

MANUAL DEL USUARIO

# Estudio de Inmigración en la Macrozona Norte

En base a la Encuesta Nacional  
de Opinión Pública CEP N° 90  
Septiembre-Noviembre 2023



El presente Manual del Usuario contiene los antecedentes relativos al Estudio de inmigración en la macrozona norte, realizado en base a la Encuesta Nacional de Opinión Pública CEP N° 90 (septiembre-noviembre 2023) llevada a cabo por el Centro de Estudios Públicos.

Este manual entrega información general del presente estudio, lo que ayudará a la comprensión de los distintos temas relativos al estudio. El **Manuel del Usuario** incluye:

- La presentación y justificación del estudio
- El diseño metodológico de la encuesta

## Ficha Técnica

Nombre oficial:	Estudio de inmigración con sobremuestra en la macrozona norte, realizado en base a la Encuesta Nacional de Opinión Pública CEP N° 90
Organismo responsable:	Centro de Estudios Públicos (CEP)
Período de la encuesta:	24 de septiembre a 18 de noviembre 2023
Cobertura geográfica:	Nacional
Tamaño de la muestra:	1.848 (1.478 casos de la encuesta CEP 90 + 370 casos de la sobremuestra en la macrozona norte)
Error muestral:	±2,3 puntos porcentuales con un nivel de confianza de un 95%, varianza máxima y un efecto del diseño estimado (deff) de 1,2
Número de variables:	86
Tasa de respuesta:	61,8%
Porcentaje de cobertura:	97%
Número de comunas:	128
Población objetivo:	Población de 18 años y más residente en todo el país (a excepción de las zonas de difícil acceso: Isla de Pascua, Palena y Antártica Chilena)
Número de UMP:	365
Modo de aplicación:	Presencial, cara a cara
Número de encuestadores promedio durante aplicación:	14
Unidad de análisis:	Individuo

## Ponderación de los Resultados

Las dificultades prácticas en la realización de encuestas pueden introducir un elemento de error en los resultados, fuera del error de muestreo. Para corregir estos errores es necesario aplicar una ponderación a los datos, de modo que se garantice una representación exacta de los adultos en todo el país. El ponderador para esta encuesta corresponde a la variable *pond*, y debe aplicarse antes del análisis de los datos.

## Diseño complejo

Las variables de diseño corresponden a *estrato* para los estratos y *secu* para los conglomerados. Además de *pond* que corresponde al ponderador.

## Manual del Usuario

### Estudio de inmigración en la macrozona norte

### En base a la Encuesta Nacional de Opinión Pública CEP N° 90

Septiembre-Noviembre 2023 | 24/04/2024

## 1. Presentación

El Estudio de inmigración en la macrozona norte busca identificar posibles barreras a la inclusión de la población inmigrante provenientes de las actitudes y percepciones de la sociedad chilena. Basándonos en la literatura y en reuniones con expertos en el tema, incorporamos preguntas que capturan los factores que inciden en las actitudes respecto de la inmigración —percepciones de amenaza, tanto a nivel económico como simbólico (cultura, identidad)— así como sus beneficios<sup>1</sup>.

A diferencia de encuestas CEP anteriores, en este estudio se amplió la muestra de la macrozona norte de Chile (Arica y Parinacota, Tarapacá, Antofagasta y Atacama), como se sabe, la mayormente afectada por la ola migratoria en los últimos cinco años. Esta sobremuestra fue efectuada de forma paralela a la Encuesta CEP N° 90 nacional. La muestra esperada para la macrozona norte en la encuesta nacional es de aproximadamente 180 casos. Con el fin de recoger las percepciones, actitudes y vivencias particulares de los residentes en esta zona, se decidió llevar a cabo 350 encuestas adicionales en la zona para así obtener un total de aproximadamente 500 casos y de este modo incrementar el poder estadístico para realizar inferencias a nivel de la macrozona norte. El cuestionario utilizado para los casos de sobremuestra incluyó el mismo módulo especial sobre inmigración de la encuesta nacional CEP N° 90, con algunas variables del módulo socioeconómico.

El trabajo de terreno y recolección de información se desarrolló entre los días 24 de septiembre a 18 de noviembre entrevistando a 1.848 personas. El universo de estudio se definió como la población de 18 años y más (urbana y rural) a lo largo de todo el país, excluyendo las provincias definidas como ADA por el INE (difícil acceso). Éstas son Isla de Pascua, Palena y Antártica Chilena<sup>2</sup>. La muestra cubrió el 97% de la población objetivo<sup>3</sup>.

---

<sup>1</sup> Un listado de la totalidad de estudios de opinión pública realizados por el Centro de Estudios Públicos se encuentra en la página web de las encuestas CEP del Programa de Opinión Pública del CEP, ver <http://bit.ly/encuestas-cep>. En este total no se incluyen las encuestas especiales en el marco del Consorcio de Centro de Estudios.

<sup>2</sup> Estas provincias incluyen las comunas de General Lagos, Colchane, Ollagüe, Juan Fernández, Isla de Pascua, Cochamó, Chaitén, Futaleufú, Hualaihué, Palena, Lago Verde, Guatecas, O'Higgins, Tortel, Laguna Blanca, Río Verde, San Gregorio, Cabo de Hornos (Ex Navarino), Antártica, Primavera, Timaukel, Torres del Paine.

<sup>3</sup> La realización del estudio que en este documento se presenta ha sido posible gracias a la colaboración de: CADEM S.A., que tuvo a su cargo la recolección de datos a través del trabajo en terreno, y DATAVOZ/STATCOM, que realizó el diseño y selección de la muestra, la supervisión externa del trabajo en terreno, así como la validación y cálculo de los ponderadores.

Para la aplicación del cuestionario se utilizaron dispositivos móviles mediante el sistema CAPI (Computer Assisted Personal Interview, Entrevistas Personales Asistidas por Computador en español). Para la programación del cuestionario se utilizó el software MobiNet. Este es un software multiplataforma, que permite gestionar el trabajo de campo y, al mismo tiempo, realizar las encuestas por parte de los encuestadores. La programación del cuestionario reduce los llamados “errores de salto”, que pueden surgir cuando el encuestador equivoca la pregunta siguiente en una secuencia de preguntas que dependen de una respuesta anterior. Además, esta programación no permite al encuestador avanzar a la siguiente pregunta si no se ha ingresado una respuesta. De esta forma, el uso de dispositivos móviles reduce significativamente los errores cuando se compara con la aplicación del cuestionario en papel<sup>4</sup>.

## 2. Consideraciones generales

Este estudio representa un estudio académico y objetivo de las actitudes y predisposiciones de todos los habitantes del país, en las fechas señaladas.

A través de este tipo de estudios se desea conocer las preocupaciones, preferencias y necesidades de la población de 18 años o más con respecto a la inmigración en el país. Así, la finalidad del estudio es proporcionar información precisa y confiable a los principales actores del país, necesaria para una correcta toma de decisiones en los diferentes ámbitos del quehacer nacional. De esta forma, la encuesta de opinión pública sirve de puente, de vehículo de comunicación entre la población chilena y quienes se interesan en comprenderla, contribuyendo con ello al fortalecimiento de la democracia.

En este estudio se han utilizado los procedimientos más rigurosos en la recolección y análisis de la información que se presenta. Los resultados que se exponen reflejan el estado de la opinión pública dentro de límites estadísticos establecidos, durante un momento determinado del tiempo. En este sentido, la encuesta no constituye una predicción de opiniones futuras.

---

<sup>4</sup> Ver por ejemplo de Leeuw, E. D. (1994). Computer Assisted Data Collection, Data Quality and Costs: A Taxonomy And Annotated Bibliography. *BMS: Bulletin of Sociological Methodology / Bulletin de Méthodologie Sociologique*, 44, 60-72; Caeyers, B., Chalmers, N. y De Weerd, (2010) J. A Comparison of CAPI and PAPI Through a Randomized Field Experiment. *SSRN Electronic Journal* - November 2010; y Schräpler, J.-P., Schupp, J., & Wagner, G. G. (2010). Changing from PAPI to CAPI: Introducing CAPI in a Longitudinal Study. *Journal of Official Statistics*, 26(2), 233-269.

### 3. Metodología

#### A. Introducción

El principio metodológico que guió este estudio es que la investigación eficaz y precisa debe basarse en una muestra verdaderamente representativa del universo de interés. Para el CEP, este universo lo constituyen los habitantes de Chile de 18 años y más. El proceso de selección de la muestra, como se detallará más adelante, es probabilístico, por conglomerados y trietápica.

El año 2010, el CEP decidió introducir una modificación al diseño muestral, relacionada con el tratamiento de la no respuesta. A partir de la primera medición de ese año, se eliminó el sistema de reemplazos y las pérdidas por no respuesta se compensan mediante una sobredimensión de la muestra. Esta sobredimensión está calculada en base a las tasas de respuestas observadas a través del tiempo en las encuestas del CEP. Esta modificación metodológica representa una mejora en el proceso de estimación, elevando los estándares técnicos del estudio, y equiparando las características técnicas con aquellas utilizadas en las encuestas más importantes a nivel internacional.

En este estudio, se seleccionó una muestra probabilística por conglomerados multietápica que incluyó una muestra total de 3.008 personas. El número de entrevistas personales concluidas y utilizables corresponde a 1.848, lo que implica una tasa de respuesta de un 61,8%.<sup>5</sup> De estas 1.848 encuestas, 499 están ubicadas en la macrozona norte. Esto equivale a un 27% de la muestra total encuestada. Las entrevistas fueron realizadas a lo largo de todo el país entre los días 24 de septiembre y 18 de noviembre de 2023.

El margen de error para muestras de este tamaño y diseño (probabilidad aleatoria por conglomerados) es de  $\pm 2,3$  puntos porcentuales con un nivel de confianza de un 95%. Lo anterior significa que, en 95 de cada 100 muestras de este tipo, frente a una pregunta de dos opciones donde cada una obtiene un 50% (varianza máxima), para una determinada pregunta, el valor poblacional se encuentra en el intervalo  $\pm 2,3$  puntos porcentuales del valor muestral. Para un mismo tamaño muestral, a medida que el nivel de respuesta se aleja del 50%, el margen de error disminuye. Por otro lado, para un mismo nivel de respuesta, al disminuir el tamaño muestral el margen de error aumenta.

---

<sup>5</sup> Siguiendo las definiciones desarrolladas por AAPOR (American Association of Public Opinion Research), la tasa de respuesta (RRI) corresponde a la proporción de entrevistas completas realizadas respecto del total de la muestra diseñada. La fórmula es:  $RRI = \frac{\text{Entrevistas completadas}}{\text{Entrevistas completadas} + \text{Entrevistas parciales} + \text{Rechazos} + \text{Otra falta de respuesta}}$ .

## B. Diseño de la muestra

En este estudio la muestra es probabilística, de modo que existe una probabilidad calculable y conocida de que todos los adultos de la población objetivo a lo largo del país sean seleccionados. La distribución de la muestra es proporcional a la distribución de la población de 18 años y más del país, con una sobremuestra en las regiones de Arica y Parinacota, Tarapacá, Antofagasta y Atacama. Esta distribución regional de la población se obtuvo de las proyecciones de población del INE al 30 de junio de 2023 en base al Censo 2017. Esta información permitió establecer, en primer lugar, la estratificación regional y en segundo lugar, la estratificación según zona, urbana y rural. En el **Cuadro N° 1** se presenta la distribución por región de la población mayor de 18 años.

**CUADRO N° 1.** POBLACIÓN OBJETIVO SEGÚN REGIÓN Y ZONA

Código	Regiones	POBLACIÓN		
		URBANO	RURAL	TOTAL
1	TARAPACÁ	278.946	16.735	295.681
2	ANTOFAGASTA	519.467	20.773	540.240
3	ATACAMA	215.754	21.177	236.931
4	COQUIMBO	541.332	122.346	663.678
5	VALPARAÍSO	1.449.220	129.462	1.578.682
6	LIBERTADOR GENERAL BERNARDO O'HIGGINS	588.513	197.711	786.224
7	MAULE	664.392	233.689	898.081
8	BIOBÍO	1.162.484	144.944	1.307.428
9	LA ARAUCANÍA	566.107	222.886	788.993
10	LOS LAGOS	519.209	183.582	702.791
11	AYSÉN DEL GENERAL CARLOS IBÁÑEZ DEL CAMPO	65.470	15.998	81.468
12	MAGALLANES Y DE LA ANTÁRTICA CHILENA	131.865	11.973	143.838
13	METROPOLITANA DE SANTIAGO	6.309.905	234.939	6.544.844
14	LOS RÍOS	232.852	87.542	320.394
15	ARICA Y PARINACOTA	180.728	16.354	197.082
16	ÑUBLE	285.731	120.520	406.251
TOTAL		13.711.975	1.780.631	15.492.606

**Fuente:** Población de 18 años y más de acuerdo a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017, Instituto Nacional de Estadísticas.

El diseño muestral que se utiliza corresponde a uno estratificado por conglomerados en tres etapas, cada una de ellas definidas de la siguiente forma:

- Unidad de Muestreo Primaria (UMP) : Manzana/Entidad<sup>6</sup>
- Unidad de Muestreo Secundaria (UMS) : Vivienda
- Unidad de Muestreo Terciaria (UMT) : Personas de 18 años y más

En tanto la estratificación está dada por las variables región y zona geográfica.

### C. Etapas de muestreo

#### *Primera etapa*

En la primera etapa del proceso de muestreo se fijó en 8 el número de entrevistas por conglomerado, entendido este como una manzana o entidad poblada (con 3 viviendas o más). Este número (8 entrevistas por conglomerado) fue escogido porque, desde un punto de vista logístico, aumenta al máximo la eficiencia en función del costo del entrevistador dentro de un conglomerado específico, asimismo reduce al mínimo la presencia incómoda del entrevistador en el vecindario, la cual podría afectar la integridad de los datos de la encuesta. Al aplicarse **8** entrevistas por conglomerado al número total de entrevistas en la muestra, esta es de **3.008**, para las **376 unidades de muestreo primarias (UMP)** a ser seleccionadas en la primera etapa.

El número de UMP se eligió considerando que existe una tasa de no respuesta regular. Esto implica un nivel de sobredimensión para la muestra de modo que se pueda obtener un tamaño muestral efectivo aproximado de 1.850 entrevistas. Para ello se analizaron las tasas de reemplazo de la medición de octubre-noviembre 2022, según región y zona urbana rural. Así también en aquellos estratos en que el número de UMP fuese menor a 2 unidades se exigió un tamaño mínimo de 2. En el **Cuadro N° 2** se incluyen la tasa de respuesta promedio de tales mediciones, insumo clave para el proceso de ajuste de sobre-muestra en la medición actual.

---

<sup>6</sup> Según definición INE, Manzana/Entidad corresponde a conglomerado de viviendas en Zonas Urbanas y Rurales, respectivamente.



## CUADRO N° 2. TASA DE RESPUESTA SEGÚN REGIÓN

Código	Regiones	TOTAL
1	TARAPACÁ	69,2%
2	ANTOFAGASTA	58,2%
3	ATACAMA	66,1%
4	COQUIMBO	58,2%
5	VALPARAÍSO	46,9%
6	LIBERTADOR GENERAL BERNARDO O'HIGGINS	63,1%
7	MAULE	65,0%
8	BIOBÍO	69,3%
9	LA ARAUCANÍA	63,5%
10	LOS LAGOS	63,9%
11	AYSÉN DEL GENERAL CARLOS IBÁÑEZ DEL CAMPO	67,0%
12	MAGALLANES Y DE LA ANTÁRTICA CHILENA	71,4%
13	METROPOLITANA DE SANTIAGO	64,0%
14	LOS RÍOS	62,5%
15	ARICA Y PARINACOTA	56,3%
16	ÑUBLE	59,5%
TOTAL		61,8%

**Fuente:** Elaboración Propia en base a Disposición Final de Casos de la Muestra.

Las UMP fueron distribuidas según región y zona urbana/rural y la tasa de no respuesta de la medición de septiembre-noviembre 2023 (muestra sobredimensionada), en todas las regiones del país sobre la base de la contribución de la región a la población tanto urbana como rural, tal como se describe en el **Cuadro N° 3**.

### CUADRO N°3. NÚMERO DE CONGLOMERADOS POR REGIÓN Y ZONA

Código	Regiones	URBANO	RURAL	TOTAL
1	TARAPACÁ	26	2	28
2	ANTOFAGASTA	33	2	35
3	ATACAMA	20	2	22
4	COQUIMBO	13	2	15
5	VALPARAÍSO	29	2	31
6	LIBERTADOR GENERAL BERNARDO O'HIGGINS	10	2	12
7	MAULE	13	3	16
8	BIOBÍO	21	2	23
9	LA ARAUCANÍA	10	3	13
10	LOS LAGOS	11	2	13
11	AYSÉN DEL GENERAL CARLOS IBÁÑEZ DEL CAMPO	2	0*	2
12	MAGALLANES Y DE LA ANTÁRTICA CHILENA	2	0*	2
13	METROPOLITANA DE SANTIAGO	122	3	125
14	LOS RÍOS	6	2	8
15	ARICA Y PARINACOTA	11	2	13
16	ÑUBLE	5	2	7
TOTAL		334	31	365

**Notas:** Aquellas UMP que registraron menos encuestas del mínimo esperado, es decir, aquellas que registraron menos de dos encuestas, se unieron con alguna otra de su mismo estrato para formar una unidad de análisis complejo de los datos, que están incluidas de esa forma en la tabla.

**(\*)** No se incluyen UMP pues el tamaño de muestra esperado es igual o inferior a 2 para el tamaño total esperado.

**Fuente:** Elaboración Datavoz

Utilizando la mejor información digital disponible, es decir, los datos del Censo Nacional, se elaboró un listado acumulativo de población por provincias, comunas, distrito, zona y manzana en el caso urbano, y de provincias, comunas, distrito, localidad y entidad en el rural (ordenado geográficamente).

Tanto en el caso urbano como rural, se estableció un intervalo fijo para cada región al dividir la población total de la región por el número de UMP asignadas a la misma. Dentro de cada una de ellas se siguió un proceso de selección puramente aleatoria de modo que existiera una probabilidad calculable de que cada individuo (representado por una estadística de población) fuera escogido como punto de referencia para una UMP.

Lo anterior se realizó mediante un proceso computarizado, aleatorio y proporcional a la población de 18 años y más para seleccionar las manzanas en las zonas urbanas y las entidades en el caso rural. Se construyó un programa computacional para seleccionar las 334 manzanas urbanas y las 31 entidades rurales para la muestra. Todas las manzanas y entidades fueron identificadas por un número y localizadas en un mapa censal.

El marco muestral utilizado corresponde a las bases de datos y cartografía asociada del Censo 2017. Dicha información es la más actualizada a la fecha, lo que permite obtener una muestra cuya cobertura es mayor y más precisa.

### *Segunda etapa*

La segunda etapa del proceso de muestreo fue la selección de grupos familiares (viviendas) al interior de las UMP. Se proporcionaron a los encuestadores pautas para la selección de las viviendas en las manzanas y entidades seleccionadas, de modo que se pudieran escoger al azar las viviendas por seleccionar en cada conglomerado. Después de empadronar cada manzana y entidad seleccionada, se siguió un procedimiento de trayecto aleatorio o muestreo sistemático en que cada enésima vivienda fue incluida en la muestra hasta que se identificaba un total de 8 familias (Unidades de Muestreo Secundarias).

### *Tercera etapa*

La tercera etapa consistió en elegir en cada vivienda seleccionada a una persona para ser encuestada. Los encuestadores recibieron instrucciones de aplicar un proceso de selección aleatoria para identificar a la persona que sería entrevistada.

Tanto en la segunda como en la tercera etapa de muestreo el entrevistador debe realizar un mínimo de tres visitas en tres días diferentes para intentar conseguir la vivienda o la persona originales a ser entrevistada. En esas visitas, el entrevistador debe dejar en la vivienda (o al sujeto seleccionado según sea el caso) una carta firmada por el Director del CEP en la cual se explica la naturaleza del CEP y los objetivos del estudio que se está realizando.

## **4. Factores de Ponderación y ajuste post-encuesta**

El proceso de ponderación y ajuste post-encuesta considera la creación de factores que incorporen en su cálculo las probabilidades de selección en cada una de las etapas y los ajustes de No respuesta. Dicho factor permite que la base de datos final a utilizar represente adecuadamente al universo en estudio y corrija los sesgos asociados a la sub o sobre representación de ciertos grupos poblacionales por aspectos asociados a no respuesta. Se construyó un ponderador diferente a la Encuesta CEP N° 90 nacional que contempla el aumento de la muestra en la macrozona norte, por lo tanto, se ajusta a la población de cada macrozona.

El procedimiento de creación del factor de ponderación contempla:

- Las probabilidades de selección en cada una de las etapas del diseño (3).
- La corrección de estas por factores asociados a NO respuesta o por sub o sobre representaciones de la muestra final.

En el primer caso, se obtiene por el producto de los inversos de las probabilidades de selección, en tanto para la obtención de los segundos, factores de corrección por no respuesta se aplica el procedimiento que sigue:

$$\omega_i = \left( n * \frac{N_i}{\sum_i N_i} \right) / n_i \quad \text{para el } i\text{-ésimo estrato.}^7$$

En que:

- $N_i$ : tamaño población al del  $i$ -ésimo estrato.
- $n_i$ : tamaño muestral final del  $i$ -ésimo estrato.
- $n$ : tamaño de la muestra total.

Dichas fórmulas cumplen que  $\sum_j \sum_i \omega_{ij} * n_{ij} = n$  Dichas fórmulas cumplen que  $\sum_j \sum_i N_{ij} = N$ , es el tamaño poblacional del universo objetivo.

Estos factores son creados considerando la distribución de la población objetivo, utilizando los datos que proporcionan las proyecciones de población basadas en el **Censo 2017** para la población objetivo de Chile por región, urbano - rural, género y tramos de edad. En el **Cuadro N° 4-1 A, N° 4-1 B, N° 4-1 C y N° 4-1 D** se incluyen dichos factores.

#### **CUADRO N° 4-1 A.**

##### **MACROZONA NORTE: FACTORES DE AJUSTE POR NO RESPUESTA SEGÚN ZONA, SEXO Y EDAD**

Zona	Sexo	Edad	Factores
URBANO	HOMBRE	18-29	1,1845
		30-44	1,7336
		45-54	1,3913
		55-64	1,0510
		65 y más	0,6895
	MUJER	18-29	1,2872
		30-44	0,9332
		45-54	0,8318
		55-64	0,6329
		65 y más	0,6156
RURAL	HOMBRE	18-29	0,9741
		30-44	1,2378
		45-54	1,6639
		55-64	0,1805
		65 y más	0,3420
	MUJER	18-29	0,3407
		30-44	0,5251
		45-54	0,1252
		55-64	0,2057
		65 y más	0,1341

**Fuente:** Elaboración Datavoz

<sup>7</sup> Corresponde a la celda que se genera del cruce las variables Género, Edad y Zona urbana/rural.

**CUADRO N° 4-1 B.****MACROZONA CENTRO: FACTORES DE AJUSTE POR NO RESPUESTA  
SEGÚN ZONA, SEXO Y EDAD**

Zona	Sexo	Edad	Factores
URBANO	HOMBRE	18-29	1,1932
		30-44	1,5721
		45-54	1,1979
		55-64	1,9636
		65 y más	0,7579
	MUJER	18-29	1,1094
		30-44	0,9072
		45-54	0,4918
		55-64	0,7020
		65 y más	1,0993
RURAL	HOMBRE	18-29	2,3227
		30-44	1,5231
		45-54	1,0459
		55-64	1,3456
		65 y más	1,1190
	MUJER	18-29	0,3962
		30-44	0,6868
		45-54	0,5363
		55-64	0,5332
		65 y más	0,4868

Fuente: Elaboración Datavoz

**CUADRO N° 4-1 C.****MACROZONA SUR: FACTORES DE AJUSTE POR NO RESPUESTA  
SEGÚN ZONA, SEXO Y EDAD**

Zona	Sexo	Edad	Factores
URBANO	HOMBRE	18-29	1,4348
		30-44	1,6350
		45-54	1,9152
		55-64	1,4571
		65 y más	0,6301
	MUJER	18-29	1,9887
		30-44	1,2779
		45-54	0,8687
		55-64	0,8348
		65 y más	0,6078
RURAL	HOMBRE	18-29	0,9248
		30-44	1,7627
		45-54	4,2129
		55-64	1,6533
		65 y más	1,1385
	MUJER	18-29	0,7299
		30-44	0,6503
		45-54	0,5757
		55-64	0,6418
		65 y más	1,3915

Fuente: Elaboración Datavoz

**CUADRO N° 4-1 D.****MACROZONA METROPOLITANA : FACTORES DE AJUSTE  
POR NO RESPUESTA SEGÚN ZONA, SEXO Y EDAD**

Zona	Sexo	Edad	Factores
URBANO	HOMBRE	18-29	1,4591
		30-44	1,6860
		45-54	1,2220
		55-64	1,2752
		65 y más	0,7138
	MUJER	18-29	1,0086
		30-44	1,0771
		45-54	0,8293
		55-64	0,5684
		65 y más	0,6123
RURAL	HOMBRE	18-29	0,7992
		30-44	1,2266
		45-54	1,7157
		55 y más*	2,4615
			0,4602
	MUJER	18-29	1,7090
		30-44	0,6777
		45-54	0,3673
		55-64	0,9398
		65 y más	

**Fuente:** Elaboración Datavoz

\* Se agruparon las categorías 55-64 y 65 y más, dada la no respuesta de casos entre 55 a 64 años.

Macrozona	Región	Factores
Norte	1. Tarapacá	0,9792
	2. Antofagasta	1,0192
	3. Atacama	0,6508
	15. Arica y Parinacota	1,0148
Centro	4. Coquimbo	0,6675
	5. Valparaíso	0,9401
	6. Libertador Bernardo O'Higgins	1,2326
	7. Maule	0,9757
	16. Ñuble	0,9214
RM	13. Metropolitana de Santiago	0,9878
Sur	8. Biobío	1,1531
	9. Araucanía	1,0009
	10. Los Lagos	1,3046
	11. Aysén	1,3227
	12. Magallanes y la Antártida Chilena	2,1671
	14. Los Ríos	0,6640

**Fuente:** Elaboración Datavoz

El resultado final de esta acción queda reflejado en un campo adicionado a la base de datos de trabajo denominado pond.



## 5. Variables de Diseño Complejo

Como se explica en las secciones anteriores, el diseño muestral de la encuesta CEP corresponde a un diseño complejo estratificado por conglomerados y multietápico. Para el análisis riguroso de la información, y en particular, para la inferencia estadística, es importante tomar en cuenta su diseño complejo. Para esto, en la base de datos se ha incorporado la variable estrato, que como su nombre lo indica, da cuenta de los estratos utilizados para la estratificación, y la variable secu que corresponde al pseudo-conglomerado. Esto ya que en casos en que solo se completó una encuesta efectiva a una UMP, para asegurar un número mínimo de encuestas realizadas en un conglomerado para permitir el cálculo de la varianza, se ha construido un pseudo-conglomerado agregando la UMP con una sola observación a otra UMP de la misma comuna. En el Anexo 2 se incluye el código para incorporar el diseño complejo en Stata y R.

## 6. Resultados levantamiento de la encuesta

### A. Validación de la Muestra

Los **Cuadros N° 5-1A a N°5-1D** muestran los valores censales, muestrales y muestrales ponderados para un conjunto de variables relevantes para la descripción de la población objetivo, definida en este estudio.

### CUADRO N° 5-1 A.

#### MACROZONA NORTE VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD (en porcentaje)

Característica demográfica		Valor Censal	Muestra encuestada	Muestra ajustada
Urbanidad	Urbano	94,1	87,4	93,7
	Rural	5,9	12,6	6,3
Sexo	Masculino	50,1	38,9	50,1
	Femenino	49,9	61,1	49,9
Edad	18-29	24,6	19,0	25,0
	30-44	32,6	26,5	32,2
	45-54	16,5	15,8	16,5
	55-64	12,9	17,0	12,7
	65 y más	13,3	21,6	13,6

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

**CUADRO N° 5-1 B.****MACROZONA CENTRO VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD**  
(en porcentaje)

Característica demográfica		Valor Censal	Muestra encuestada	Muestra ajustada
Urbanidad	Urbano	81,5	79,6	81,5
	Rural	18,5	20,4	18,5
Sexo	Masculino	48,5	36,1	47,9
	Femenino	51,5	63,9	52,1
Edad	18-29	20,4	17,4	20,2
	30-44	27,4	24,8	27,9
	45-54	16,6	19,4	15,7
	55-64	15,7	14,1	16,2
	65 y más	19,8	24,3	20,0

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

**CUADRO N° 5-1 C.****MACROZONA SUR VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD**  
(en porcentaje)

Característica demográfica		Valor Censal	Muestra encuestada	Muestra ajustada
Urbanidad	Urbano	80,1	78,1	82,0
	Rural	19,9	21,9	18,0
Sexo	Masculino	48,5	38,1	48,1
	Femenino	51,5	61,9	51,9
Edad	18-29	20,8	14,4	21,5
	30-44	27,8	23,1	27,5
	45-54	17,2	16,9	17,2
	55-64	15,6	17,5	16,2
	65 y más	18,5	28,1	17,5

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

**CUADRO N° 5-1 D.****MACROZONA CENTRO VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD**  
(en porcentaje)

Característica demográfica		Valor Censal	Muestra encuestada	Muestra ajustada
Urbanidad	Urbano	96,4	96,0	96,4
	Rural	3,6	4,0	3,6
Sexo	Masculino	49,1	36,2	49,1
	Femenino	50,9	63,8	50,9
Edad	18-29	22,8	17,3	22,8
	30-44	32,0	24,5	32,0
	45-54	16,1	16,1	16,1
	55-64	13,6	17,3	13,3
	65 y más	15,5	25,0	15,8

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

**CUADRO N° 5-2 A.****MACROZONA NORTE VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD**  
(en porcentaje)

Región	Valor Censal	Muestra encuestada	Muestra ajustada
1. TARAPACÁ	23,3	27,1	23,2
2. ANTOFAGASTA	42,5	37,1	40,1
3. ATACAMA	18,7	21,4	20,2
15. ARICA Y PARINACOTA	15,5	14,4	16,5

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

**CUADRO N° 5-2 B.****MACROZONA CENTRO VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD**  
(en porcentaje)

Región	Valor Censal	Muestra encuestada	Muestra ajustada
5. VALPARAÍSO	23,3	27,1	23,2
6. LIBERTADOR BERNARDO O'HIGGINS	42,5	37,1	40,1
7. MAULE	18,7	21,4	20,2
16. ÑUBLE	15,5	14,4	16,5

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

**CUADRO N° 5-2 C.****MACROZONA SUR VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD**  
(en porcentaje)

Región	Valor Censal	Muestra encuestada	Muestra ajustada
8. BIOBÍO	39,1	36,6	39,9
9. ARAUCANÍA	23,6	23,4	22,7
10. LOS LAGOS	21,0	20,0	20,7
11. AYSÉN	2,4	3,1	2,3
12. MAGALLANES Y ANTÁRTIDA CHILENA	4,3	2,8	5,1
14. LOS RÍOS	9,6	14,1	9,3

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

## CUADRO N° 5-2 D.

### MACROZONA SUR VALIDACIÓN DE LA MUESTRA POR ZONA, SEXO Y EDAD

(en porcentaje)

Región	Valor Censal	Muestra encuestada	Muestra ajustada
13. METROPOLITANA DE SANTIAGO	100,0	100,0	100,0

**Nota:** Los valores censales están calculados en base a las proyecciones de población al 30 de junio de 2023, con base en Censo 2017.

**Fuente:** Elaboración Datavoz.

Sobre la base de la metodología muestral empleada y de la validación de la muestra mediante comparaciones con parámetros estadísticos conocidos del censo, podemos concluir que este estudio representa con exactitud las actitudes y opiniones de adultos chilenos respecto de las preguntas de la encuesta dentro del margen de error estipulado ( $\pm 2,8\%$ ) y en las condiciones antes descritas.

#### B. Resultados del Trabajo en Terreno

Tal como se señalaba anteriormente, el proceso de selección de la muestra definió en esta ocasión una muestra de 3.008 personas. A partir de esta muestra se obtuvieron un total de **1.848** entrevistas efectivas. Siguiendo las definiciones desarrolladas por AAPOR (American Association of Public Opinion Research) en 2023, es posible evaluar la efectividad del trabajo de terreno.

Utilizando los procedimientos definidos por AAPOR, entonces, los resultados del levantamiento de la encuesta entregan los siguientes resultados:

**Tasa de Respuesta (RR1):**<sup>8</sup> Corresponde a la proporción de entrevistas completas realizadas respecto del total de la muestra diseñada. En esta ocasión, la tasa de respuesta alcanzó a un **61,8%**.

**Tasa de Rechazo (REF1):** Corresponde a la proporción de casos en los que se produjo rechazo a contestar la entrevista, ya sea a nivel de hogar como a nivel de personas seleccionadas. En esta ocasión el rechazo total alcanzó a un **19,3%**.

**Tasa de Cooperación (COOP1):** Corresponde al número de entrevistas completas realizadas, dividido por el número de encuestas realizadas más aquellas donde hubo rechazo a contestar la encuesta, ya sea a nivel de hogar o a nivel de la persona seleccionada. En esta ocasión, la tasa de cooperación alcanzó un **76,1%**.

Las principales razones para la no realización de las entrevistas se detallan a continuación:

<sup>8</sup> Se incluyen en el denominador los 13 casos no elegibles.

**CUADRO N° 6.**

## RESULTADOS DEL LEVANTAMIENTO DE LA ENCUESTA

Región	Número de casos	Porcentaje sobre Total Muestra
<b>Entrevistas realizadas</b>		
Completas	1.848	61,8%
Parcial	0	0,0%
<b>No entrevistado, elegible</b>		
Rechazo en el hogar	532	17,7%
Entrevista interrumpida	4	0,1%
Acceso impedido	2	0%
Nadie en el hogar	517	17,2%
No puede atender	22	0,7%
Enfermo	4	0,1%
Rechazo individuo	48	1,6%
Otra razón	4	0,1%
Anuladas	4	0,1%
<b>No entrevistado, elegibilidad desconocida</b>		
Área peligrosa o de difícil acceso	0	0,0%
<b>No entrevistado, no elegible</b>		
Empresa, oficina de gobierno u otra organización	1	0%
Vivienda particular desocupada	6	0,2%
Vivienda de veraneo o de uso temporal	0	0,0%
Otros no elegible	6	0,2%

Fuente: Elaboración propia.

## 7. Cuestionario

El cuestionario utilizado en el Estudio de inmigración en la macrozona norte incluyó el mismo módulo especial sobre inmigración de la encuesta nacional CEP N° 90, con algunas variables del módulo socioeconómico. Este se aplicó de manera presencial por los encuestadores/as mediante dispositivos electrónicos (tablets o teléfonos smartphones). El cuestionario fue programado de manera tal que el orden de aparición depende de las respuestas entregadas a las preguntas precedentes. El cuestionario completo se encuentra en el siguiente enlace: [Cuestionario Estudio de inmigración en la macrozona norte](#).

El cuestionario del Estudio de inmigración en la macrozona norte tiene los siguientes módulos:

- Módulo Inmigración (MI)
- Módulo Socioeconómico (MSE)

## 8. Acceso y uso de la información

El Centro de Estudios Públicos pone a disposición pública todo el material recopilado en sus estudios. La base de datos, el cuestionario completo, el diccionario de variables, el Manual del Usuario (documentación técnica) y los resultados principales presentados en la conferencia de prensa son de libre acceso y están disponibles en la página web para su descarga.

En caso de requerir información adicional no publicada, solicitar aclaración respecto a la metodología, o tener dudas respecto a cómo utilizar la base de datos, puede ponerse en contacto con el Área de Opinión Pública del CEP a través de [encuestas@cepchile.cl](mailto:encuestas@cepchile.cl).

Cualquier publicación que incluya para su realización el uso de esta base de datos debe citar la fuente de la siguiente forma:

Centro de Estudios Públicos. Estudio de inmigración en la macrozona norte, septiembre-noviembre 2023 [Computer File]. CEP0088-v1. Santiago: Centro de Estudios Públicos.

## Anexo 1

### Diseño de la muestra

<b>REGIÓN/COMUNA</b>	<b>URBANA</b>	<b>Rural</b>
<b>ARICA Y PARINACOTA</b>	<b>11</b>	<b>2</b>
ARICA	11	2
<b>TARAPACÁ</b>	<b>27</b>	<b>2</b>
ALTO HOSPICIO	7	
IQUIQUE	20	2
<b>ANTOFAGASTA</b>	<b>33</b>	<b>2</b>
ANTOFAGASTA	22	
CALAMA	9	2
TOCOPILLA	2	
<b>ATACAMA</b>	<b>21</b>	<b>2</b>
CHAÑARAL	1	
COPIAPÓ	13	
FREIRINA	2	1
HUASCO	1	
TIERRA AMARILLA	1	
VALLENAR	3	1
<b>COQUIMBO</b>	<b>14</b>	<b>2</b>
CANELA		1
COMBARBALÁ	1	
COQUIMBO	4	
ILLAPEL	3	
LA SERENA	4	1
OVALLE	1	
PUNITAQUI	1	
<b>VALPARAÍSO</b>	<b>30</b>	<b>2</b>
CALERA	2	
CASABLANCA	2	
CONCÓN	1	
HIJUELAS		1
LA LIGUA	1	
LOS ANDES	1	
OLMUÉ		1
QUILPUÉ	4	
QUINTERO	1	
SAN ANTONIO	1	
SAN FELIPE	2	
VALPARAÍSO	7	
VIÑA DEL MAR	8	
<b>METROPOLITANA</b>	<b>126</b>	<b>3</b>



BUIN	1	
CERRO NAVIA	2	
COLINA	1	1
CONCHALÍ	4	
EL BOSQUE	4	
ESTACIÓN CENTRAL	3	
HUECHURABA	3	
INDEPENDENCIA	2	
ISLA DE MAIPO	1	
LA CISTERNA	2	
LA FLORIDA	5	
LA GRANJA	1	
LA PINTANA	1	
LA REINA	1	
LAMPA	2	1
LAS CONDES	4	
LO BARNECHEA	3	
LO ESPEJO	2	
LO PRADO	4	
MACUL	1	
MAIPÚ	6	
MARÍA PINTO		1
MELIPILLA	3	
ÑUÑO A	2	
PADRE HURTADO	2	
PEDRO AGUIRRE CERDA	2	
PEÑALOLÉN	7	
PROVIDENCIA	3	
PUDAHUEL	5	
PUENTE ALTO	19	
QUILICURA	3	
QUINTA NORMAL	4	
RECOLETA	4	
RENCA	1	
SAN BERNARDO	4	
SAN MIGUEL	3	
SAN RAMÓN	3	
SANTIAGO	5	
TALAGANTE	1	
VITACURA	2	
<b>LIBERTADOR BERNARDO O'HIGGINS</b>	<b>11</b>	<b>2</b>
GRANEROS	1	

LAS CABRAS		1
LITUECHE	1	
MOSTAZAL	1	
OLIVAR	1	
RANCAGUA	4	
RENGO	1	1
SAN FERNANDO	2	
<b>MAULE</b>	<b>13</b>	<b>3</b>
CURICÓ	3	
LINARES	1	
LONGAVÍ		1
MOLINA	1	
RETIRO		1
RÍO CLARO	1	
SAN CLEMENTE		1
SAN JAVIER	1	
TALCA	6	
<b>BIOBÍO</b>	<b>21</b>	<b>2</b>
CABRERO	1	
CHIGUAYANTE	1	
CONCEPCIÓN	3	1
CORONEL	1	
CURANILAHUE	1	
HUALPÉN	1	
HUALQUI	1	
LAJA	1	
LOS ÁNGELES	2	
PENCO	2	
QUILLECO		1
SAN PEDRO DE LA PAZ	4	
TALCAHUANO	2	
TOMÉ	1	
<b>ÑUBLE</b>	<b>6</b>	<b>2</b>
BULNES	1	
CHILLÁN	2	1
CHILLÁN VIEJO	1	
COIHUECO	1	
EL CARMEN		1
SAN CARLOS	1	
ARAUCANÍA	11	3
CARAHUE		1
LAUTARO	1	
MELIPEUCO	1	

PADRE LAS CASAS	2	
PURÉN		1
TEMUCO	6	
VICTORIA	1	
VILLARRICA		1
<b>LOS LAGOS</b>	<b>11</b>	<b>2</b>
CALBUCO		1
CASTRO	1	1
OSORNO	6	
PUERTO MONTT	3	
PURRANQUE	1	
<b>LOS RÍOS</b>	<b>6</b>	<b>2</b>
LOS LAGOS	1	
PAILLACO		1
VALDIVIA	5	1
<b>AYSÉN</b>	<b>2</b>	
AYSÉN	1	
COYHAIQUE	1	
<b>MAGALLANES</b>	<b>2</b>	
PUNTA ARENAS	2	

## Anexo 2

### Procesamiento de la información

---

#### A. Variables de diseño complejo

Para el análisis de los datos se debe utilizar un programa estadístico que permita incorporar tanto el ponderador como las variables de diseño complejo en el análisis. Como se detalla anteriormente, el ponderador es necesario para el correcto análisis e interpretación de los resultados. Este corresponde a la variable *pond* en la base de datos consolidada.

El diseño muestral de las encuestas CEP es del tipo complejo, por lo que se requiere incorporar las características del diseño para una correcta estimación del error estándar. En particular, además del ponderador, se debe utilizar programas estadísticos que permitan incorporar la información sobre estratos y conglomerados, o UMP.

En concreto, las variables que toman en cuenta el diseño complejo son:

- **estrato:** considera los estratos definidos según región y zona. Debido a que en algunos años solo se consideró una UMP por estrato, siguiendo recomendaciones en la literatura especializada<sup>9</sup>, se han agregado en un solo estrato.
- **secu:** corresponde al conglomerado, el cual, por motivos de confidencialidad de la información, así como para asegurar una correcta estimación de la varianza, se han identificado con números correlativos y en algunos casos, fusionados con otros conglomerados similares.

Estas variables solo están disponibles desde la encuesta 29 en adelante, por lo que, para las encuestas anteriores, se debe incorporar solo el ponderador para el análisis.

Es importante destacar que los estratos son equivalentes entre encuestas, mientras las *secu* tienen un número único, que no se repite entre encuestas.

Por último, se debe recordar que para analizar subgrupos dentro de una misma encuesta es necesario identificarla como una subpoblación y no se debe filtrar la información, ya que, si se filtra, se subestima la varianza del error.

---

<sup>9</sup> Heeringa, Steven G.; West, Brady T.; Berglund, Patricia A. (2010). Applied Survey Data Analysis (Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences) (p. 104). CRC Press.

## B. Ejemplos análisis de la información en STATA

Recomendamos fijar el número máximo de variables permitidas en su máximo antes de cargar la base y comenzar el análisis:

```
clear all  
set maxvar 32767
```

Antes de comenzar el análisis, se declara el diseño muestral:

```
svyset secu [pweight=pond], strata(estrato) singleunit(certainty) vce(linearized)
```

### Ejemplo 1:

Estimación del porcentaje de aprobación por sexo

#### Pregunta:

Independiente de su posición política, usted ¿aprueba o desaprueba la forma como está conduciendo el PRESIDENTE el gobierno?

#### Primera opción:

```
svy: prop eval_gob_1, over(sexo)
```

#### Segunda opción:

```
svy: tab sexo eval_gob_1, row se
```

### Ejemplo 2:

Estimación del porcentaje de aprobación de las personas que viven en zonas urbanas, por sexo

```
gen durb=[zona_u_r==1]
```

#### Primera opción:

```
svy, subpop(durb): prop eval_gob_1, over(sexo)
```

#### Segunda opción:

```
svy, subpop(durb): tab sexo eval_gob_1, row se
```

## C. Ejemplos análisis de la información en R

Para el análisis de la información, se sugiere utilizar el paquete `srvyr`.<sup>10</sup>

**Nota:** cualquier procesamiento de la información, tal como creación de nuevas variables, debe ser realizado antes de declarar el diseño. Se declara el diseño generando nuevo objeto, en el cuál es difícil la manipulación de variables.

```
install.packages("srvyr")
library(srvyr)
```

Se declara el diseño de la encuesta

```
svy_cep <- base_89 %>%
  as_survey(weights=pond, strata=estrato, ids=secu, nest=T)
```

### Ejemplo 1:

Estimación del porcentaje de aprobación por sexo

#### Pregunta:

Independiente de su posición política, usted ¿aprueba o desaprueba la forma como está conduciendo el PRESIDENTE el gobierno?

```
svy_cep %>%
  group_by(sexo, eval_gob_1) %>%
  summarise(proportion = survey_prop())
```

### Ejemplo 2:

Estimación del porcentaje de aprobación de las personas que viven en zonas urbanas, por sexo

```
svy_cep %>%
  filter(zona_u_r==1) %>%
  group_by(sexo, eval_gob_1) %>%
  summarise(proportion = survey_prop())
```

---

<sup>10</sup> Freedman Ellis, Greg et al. (2022). "srvyr: 'dplyr'-Like Syntax for Summary Statistics of Survey Data". R package version 1.1.2 <https://CRAN.R-project.org/package=srvyr>.