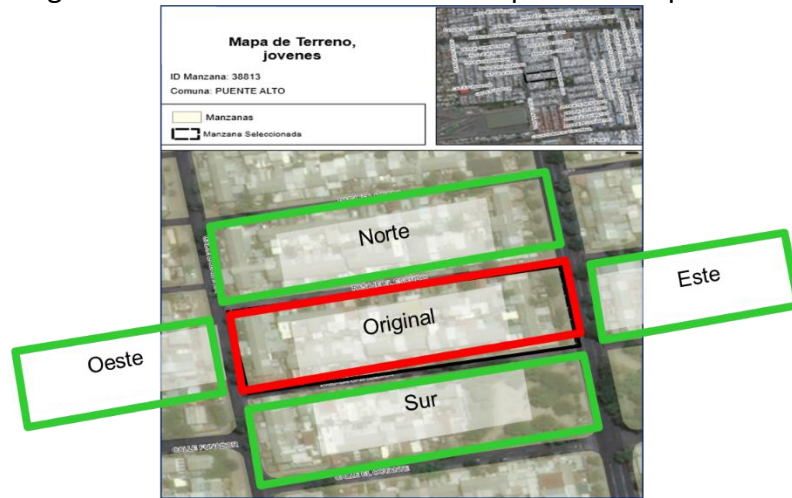
Fuente: Décima Encuesta Nacional de Juventudes 2022.

- Segunda alternativa de reemplazo: las manzanas de reemplazo.

En el caso de haber visitado las viviendas de reemplazo de la manzana originalmente sorteada, pero aun así no se logró completar la muestra de 8 encuestas, se recurrió a las manzanas que rodean a la manzana sorteada, las que operaron como manzanas de reemplazo, comenzando siempre por la manzana del norte como primera posibilidad; y de haber sido necesario, estaba como posibilidad ocupar la manzana ubicada al este, luego la del sur y finalmente la del oeste.

Estas manzanas estaban marcadas en cada mapa como se muestra en la figura a continuación. Y, al igual que la manzana original, ya tenían las 8 direcciones presorteadas con la información de las viviendas a visitar en cada manzana de reemplazo.

Figura 8. Vista de manzanas de reemplazo en mapa de terreno

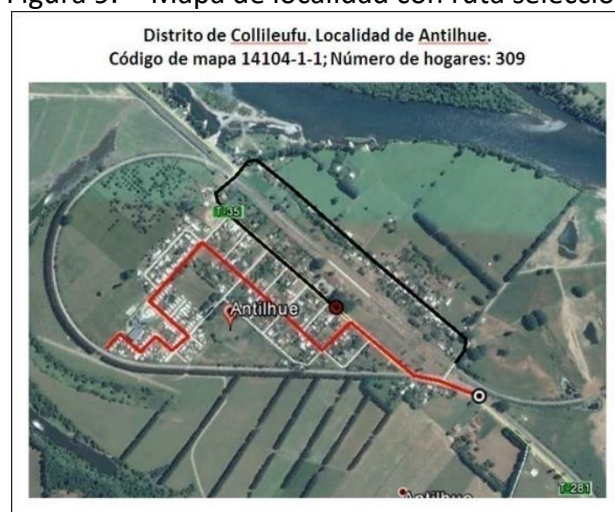


El orden de manzanas a visitar fue primero recorrer la manzana norte con los 8 reemplazos seleccionados, luego la manzana este con sus 8 direcciones de reemplazo, en tercer lugar, la manzana sur con sus 8 direcciones de reemplazo y, finalmente, la manzana oeste con sus 8 direcciones de reemplazo.

Criterios de selección de viviendas en zonas residenciales rurales:

En las zonas rurales elegidas en la segunda selección del estudio, se trabajó de la siguiente forma en terreno. En primer lugar, cada mapa de localidades urbanas tuvo una ruta predefinida para la obtención de las encuestas, como se muestra en la imagen a continuación.

Figura 9. Mapa de localidad con ruta seleccionada.



Fuente: Ipsos Chile

Para seleccionar las viviendas se ocupó la forma de selección que se aplicó en la Novena Encuesta Nacional de Juventudes 2018. Dos personas encuestadoras iniciaron en forma paralela al inicio de la ruta y avanzaron por cada vía de la calle. Cada una visitó la primera vivienda de su vía, correspondiendo ésta a la primera vivienda sorteada.

Independientemente del resultado, se procedió a aplicar un salto sistemático de 2 avanzando por su vía designada y visitó la cuarta vivienda del recorrido. Esa vivienda correspondió a la segunda vivienda sorteada. Así, se continuó con ese salto sistemático para sortear las siguientes viviendas. Este mismo sorteo lo realizó cada persona encuestadora por su vía asignada.

Todo el proceso quedó registrado en la hoja de ruta digital, donde se consignaron las direcciones de las viviendas sorteadas y las viviendas saltadas.

Reemplazo de viviendas en zonas rurales:

Una vivienda era descartada sólo después de un tercer intento de visita o cuando se trató de un rechazo efectivo. En ese caso se visitó a las dos viviendas adyacentes a la vivienda a reemplazar, manteniendo el mismo sentido de la ruta que cada uno de los dos encuestadores en terreno siguieron. En caso de no haber resultado en encuestas efectivas, se debió agregar un nuevo salto al final de la ruta.

Es decir, si por ejemplo la vivienda número 1 finalmente no resultó en encuesta efectiva después de 3 visitas, se debió aplicar un nuevo salto a partir de la última vivienda sorteada. En este sentido, se consideró aplicar la cantidad de saltos necesarios para las encuestas faltantes:

1 encuesta faltante = 1 nuevo salto
2 encuestas faltantes = 2 nuevos saltos
3 encuestas faltantes = 3 nuevos saltos
4 encuestas faltantes = 4 nuevos saltos
5 encuestas faltantes = 5 nuevos saltos
6 encuestas faltantes = 6 nuevos saltos
7 encuestas faltantes = 7 nuevos saltos
8 encuestas faltantes = 8 nuevos saltos

Tanto para zonas urbanas como rurales, en caso de registrarse más de un hogar en una vivienda, siempre fue encuestado el hogar principal, admitiendo sólo una encuesta por vivienda.

- Tercer sorteo aleatorio (fase 1 de precontacto): sorteo de respondiente

Tanto en hogares urbanos como en hogares rurales, una vez sorteada la vivienda se procedió a sortear al respondiente. Para tal efecto, se registró la información del tamaño de hogar, de la cantidad de integrantes que cumplen con los criterios de inclusión en el universo y finalmente se sorteó al respondiente.

Como mecanismo de sorteo se utilizó la elección en base a “el último cumpleaños”. Este tipo de sorteo es de fácil aplicación para el encuestador y fácil supervisión posterior (¿Quién era el último que había cumplido años al día XXX, cuando se le aplicó la encuesta?). Además, correspondía al sorteo aplicado en la Octava Encuesta de Juventudes 2015 y en la Novena Encuesta de Juventudes 2018.

La probabilidad de selección por respondiente de una manzana i en la muestra de una comuna h se calculó de la siguiente forma:

Figura 10. Fórmula de probabilidad de selección de una persona

$$f_3 = \frac{1}{k_{hij}}$$

Fuente: Ipsos Chile

Donde:

F_3 = Probabilidad de selección de una persona

k_{hij} = Número de personas seleccionables en el hogar

Una vez sorteada la persona se le solicitó su número telefónico y se avanzó en el agendamiento de la cita para la aplicación de la encuesta telefónica.

Criterios del reemplazo de respondiente:

No se permitió ningún reemplazo dentro del hogar, por cuanto se corría el riesgo de desbalancear la muestra obteniendo más mujeres o más dueñas de casa que la proporción que se obtendría aleatoriamente (es más alta la probabilidad de encontrar ese perfil dentro del hogar). Por tanto, el criterio fue reemplazar la vivienda bajo los criterios mencionados en la segunda selección correspondiente a viviendas.

- La hoja de ruta digital

Toda la información de cada visita fue registrada en una hoja de ruta digital que era la principal fuente de información para la generación de informes de avance, y se cargó en el sistema de reportería online para la información en dashboard.

La hoja de ruta digital permitió recolectar la información respecto de la tasa de respuesta, cooperación y reemplazos necesarios para lograr la muestra. La tabla de códigos para la clasificación del resultado de la visita a cada vivienda eran los siguientes:

Cuadro 11. Códigos de resultado de la visita a viviendas

		Códigos de columna R	
	Encuesta aplicada	Completa	1
		Parcial (Con algunas preguntas o módulos sin responder, pero terminada)	2
Elegible, sin entrevista	Rechazos y quiebres	Rechazo de la persona seleccionada (antes de iniciar la encuesta)	3
		Rechazo en la vivienda seleccionada (antes de iniciar la encuesta)	4
		Quiebre de la entrevista (El encuestado no quiere seguir respondiendo)	5
	No hay contacto	Acceso impedido	6
		Nadie en el hogar, pero habitado	7
		Sujeto seleccionado no ubicable, de viaje o no puede atender	8
	Otros	Impedido de contestar por enfermedad (física o mental)	9
		Problemas de lenguaje o idioma	10
		Otros	11
Elegibilidad desconocida	Se desconoce la existencia de hogares	El sector no se trabajó (sin empadronamiento)	12
		Áreas peligrosas o imposibles de acceder	13
		Imposible localizar la dirección	14
	Se desconoce si hay entrevistados elegibles	El hogar no se trabajó, se desconoce si hay elegibles.	15
No elegible	Fuera del marco muestral	Casos fuera del marco muestral (no viven jóvenes de 15 a 29 años)	16
	No hay vivienda	Empresas, oficinas, otras organizaciones	17
		Instituciones (Hospitales, cárceles...)	18
		Recintos militares o de entrenamiento	19
		Otros	20
	Vivienda desocupada	Viviendas regulares desocupadas	21
		Residencias de vacaciones o con ocupación estacional	22
		Otras	23

Fuente: Décima Encuesta Nacional de Juventudes 2022.

Estos códigos y su trabajo posterior fueron hechos en base a la clasificación de American Association for Public Opinion Research (AAPOR). La AAPOR ofrece una estandarización de códigos para agrupar incidencias de acuerdo a los tipos de resultados obtenidos.

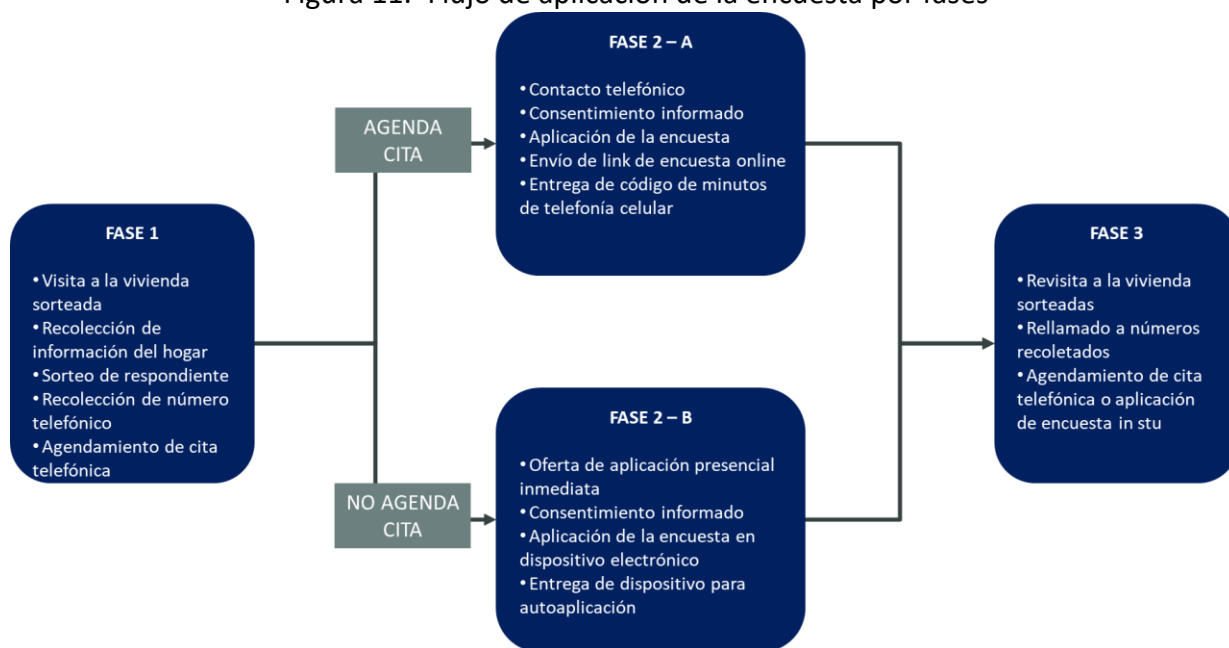
La hoja de ruta y su tratamiento permitieron dimensionar la cantidad de contactos necesarios para el logro de una encuesta y la efectividad del muestreo. Los códigos son un importante nivel de detalle de las incidencias consideradas, lo cual facilitó la codificación de una amplia gama de resultados posibles en el trabajo de campo.

Esta información permitió calcular la tasa de respuesta, tasa de cooperación, tasa de rechazo y tasa de contacto del muestreo, informaciones claves para la elaboración de los informes de terreno y el informe final del estudio.

6.1 Fase 2: Mix-mode, Encuesta telefónica y Encuesta online a personas sorteadas en hogares como técnica de levantamiento.

Una vez aplicado el formulario de precontacto en la fase 1 y habiendo sorteado a la persona que respondía la encuesta, tanto en la muestra de jóvenes como en la muestra de personas adultas, se pasó a la fase 2 de aplicación de encuestas. En este punto es importante establecer la conexión entre la fase 1 y la fase 2 de la aplicación de la encuesta y la posterior fase 3. Para tal efecto, se presenta el siguiente diagrama de flujo de la aplicación de la encuesta.

Figura 11. Flujo de aplicación de la encuesta por fases



Fuente: Décima Encuesta Nacional de Juventudes 2022.

En términos generales, la fase 1 permitió sortear al respondiente de cada encuesta, recolectar su número telefónico -necesario tanto para la aplicación de la encuesta como para la supervisión telefónica- y el agendamiento de cita para la aplicación de la encuesta mix-mode.

La técnica que correspondía a la encuesta mix-mode, integra las técnicas de la encuesta telefónica y la encuesta online.

La encuesta telefónica es la estrategia primaria de recolección de datos, la que será aplicada con dispositivo electrónico por el equipo de encuestas telefónicas. Para tal efecto, en la fase 1 de aplicación de precontacto se solicitó el número de teléfono de la persona sorteada, el que se ingresó a la información de la hoja de ruta digital y formó parte de la base de datos telefónica a contactar.

Los momentos de la aplicación de la encuesta telefónica fueron los siguientes:

1. Confirmación de los datos recolectados en la fase 1 de sorteo.
2. Entrega de información del estudio y confirmación de la participación mediante aceptación grabada del consentimiento informado.
3. Aplicación de la encuesta por parte del equipo de encuestadores/as telefónicos capacitado para el estudio.
4. Entrega de link de encuesta online para responder sección auto aplicada.
5. Confirmación en línea del cierre de la sección auto aplicada.
6. Entrega de código de minutos de telefonía celular.
7. Cierre de la entrevista.

Al respecto, cada persona encuestadora telefónica aplicó el protocolo previamente relatado, implementando la encuesta en forma telefónica, la que fue grabada en formato de audio a partir de la aceptación de aplicación posteriormente a la lectura de consentimiento informado. Sin embargo, el cuestionario es de naturaleza mixta en la medida que incorpora dos secciones de diferente formato de aplicación:

- Sección aplicada por persona encuestadora, la que se estima conforma cerca del 80% de instrumento, correspondiente a la primera sección. Esta parte del instrumento contiene desde la presentación de la encuesta, las secciones de indicadores a estudiar como parte estudio y la caracterización socioeconómica del hogar.
- Sección auto aplicada. Esta segunda sección, estimada en el 20% de la encuesta, correspondiente a las secciones de sexualidad, drogas y violencia. Como se declara en el Informe de la VII Encuesta de Juventudes, “Consultar sobre temas sensibles representa siempre una dificultad para las encuestas. Esto es, temas sobre los cuales el entrevistado no está dispuesto a entregar públicamente información. Tal es el caso de las conductas sexuales, el consumo de drogas, las conductas violentas, entre otras. Tourangeau et al. (2000) revisan la evidencia al respecto e indican que las preguntas pueden ser sensibles en al menos tres dimensiones: deseabilidad social de las respuestas, invasión de la privacidad y preocupación por la confidencialidad. Dada las probabilidades de ocurrencia de las sensibilidades antes expuestas, en la VII Encuesta de Juventud del año 2012 se procedió a separar esta etapa de aplicación (posteriormente a la revisión de funcionamiento en el pretest) de la aplicación clásica por encuestador.

En este caso, la persona encuestada podía responder la segunda sección de la encuesta en privacidad. Para tal efecto se aplicó la técnica de recolección secundaria correspondiente a la encuesta online. El procedimiento era que, llegado el momento de pasar a la sección auto aplicada, la persona encuestadora enviaba el link de ingreso a la encuesta online que se habilitaba con un código que permitía en la base de datos integrar las respuestas de las dos secciones. Esta información se mantiene en línea mientras él o la respondiente accedía a la encuesta, respondía y concluía con el envío de la misma. De esta forma se minimizaba la pérdida de casos por no respuesta de la sección auto aplicada.

Plan de contingencia para disminuir la pérdida de personas seleccionadas y evitar el aumento de la tasa de no respuesta:

Una estrategia desarrollada en terreno en la fase de precontacto que Ipsos implementó exitosamente es la opción de aplicación en el hogar posteriormente al sorteo de respondiente. Es decir, ofrecer la fase 2 *in situ*. Esto se implementaba cuando se detectaba que la persona seleccionada no estaba interesada en la aplicación telefónica, tiene dudas al respecto o indica directamente su disponibilidad inmediata para la aplicación de la encuesta.

Para tal efecto, las personas encuestadoras aplicaban la encuesta en dispositivo electrónico, soporte clasificado dentro del sistema de recolección CAPI (Computer Assisted Personal Interviewing). Además, era requisito en terreno el uso de baterías externas y opciones de recarga para evitar la pérdida de casos por problemas de carga.

En esta modalidad, la sección auto aplicada de la encuesta fue ofrecida de ser respondida en forma inmediata en el mismo dispositivo de la persona encuestadora, no incluyendo la entrega del link de ingreso de la encuesta online. Esta modalidad es la que se ocupó en la totalidad de los casos de la Novena Encuesta de Juventudes 2018, por lo que se trataba de una modalidad ya probada y validada. Esta estrategia se sustentaba en la alta probabilidad de abandono de la encuesta por tratarse de casos que ya han manifestado su voluntad de responder *in situ* y en forma inmediata, lo que se traduce en una menor disponibilidad de respuesta independiente y posterior. Por tanto, no se requería del uso de dispositivos electrónicos de propiedad del respondiente o en gasto de sus gigas o cuenta celular. Por lo que la última variable de la hoja de ruta digital era si la persona accede a responder telefónicamente la encuesta u opta por responderla presencialmente *in situ*.

La modalidad mix-mode propuesta, la cual combinaba las diversas técnicas en la presente propuesta, permitía que el estudio obtuviera las siguientes ganancias en el trabajo de campo:

- Validación a través de sistema de información incorporado por programación de la encuesta que disminuye los errores por ingreso y evita las preguntas no aplicadas por el encuestador.
- Optimización del tiempo de la persona encuestadora en terreno y menor carga de material.
- Eliminación de dos unidades de la cadena de producción: el equipo de revisadores y el equipo de digitación, con el consiguiente ahorro de tiempo y disminución de fuentes de errores no muestrales.
- Eliminación completa del papel, lo que es un aporte desde el enfoque de sostenibilidad del estudio, manteniendo el estándar ya logrado en el año 2018. Además, es un mecanismo de protección frente al contagio de SARS-CoV-2 para personas encuestadoras y encuestadas.

6.2 Fase 3: Recorrido de casos pendientes de las fases 1 y 2

La tercera fase de terreno correspondió a la recontactación de casos pendientes ya sea por no contactación por no acceso en la fase 1, los casos en los que no se sortearon respondientes en la fase 1, los casos que fueron sorteados en la fase 1 pero con se concretaron en encuestas aplicadas en la fase 2 o que abandonaron o no concluyeron la totalidad de la encuesta en la fase 2.

La forma de recontactación dependía del estatus del caso:

- a) Revisita de la vivienda en terreno para casos no contactados por no acceso y los casos en los que no se sortearon respondientes en la fase 1
- b) Rellamada al respondiente en los casos que fueron sorteados en la fase 1 pero no se concretaron en encuestas aplicadas en la fase 2 o que abandonaron o no concluyeron la totalidad de la encuesta en la fase 2.

En ambos casos se reiniciaba el proceso en la etapa pendiente y se ofrecía la respuesta telefónica para el cierre de la encuesta o de la visita en la vivienda para aplicación *in situ* de la encuesta para finalizar el caso pendiente.

7. Validación, cálculo de expansores y ponderación de las bases de datos

Previo al procesamiento de cada una de las bases de datos, se crearon los factores de expansión para la muestra de jóvenes entre 15 y 29 años y para la muestra de adultos entre 30 y 59 años.

Las unidades de primera etapa corresponden a las manzanas en las zonas urbanas y a las secciones en las zonas rurales. Por otra parte, un estrato para el factor de expansión correspondió a la combinación de Región-Zona, por lo cual se totalizan 32 grupos según el factor a considerar, los cuales surgen a partir de la combinación de 16 regiones y 2 zonas (urbano-rural), esto debido a que la asignación de la muestra se realizó a priori para cada uno de estos estratos, sin que esta sea dejada al azar. Este diseño dio lugar a un estudio con un nivel de representatividad que abarca el nivel nacional, regional y nacional urbano-rural. A su vez, las unidades de segunda etapa corresponden a las viviendas, las cuales fueron seleccionadas al interior de cada una de las unidades de primera etapa (manzana o sección), y de ellas se obtuvieron los informantes, quienes corresponden a la unidad de tercera etapa dentro del muestro polietápico de la Encuesta Nacional de la Juventud.

La selección de unidades de primera etapa, en concordancia con el método bietápico, se realizó en cada estrato (región-zona) con probabilidad proporcional al tamaño de la unidad (sección o

manzana) en términos del número de viviendas que la componen. Así, la probabilidad de inclusión de la i-ésima unidad de primera etapa correspondió a:

Figura 12. Probabilidad de inclusión de i-ésima unidad de primera etapa

$$P_{hi} = \frac{n_h M_{hi}}{M_h}$$

Donde

- h: corresponde al estrato (conformado por Región-Zona)
- i: corresponde a la “i-ésima” unidad de primera etapa (manzana o sección)
- n_h : número de unidades de primera etapa en la muestra del estrato “h”.
- M_{hi} : Número de viviendas en la i-ésima unidad de primera etapa del estrato “h” (información obtenida del marco muestral)
- M_h : Número de viviendas en el estrato “h” (información obtenida del marco muestral)

La selección de unidades de segunda etapa (viviendas dentro de cada unidad de primera etapa seleccionada), se realizó con probabilidad igual para todas las viviendas dentro de la respectiva unidad, mediante una selección sistemática. Dicha probabilidad corresponde a:

Figura 13. Probabilidad de inclusión de unidad de segunda etapa

$$P_{j/hi} = \frac{m_{hi}}{M'_{hi}}$$

Donde

- j/hi: j-ésima vivienda seleccionada al interior de la i-ésima unidad de primera etapa (manzana o sección) seleccionada dentro del estrato “h”.
- m_{hi} : número de viviendas logradas dentro de la i-ésima unidad de primera etapa (manzana o sección) seleccionada dentro del estrato “h”.
- M'_{hi} : número de viviendas en la sección o actualizadas de la manzana en la i-ésima unidad de primera etapa seleccionada dentro del estrato “h”.

Aunque se consideró encuestar a 8 viviendas en la manzana/sección seleccionada. Es decir, $m_{hi}=8$. Sin embargo, existen posibilidades de reemplazo (viviendas adyacentes y manzanas adicionales de reemplazo) para completar las 8 viviendas, por lo cual, m_{hi} : corresponde al número de viviendas logradas en cada manzana/sección seleccionada en el estrato. Al utilizar este método, no se requirió corregir por la no respuesta, ya que esta queda implícita.

El Factor de Expansión se interpreta como la cantidad de viviendas en la población que representa una vivienda en la muestra. Lo que se transmite a las personas que viven en la vivienda.

Dicho factor de acuerdo con el diseño corresponde al aplicado a una muestra de viviendas realizada en dos etapas, a la cual se suma posteriormente el sorteo de la tercera unidad

equivalente a los respondientes, la cual se realiza por medio del sorteo del último cumpleaños. Este depende muestralmente del número de viviendas que tienen las unidades de primera etapa (manzanas o secciones) y de la cantidad de éstas en cada estrato, y, poblacionalmente, del número de viviendas en el estrato y en la sección o manzana seleccionada.

El factor de expansión teórico $F1_{hi}$ es calculado como el inverso de la probabilidad de selección de cada vivienda. Por tanto, este factor es constante por manzana o sección y se aplica sobre todas las viviendas y por ende a todas las personas que pertenecen a la manzana o sección “i” del estrato “h”. Este factor se define como sigue:

Figura 14. Factor de expansión teórico $F1_{hi}$

$$F1_{hi} = \frac{M_h}{M_{hi}n_h} \times \frac{M'_{hi}}{m_{hi}}$$

Donde

- M_h : Número de viviendas en el estrato “h” (información obtenida del marco muestral)
- M'_{hi} : número de viviendas de la sección o actualizadas de la manzana en la i-ésima unidad de primera etapa seleccionada dentro del estrato “h”.
- M_{hi} : Número de viviendas en la i-ésima unidad de primera etapa del estrato “h” (información obtenida del marco muestral)
- n_h : número de unidades de primera etapa en la muestra del estrato “h”.
- m_{hi} : número de viviendas logradas dentro de la i-ésima unidad de primera etapa (manzana o sección) seleccionada dentro del estrato “h”.

Cabe consignar que el uso del factor de expansión teórico descrito en el párrafo precedente permite “reproducir” los totales (corregidos) relativos al número de viviendas por estratos. Esta corrección se produce por el proceso de actualización del tamaño de cada unidad de primera etapa seleccionada (manzana o sección). No obstante, en la presente encuesta no se llevó a cabo una actualización del número de viviendas de las manzanas o secciones seleccionadas debido a que se optó por un sorteo tanto de manzanas como de viviendas realizado en oficina, descartándose la opción de un empadronamiento de las manzanas en terreno. Esta situación también se replica en las zonas rurales, en donde no se realizó un empadronamiento de los sectores rurales encuestados.

- Selección del informante

Una vez obtenido la muestra de viviendas de acuerdo con el diseño descrito previamente, para cada una de las unidades de la muestra se procedió a identificar a los elegibles (sean adultos o jóvenes respectivamente) y se procedió a “seleccionar / sortear” de acuerdo con el procedimiento del último cumpleaños. Cabe mencionar que se selecciona solo UN INFORMANTE por unidad muestral (vivienda). Técnicamente, al incorporar esta etapa en el muestreo, el diseño corresponde a un muestreo polietápico Manzana (o sección) – Vivienda – Informante. La probabilidad de inclusión es absorbida por el ajuste por sexo y tramo de edad que se describe en los siguientes párrafos.

- Ajuste Urbano – Rural

El primer factor de corrección que se aplicó al factor de expansión anterior correspondió a la corrección de acuerdo a la composición urbano – rural. Esta información se obtuvo de la fuente oficial del Censo 2017 del Instituto Nacional de Estadística.

Por lo tanto, el factor de expansión F2 resulta del producto entre el factor de expansión F1 y el factor de corrección por zona FC1. Es decir, $F2 = F1 \times FC1$.

- Ajuste por sexo y tramo de edad

Al disponer de información actualizada y confiable respecto a la composición de la población según sexo y tramo de edad, el factor de expansión puede incluir un ajuste de población de personas por estrato, donde en cada estrato se dispone de estimaciones de proyecciones de población por sexo y tramo de edad. Este factor corregido permite aumentar la precisión de las estimaciones y permite resultados consistentes con los stocks poblacionales de las proyecciones de población. Los datos para realizar esta corrección se obtuvieron de la fuente oficial del Censo 2017 del Instituto Nacional de Estadística.

Así, para cada sección o manzana del estrato “h” se tienen 2xt factores de expansión (2 – por masculino y femenino, y t – número de tramos de edad) que se aplican para todas las personas dentro de la sección que sean hombres o mujeres de acuerdo con el tramo de edad.

El cálculo de factor de expansión final corresponde a:

Figura 15. Factor de expansión final

$$F3'_{hi} = \frac{M_h}{M_{hi}n_h} \times \frac{M'_{hi}}{m_{hi}} \times FC1 \times FC2, \text{ donde } FC2 = \frac{P_h}{\widehat{P}_h'}$$

Donde

- FC1: factor de corrección según composición Urbano – Rural.
- FC2: factor de corrección según Sexo y Tramo de Edad.
- Ph: proyección de población total por sexo y tramo de edad en el estrato “h”
- \widehat{P}_h' : número total de personas estimadas con el factor teórico, por sexo y tramo de edad en el estrato “h”.

Con

$$\widehat{P}_h' = \sum_i \sum_j \frac{M_h}{M_{hi}n_h} \times \frac{M'_{hi}}{m_{hi}} \times P_{hij} = \sum_i \sum_j F2_{hi} \times P_{hij}$$

Donde

- P_{hij} : número total de personas elegibles en la j-ésima vivienda de la i-ésima unidad primaria (manzana o sección) en el estrato “h”.

- Factor de expansión – Tasa de no respuesta

En las encuestas con diseño muestral probabilístico complejo – dos o más etapas - se debe utilizar un ponderador para que las estimaciones de las variables de interés tengan validez respecto a la población objetivo. El factor de expansión permite reflejar directamente las probabilidades de selección de las distintas unidades de muestreo en cada una de las etapas.

Factor de expansión teórico presentado previamente, $F1_{hi}$, incorpora la tasa de no respuesta. Recuérdese que

$$F1_{hi} = \frac{M_h}{M_{hi}n_h} \times \frac{M'_{hi}}{m_{hi}}$$

Donde

- M_h : Número de viviendas en el estrato “h” (información obtenida del marco muestral)
- M'_{hi} : número de viviendas de la sección o actualizadas de la manzana en la i-ésima unidad de primera etapa seleccionada dentro del estrato “h”.
- M_{hi} : Número de viviendas en la i-ésima unidad de primera etapa del estrato “h” (información obtenida del marco muestral)
- n_h : número de unidades de primera etapa en la muestra del estrato “h”.
- m_{hi} : número de viviendas logradas dentro de la i-ésima unidad de primera etapa (manzana o sección) seleccionada dentro del estrato “h”.

Como se indica, al utilizar este método no se requiere corregir por la no respuesta, ya que está implícita. Lo anterior se debe a que el factor de expansión definido para cada conglomerado (en este caso manzanas en el caso de zonas urbanas y secciones en el caso de las zonas rurales) incorpora el ajuste de no respuesta. Lo que implica que la ponderación de las unidades que no responden se distribuye uniformemente en las unidades del conglomerado que si responden (en términos equivalente a lo descrito en la ENE y como se aprecia en la ecuación correspondiente).

- Necesidad de Ranking

El algoritmo previamente descrito permite obtener los ponderadores y factores de expansión adecuados para todas las combinaciones, siempre y cuando se cuente con casos muestrales para cada una de las celdas.

Sin embargo, el procedimiento presenta problemas en los casos donde hay muy pocas o ninguna observación en la respectiva celda: **estrato-sexo-tramo de edad**.

De hecho, el caso de la encuesta de Adultos – zona rural existen 10 celdas sin observaciones y más de una docena con solo un caso (ver tabla 1). En el caso de la encuesta de Jóvenes existen, en la zona rural, dos celdas sin observaciones. Lo anterior implica que el procedimiento de *weighting* no es recomendable dado que hay celdas en la cual no es aplicable, y por tanto se recomienda aplicar *raking*.

Tabla 1 – Número de casos – Encuesta Adultos – según región-sexo-tramo etario – Zona RURAL

REGION	Hombre				Mujer				Total
	30-39	40-49	50-59	Tot H	30-39	40-49	50-59	Tot M	
1. I Región (Tarapacá)		4	3	7	2	4	2	8	15
2. II Región (Antofagasta)	3		3	6	3	5	1	9	15
3. III Región (Atacama)	2	2	1	5	5	4	1	10	15
4. IV Región (Coquimbo)	1	2	3	6	2	2	5	9	15
5. V Región (Valparaíso)	1		1	2	6	4	3	13	15
6. VI Región (Libertador B. O'Higgins)	4	1		5	3	4	3	10	15
7. VII Región (Maule)		2	4	6	1	4	4	9	15
8. VIII Región (Biobío)	2	1	2	5	2	3	5	10	15
9. IX Región (La Araucanía)	1	1	1	3	2	4	6	12	15
10. X Región (Los Lagos)	4	5		9	5		1	6	15
11. XI Región (Aysén)	3	3	3	9	1	4	1	6	15
12. XII Región (Magallanes)	1	2	1	4	1	3	7	11	15
13. Región Metropolitana	1	1	2	4	3	6	2	11	15
14. XIV Región (Los Ríos)		1	1	2	8	1	4	13	15
15. XV Región (Arica y Parinacota)	1	1		2	5	5	3	13	15
16. XVI Región (Ñuble)	2		3	5	4	3	3	10	15
Total general	26	26	28	80	53	56	51	160	240

Frente a este problema, en la encuesta de Adultos se optó por calibrar mediante la técnica de raking el sector rural solo por sexo. Cabe mencionar que el procedimiento permite recuperar los stocks para cada uno de los estratos. Además, de reproducir al interior de cada estrato según sexo y tramo de edad para la zona Urbana y solo según sexo para la zona Rural.

En el caso de la encuesta de jóvenes, aun para el tamaño muestral utilizado, el problema se presentó en la zona rural de dos regiones. En este caso, dada la relevancia de disponer de estimaciones para cada uno de los estratos en forma desagregada, se recurrió a la metodología raking – la cual mediante un método iterativo permite ajustar y obtener los factores de expansión. Cabe mencionar que el método permite recuperar los stocks según sexo (y no por tramo etario) para esas dos regiones en la zona rural.

Sin embargo, en la muestra obtenida para los jóvenes existen dos combinaciones en los cuales no se lograron entrevistas, tal como lo muestra el siguiente cuadro.

Tabla 2 – Número de casos – Encuesta Jóvenes – según región-zona-sexo-tramo etario

Region	Rural						Urbano						Total
	Hombre			Mujer			Hombre			Mujer			
	15-19	20-24	25-29	15-19	20-24	25-29	15-19	20-24	25-29	15-19	20-24	25-29	
XV Región (Arica y Parinacota)	8	4	1	14	2	8	68	60	47	80	59	49	400
I Región (Tarapacá)	3	1	4	6	3	2	47	38	63	78	76	79	400
II Región (Antofagasta)		2	4	3	1	2	54	36	62	50	44	142	400
III Región (Atacama)	7	1		9	9	2	66	47	40	79	68	72	400
IV Región (Coquimbo)	13	8	8	7	14	9	59	47	50	63	46	76	400
V Región (Valparaíso)	13	7	10	13	16	10	156	170	113	168	164	160	1000

VI Región (Lib. B. O'Higgins)	18	8	18	9	16	21	63	38	46	65	50	48	400
VII Región (Maule)	19	11	11	14	20	13	53	53	41	61	60	44	400
XVI Región (Ñuble)	28	5	14	16	8	26	49	38	42	49	56	69	400
VIII Región (Biobío)	14	11	13	14	13	23	147	127	127	164	160	187	1000
IX Región (La Araucanía)	18	13	12	16	13	19	51	67	27	52	62	50	400
XIV Región (Los Ríos)	12	13	10	13	18	19	36	47	60	48	46	78	400
X Región (Los Lagos)	10	7	9	17	16	20	42	34	34	57	66	88	400
XI Región (Aysén)	6	3	9	5	16	17	56	49	42	58	45	94	400
XII Región (Magallanes)	4	8	4	9	5	4	61	49	41	61	65	89	400
Región Metropolitana	11	12	13	15	12	19	458	401	295	412	398	454	2500
Total general	184	114	140	180	182	214	1466	1301	1130	1545	1465	1779	9700

Como se puede apreciar, existen DOS celdas en los cuales no existe casos muestrales. Por lo anterior, las expresiones previas para ajustar y expandir por área – sexo – tramo etario no son posibles.

Con el objetivo de determinar los correspondientes ponderadores y factores de expansión se recurre para estas dos regiones (II y III) y área Rural a la metodología de calibración denominado RAKING.

En el presente caso se aplicó **Raking** (con software R – función *anesrake*) inicialmente por región-etario y posteriormente por región-sexo. Lo anterior permite “recuperar” las participaciones “exactas” a nivel región-sexo, cercanas a nivel región-etario y obviamente, no es posible tener la precisión para región-sexo-etario.

Los resultados de la aplicación de Raking a la II y III región Rural, por sexo y tramo etario permitió obtener ponderadores y factores de expansión que se ajustan bien a nivel regional y por sexo, y para la región de atacama no fue posible lograr la cuadratura, tal como se muestra en la siguiente tabla al contrastar los resultados con las proyecciones del INE para la población 2022:

ESTRATO - Rural	RAKING	Pob - INE	Difer %
Region II - total	5101	5101	0,0%
Region III- total	4406	4406	0,0%
Region II - Sexo MASC	3523	3523	0,0%
Region II - Sexo FEME	1578	1578	0,0%
Region III - Sexo MASC	2293	2293	0,0%
Region III - Sexo FEME	2113	2113	0,0%
Region II - Etario 15-19	771	771	0,0%
Region II - Etario 20-24	1597	1597	0,0%
Region II - Etario 25-29	2733	2733	0,0%
Region III - Etario 15-19	1701	1353	25,7%
Region III - Etario 20-24	1344	1381	-2,7%
Region III - Etario 25-29	1361	1672	-18,6%

- Ranking -

Raking (rastrillo en español) es una metodología propuesta a inicio de los años 70, como resultado de las técnicas de estimación iterativas utilizada en los modelos log-lineales (iterative proportional fitting). También ha sido denominada como RIM (ajuste proporcional iterativo) que en palabras simples implica, ajustar por una variable (ej. Sexo) y luego por la siguiente variable (ej. Etario). Si los ajustes de una variable afectan demasiado a la otra variable, se itera nuevamente sobre los ponderadores previos, iterando sistemáticamente hasta converger (de ahí el término iterativo).

Es claro que el procedimiento es tedioso y complejo cuando hay tres o más variables involucradas (región área, sexo, etario). Si desean revisar un ejemplo simple y explicado paso-a-paso pueden ver: Bueno, E., Zarraga, A. y Iztueta, Z. (1999) **Ajuste de muestras con información auxiliar**, Eostat. N° 9801, pags: 31-38.

Afortunadamente, hoy en día está implementado en diferentes softwares. En el presente caso se utilizaron las funciones de la librería **anesrake** de R.

- <https://cran.r-project.org/web/packages/anesrake/anesrake.pdf>

Alternativamente, es posible utilizar las funciones de:

- STATA – el paquete ipfraking permite calibrar los datos de una muestra utilizando IPF (iterative proportional fitting)
- SPSS – existen funciones que permiten implementar RIM WEIGHTING con el procedimiento GENLOG.
- SAS – existe macros que permiten “balancear” muestras ponderadas (raking)

8. Cálculo y clasificación de niveles socioeconómicos

Para la clasificación socioeconómica de los hogares donde habitan las personas encuestadas, se utilizó el modelo original de ESOMAR, que calcula el Grupo Socioeconómico (GSE) del hogar en función de dos variables: la profesión o trabajo del principal sostenedor del hogar y su nivel educacional².

Cuando no se obtuvo información suficiente para hacer la clasificación del GSE a partir de la utilización del modelo de ESOMAR, el GSE del hogar se declaró como “No calculable”. Esto ocurrió en 447 casos para la muestra a jóvenes, y a 20 para la muestra a adultos.

Para efectos de nomenclatura, se refiere a “Grupo Socioeconómico” o “GSE” al hablar de los cinco grupos resultantes de la metodología recién citada: ABC1, C2, C3, D y E. Asimismo, se refiere a “Nivel Socioeconómico” o “NSE” al hablar de la agrupación de los anteriores grupos de la siguiente forma: alto (ABC1), medio (C2 y C3) y bajo (D y E).

² Para mayor detalle sobre las categorías de respuesta de estas dos variables y algunos ajustes al modelo que hizo la empresa ejecutora, consultar el documento de trabajo final entregado por IPSOS al INJUV.

9. Cuadro comparativo Encuesta Nacional de Juventud (1994-2022)

Cuadro comparativo de las Encuestas Nacionales de Juventud										
Encuestas	I	II	III	IV	V	VI	VII	VIII	IX	X
Años	1994	1997	2000	2003	2006	2009	2012	2015	2018	2022
Casos	3.792	3.448	3.701	7.189	6.345	7.570	8.352	9.393	9.700	9.700
Áreas de representación	Por regiones I, II, V, VIII, IX, X, XI y RM	Nacional, urbana	Nacional, urbana y rural	Nacional, urbana y rural, y regional	Nacional, urbana y rural, y regional	Nacional, urbana y rural, y regional	Nacional, urbana y rural, y regional	Nacional, urbana y rural, y regional	Nacional, urbana y rural, y regional	Nacional, urbana y rural, y regional
Población entrevistada	Jóvenes 15-29 años de las regiones I, II, V, VIII, IX, X, XI y RM	Jóvenes 15-29 años de zonas urbanas de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones	Jóvenes 15-29 años de zonas urbanas y rurales de todas las regiones
Tipo de entrevista	Cara a cara	Cara a cara	Cara a cara	Cara a cara	Cara a cara	Cara a cara	Cara a cara, más módulo autoaplicado	Cara a cara, más módulo autoaplicado	Cara a cara, más módulo autoaplicado	Cara a cara, más módulo autoaplicado 0 Mixta secuencial
Diseño muestral	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos	Muestreo estratificado por conglomerados y polietápicos
Error muestral	Sin información	Sin información	2,91%	2,10%	1,25%	1,15%	1,07%	1,01%	1,00%	1,00%
Confiabilidad usada para el error	Sin información	Sin información	95%	95%	95%	95%	95%	95%	95%	95%
Levantamiento	Papel	Papel	Papel	Papel	Papel	Papel	Papel	Papel	Tablet	Tablet/telefónica
Muestra independiente de adultos	No	No	No	No	No	No	No	Sí	Sí	Sí