



Encuesta Nacional de Participación Cultural 2017

Informe metodológico post estratificación y cálculo de la variable “Nivel Socioeconómico”

(Producto N°5.2/Versión 02)

INSTITUTO NACIONAL DE ESTADÍSTICAS

Marzo / 2018

SUBDIRECCIÓN TÉCNICA
DEPARTAMENTO DE INVESTIGACIÓN Y DESARROLLO
SECCIÓN DE ESTADÍSTICAS SOCIALES

Encuesta Nacional de Participación Cultural 2017.
Informe metodológico post estratificación y cálculo de la variable “Nivel Socioeconómico”.
Producto N° 5.2/Versión 02.
Instituto Nacional de Estadísticas, Chile.
28 de marzo de 2018.

ÍNDICE

I.	Presentación	6
II.	Método de clasificación socioeconómica.....	7
II.1.	Clasificación socioeconómica de los hogares.....	7
II.2.	El procedimiento PRINCALS.....	8
III.	Variables Seleccionadas.....	9
III.1.	Variables levantadas en la Encuesta Nacional de Participación Cultural 2017 a considerar en el PRINCALS	10
III.2.	Variables del Sistema de Información Geográfico a considerar en el PRINCALS	14
IV.	Factor de Expansión para post estratificación	20
IV.1.	Ponderador a nivel de personas para post estratificación.....	20
IV.2.	Ajuste para calibrar al total poblacional por región, sexo y tramo etario	21
V.	Configuración de método de PRINCALS	24
V.1.	Conjunto de variables consideradas en escenarios de post estratificación.....	24
V.2.	Bondad del ajuste de los escenarios evaluados de post estratificación	24
V.3.	Selección de escenario definitivo de post estratificación	25
V.4.	Importancia de las variables en el escenario de post estratificación.....	26
VI.	Utilización de los factores de expansión	27
VI.1.	Utilización de los factores de expansión	28
VI.2.	Base a nivel de registro de personas.....	31
VII.	Resultados	34
VIII.	Bibliografía.....	38
IX.	Anexo	40

ÍNDICE DE TABLAS

Tabla 1	Variables evaluadas en la implementación del método PRINCALS.	10
Tabla 2.	Distribución de las personas según Grupo ocupacional CIUO modificado.	11
Tabla 3.	Distribución de las personas según Categoría ocupacional CISE modificado.....	12
Tabla 4.	Distribución de las personas según Nivel educacional modificado.....	13
Tabla 5.	Distribución de las personas según Tenencia de vehículos modificado.....	13
Tabla 6.	Distribución de variable Tipo de vivienda modificado	14
Tabla 7.	Proyecciones de Población por sexo y región.	22
Tabla 8.	Variables empleadas en los escenarios de pos estratificación.....	24

Tabla 9. Estadísticos de bondad de escenarios de post estratificación.	25
Tabla 10 .Descripción de nivel de registro y factor de expansión de bases de datos.	27
Tabla 11. Distribución poblacional según quintil socioeconómico.	28
Tabla 12. Distribución de personas mayores según nivel educacional y quintil.	29
Tabla 13. Distribución de personas mayores según nivel educacional y quintil.	30
Tabla 14. Distribución porcentual de personas mayores según nivel educacional y quintil.	30
Tabla 15 Distribución porcentual de personas mayores según nivel educacional y quintil.	31
Tabla 16. Distribución de total personas según nivel educacional y quintil socioeconómico.	32
Tabla 17. Comparación de distribución de personas mayores según quintil	32
Tabla 18 Distribución de hogares expandidos según quintil socioeconómico.....	33
Tabla 19. Distribución porcentual de personas por quintil según grupo ocupacional CIUO modificado. ...	35
Tabla 20. Distribución porcentual de personas por quintil según nivel educativo modificado.....	36
Tabla 21.Distribución porcentual de hogares por quintil según tipo de vivienda.	37
Tabla 22. Distribución porcentual de hogares por quintil según tramo de avalúo promedio.....	37

ÍNDICE DE GRÁFICOS

Gráfico 1. Distribución de frecuencia de variable Distancia en metros a comisaría.....	15
Gráfico 2. Distribución de frecuencia de variable Distancia en metros a áreas verdes.....	16
Gráfico 3. Distribución de frecuencia de variable Avalúo por manzana (en millones de pesos).....	17
Gráfico 4. Distribución de frecuencia de Número de establecimientos educacionales.	18
Gráfico 5. Distribución de frecuencia de Número de establecimientos de Salud.....	19
Gráfico 6 . Saturaciones en componentes.	26
Gráfico 7. Distribución de frecuencia de los puntajes PRINCALS asociados a los hogares.	34

I. Presentación

En el año 2015, el Instituto Nacional de Estadísticas (INE) suscribe un convenio con el Consejo Nacional de la Cultura y las Artes (CNCA), donde se encomienda al INE, en el marco de la Encuesta Nacional de Participación Cultural 2017, realizar el diseño muestral, seleccionar las unidades muestrales, elaborar los factores de expansión, así como generar la post estratificación socioeconómica de los hogares.

Una vez culminadas las primeras tres etapas convenidas, el INE se enfoca en generar la post estratificación, cuyo objetivo es clasificar a los hogares de acuerdo a un conjunto de variables relacionadas con su nivel socioeconómico, generando un puntaje para el hogar que permita un ordenamiento y clasificación de los mismos.

En consecuencia, para dar respuesta al objetivo planteado se emplea el método multivariado PRINCALS¹, por sus siglas en inglés *PRINCipal components analysis by means of Alternating Least Squares*, que combina reducción de la dimensionalidad de los datos, con escalamiento óptimo de las categorías de las variables; procedimiento implementado mediante el uso del software SPSS.

En el presente documento se detallan los aspectos conceptuales relacionados con la obtención de la clasificación socioeconómica, entre ellos: proceso de selección y adecuación de las variables, factores de expansión utilizados, configuración del método, y, por último, presentación de los resultados obtenidos.

¹ Gifi, A. 1985. *PRINCALS*. Leiden: Department of Data Theory, Internal Report.

II. Método de clasificación socioeconómica

Existen distintos métodos y enfoques para lograr una clasificación socioeconómica de los hogares, entre ellos destacan los enfoques basados en la capacidad para generar ingresos y otros que se apoyan en el análisis multidimensional. Estos últimos, apuntan a la caracterización de los miembros del hogar en relación a ciertas variables, principalmente midiendo carencias en ámbitos de salud, educación, trabajo, seguridad social, vivienda, cultura, entretención, entre otras.

El método aplicado para este estudio, presenta un enfoque relativo de la medición del nivel socioeconómico en relación al ingreso. Este enfoque considera aquellas variables² que, según varios estudios³, están altamente correlacionadas con el nivel de ingresos, como son la ocupación (grupo y categoría), educación, territorio, propiedad de vehículos, tipo de vivienda, entre otras.

II.1. Clasificación socioeconómica de los hogares

Con la clasificación socioeconómica se busca categorizar a cada hogar de la muestra, clasificándolo por medio del puntaje asignado a las personas del hogar mediante el método PRINCALS. La clasificación de los hogares se obtiene mediante el promedio de las puntuaciones de las personas que los componen, que en este caso se “categoriza” en quintiles de acuerdo al objetivo del estudio.

² Pueden existir muchas más variables que de alguna manera están correlacionadas con el nivel de ingresos, sin embargo, se consideran solo las que fueron investigadas en la encuesta y algunas provenientes de registros administrativos como avalúo fiscal, otras de registros cartográficos como distancia del centroide de la manzana seleccionada a áreas verdes, etc.

³ Entre ellos: Casen, 2013; León y Martínez, 2001; e INE, 2011.

II.2. El procedimiento PRINCALS

El procedimiento PRINCALS extiende el análisis de componentes principales ordinario a un contexto más general, en el cual se pueden emplear tanto variables cualitativas como cuantitativas. Utiliza un doble algoritmo, en que se escalan las categorías y valores observados de las variables, para luego determinar, mediante el método de mínimos cuadrados, los parámetros del análisis de componentes principales. Este proceso logra un escalamiento óptimo en el sentido que se repite iterativamente hasta lograr minimizar la pérdida de información debido a la representación de las variables por un número pequeño de dimensiones.

En la aplicación de este procedimiento existen varias opciones, entre ellos el número de dimensiones a analizar, el método de normalización, entre otras; lo que permite generar puntuaciones óptimas que sitúan a sujetos y variables en la misma escala, de modo que dos sujetos con la misma puntuación le corresponden características similares en cuanto a las variables incorporadas.

En la aplicación de este procedimiento, se contempla analizar solo la primera dimensión ya que explica la mayor cantidad de variabilidad asociada al nivel socioeconómico del hogar. Se elige la opción normalización simétrica, que maximiza la distancia entre las categorías de las variables, al mismo tiempo que maximiza las distancias entre los hogares disímiles.

A través de método PRINCALS se obtiene un puntaje para cada miembro del hogar, con distribución normal estándar (media cero y varianza uno). Como este puntaje toma valores positivos y negativos, y para una mayor comprensión, se realiza una transformación lineal⁴ de manera que todos los puntajes sean positivos. Enseguida se calcula el puntaje promedio por hogar, el cual se categoriza en quintiles.

⁴ A cada puntaje se le suma el valor absoluto del mínimo puntaje obtenido, de esta manera, el puntaje mínimo es cero o mayor que cero.

III. Variables Seleccionadas

Para aplicar el método PRINCALS se requiere considerar información de amplio espectro que dé cuenta de las condiciones socioeconómicas asociadas a los hogares. Es por esto, que es importante considerar características de las personas y su entorno. Es por ello que se utilizan variables disponibles a nivel de personas y viviendas que provienen del levantamiento de información de la propia encuesta; así como variables obtenidas a partir del Sistema de Información Geográfica del Instituto Nacional de Estadísticas⁵, las que se encuentran a nivel de manzanas y Distrito censal. La utilización de variables en distintos niveles, permite un mejor entendimiento del contexto en que se encuentra el hogar y por tanto una mejor clasificación.

En la tabla 1, se detallan las variables consideradas para configurar el método PRINCALS, tipo de variable, nivel de desagregación y su respectiva fuente.

⁵ El Sistema de Información Geográfica procesa información de diversas fuentes como registros administrativos y levantamiento en terreno, que permiten obtener un conjunto de variables a nivel de manzanas y distrito censal.

Tabla 1 Variables evaluadas en la implementación del método PRINCALS.

Variable	Tipo	Nivel de desagregación	Fuente
Grupo ocupacional CIOU	Nominal	Persona	Nacional de Participación 2017
Categoría ocupacional CISE	Nominal	Persona	
Nivel educacional	Ordinal	Persona	
Comuna	Nominal	Vivienda	
Tenencia de vehículo	Numérica	Vivienda	
Tipo de vivienda	Nominal	Vivienda	
Distancia en metros a comisaría	Numérica	Manzana	Sistema de Información geográfico INE
Distancia en metros a áreas verdes	Numérica	Manzana	
Avalúo fiscal promedio	Numérica	Manzana	
Número de establecimientos de salud	Numérica	Distrito censal	
Número de establecimientos de educación	Numérica	Distrito censal	

Elaborado por el Instituto Nacional de Estadísticas.

III.1. Variables levantadas en la Encuesta Nacional de Participación Cultural 2017 a considerar en el PRINCALS

A continuación, se expone el procesamiento realizado sobre las variables, con el objeto de adecuarlas para la aplicación del método de post estratificación.

Con el objeto de potenciar la correlación entre las variables, sobre aquellas que presentan valor *misino* por su propia naturaleza⁶, son agregadas nuevas categorías de respuesta, que rellenan estos valores perdidos. En todo caso, estas nuevas categorías que se añaden, deben conservar en lo posible el tipo de la variable a la cual se agregan, esto es, si la variable a la que se añaden es de tipo ordinal, las categorías añadidas deben mantener el carácter ordinal de la variable.

⁶Algunas variables por su naturaleza, no aplican para todos los casos. Por ejemplo, la categoría ocupacional, presenta missing cuando la persona no trabaja, ya sea porque está desocupada, está inactiva o simplemente porque es menor de edad o no declaró la actividad.

En ese sentido, a la variable “Grupo Ocupación CIUO” (a6_2, en base de datos), se le incluyen las categorías “Ocupados sin declaración de actividad”, “desocupados”, “Inactivos” y “Menores”; ésta última categoría contempla los menores de 15 años. La distribución de esta variable se observa en la tabla 2.

Tabla 2. Distribución de las personas según Grupo ocupacional CIUO modificado.

Categoría	Grupo ocupacional CIUO modificado	Frecuencia	Porcentaje	Porcentaje acumulado
Original	Total	35.415	100	100
	Miembros del poder ejecutivo y de los cuerpos legislativo y personal	283	0,8	0,8
	Profesionales científicos e intelectuales	2.130	6	6,8
	Técnicos y profesionales de nivel medio	1.673	4,7	11,5
	Empleados de oficina	1.198	3,4	14,9
	Trabajadores de los servicios y vendedores de comercios y mercados	3.157	8,9	23,8
	Agricultores y trabajadores calificados agropecuarios y pesqueros	198	0,6	24,4
	Oficiales, operarios y artesanos de artes mecánicas y de otros oficios	2.094	5,9	30,3
	Operadores de instalaciones y máquinas y montadores	1.497	4,2	34,5
	Trabajadores no calificados	2.340	6,6	41,1
	Fuerzas armadas	50	0,1	41,3
Incluida	Ocupados sin declaración de actividad	1.017	2,9	44,2
	Desocupados	975	2,8	46,9
	Inactivos	12.373	34,9	81,8
	Menores	6.430	18,2	100

Elaborado por el Instituto Nacional de Estadísticas.

A la variable “Categoría de ocupación CISE” (“a7”, en base de datos) se incluyen las categorías “Desocupados”, “Inactivos” y “Menores”; mientras que para los valores perdidos se genera una nueva categoría denominada “Perdidos”. Así, la variable queda distribuida como indica a en la tabla 3.

Tabla 3. Distribución de las personas según Categoría ocupacional CISE modificado.

Categoría	Categoría ocupacional CISE Modificado	Frecuencia	Porcentaje	Porcentaje acumulado
Original	Total	35.399	100	100
	Empleador o patrón	1.114	3,10	3,10
	Trabajador por cuenta propia	3.716	10,50	13,60
	Asalariado del sector privado	8.094	22,90	36,50
	Asalariado del sector público	2.178	6,10	42,70
	Personal de servicio doméstico puertas adentro	41	0,10	42,80
	Personal de servicio doméstico puertas afuera	299	0,80	43,60
	Familiar o personal no remunerado	181	0,50	44,10
Incluida	Desocupados	975	2,80	46,90
	Inactivos	12.371	34,90	81,80
	Menores	6.430	18,20	100
	Perdidos	16	-	100

Elaborado por el Instituto Nacional de Estadísticas.

Mientras que la variable “Nivel educacional” (“a2”, en base de datos) se reagrupa en cinco nuevas categorías: “Sin enseñanza básica”, que incluye personas que declaran no poseer educación o solo educación preescolar; “Enseñanza Básica”, personas que declaran poseer educación básica; “enseñanza media”, para quienes cursaron o están cursando enseñanza media científica humanista o técnico profesional; “Técnico de nivel superior o profesional”, que contiene personas que declaran haber cursado o estar cursando carreras técnicas o profesionales; y por último, la categoría “Postítulo, Magíster o Doctorado”, para quienes cursan o cursaron algún programa de este tipo, ver tabla 4.

Tabla 4. Distribución de las personas según Nivel educacional modificado.

Nivel educacional modificado	Frecuencia	Porcentaje	Porcentaje acumulado
Total	35.415	100,00	100,00
Sin enseñanza básica	3.091	8,73	8,73
Enseñanza básica	9.033	25,51	34,23
Enseñanza media	14.861	41,96	76,20
Técnico de nivel superior o profesional	8.096	22,86	99,06
Postítulo, magíster o doctorado	334	0,94	100,00

Elaborado por el Instituto Nacional de Estadísticas.

En el caso de la variable “Comuna”, no fue necesaria la incorporación de ningún ajuste para poder ser incluida en el PRINCALS, se presenta la respectiva distribución de frecuencia en el anexo 1. Cabe mencionar que los resultados a este nivel deben ser evaluados con precaución, ya que el tamaño de la muestra no garantiza representatividad a este nivel, sin embargo, es considerada en el PRINCALS por su relación con la clasificación socioeconómica de los hogares.

Al igual que el caso anterior, la variable “Tenencia de vehículos” (b12, en base de datos), se utiliza de la misma forma en que está presente en la base de datos. Sin embargo, se recodifica al valor cero (“0”) aquellos casos en donde los entrevistados declaran no haber tenido vehículos en los últimos 12 meses (“b11”, en base de datos). Así, la distribución de frecuencia de la variable Tenencia de vehículos, queda como se observa en la tabla 5.

Tabla 5. Distribución de las personas según Tenencia de vehículos modificado.

N° de vehículos	Frecuencia	Porcentaje	Porcentaje acumulado
Total	35.415	100,00	100,00
0	22.207	62,71	62,71
1	10.443	29,49	92,19
2	2.214	6,25	98,44
3	421	1,19	99,63
4	79	0,22	99,86
5	51	0,14	100,00

Elaborado por el Instituto Nacional de Estadísticas.

La última variable trabajada proveniente de la base de levantamiento es “Tipo de vivienda”, la cual se utiliza con las mismas categorías que posee en la base de datos original, salvo que, se incluye la categoría “Perdidos”, que da cuenta de los registros sin valor. La distribución de frecuencia de esta variable se observa en la tabla 6.

Tabla 6. Distribución de variable Tipo de vivienda modificado

Categoría	Tipo de vivienda modificado	Frecuencia	Porcentaje	Porcentaje acumulado
Original	Total	35.400	99,96	100
	Casa aislada (no pareada)	9.785	27,63	27,63
	Casa pareada por un lado	14.103	39,82	67,45
	Casa pareada por ambos lados	7.055	19,92	87,37
	Departamento en edificio con ascensor	1.132	3,2	90,57
	Departamento en edificio sin ascensor	2.921	8,25	98,82
	Pieza en casa contigua o conventillo	181	0,51	99,33
	Mediagua o mejora	171	0,48	99,81
	Rancho, choza o ruca	6	0,02	99,83
	Vivienda precaria de materiales reutilizados	46	0,13	99,96
Incluida	Perdidos	15	0,04	100

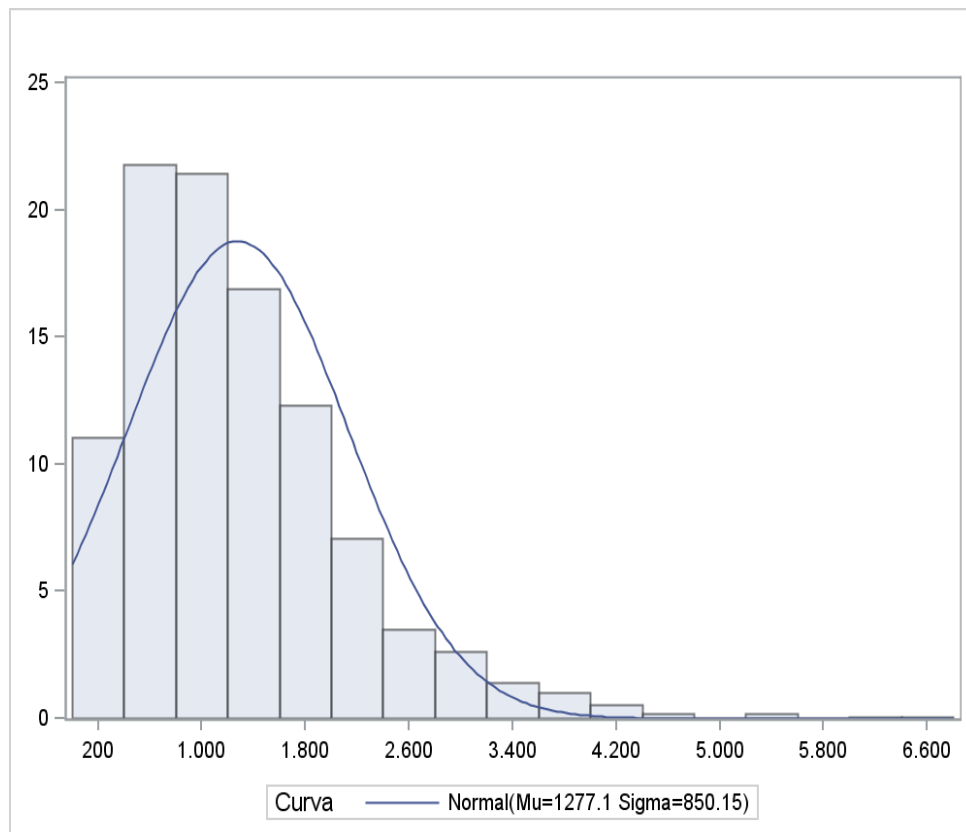
Elaborado por el Instituto Nacional de Estadísticas.

III.2. Variables del Sistema de Información Geográfico a considerar en el PRINCALS

Igualmente, también se emplean las variables provenientes del Sistema de Información Geográfica, y dado que estas variables previamente ya han pasado por un proceso de imputación y depuración, por parte del personal técnico del INE, garantizando consistencia y completitud; no se requiere ningún tratamiento adicional para ser incorporada en el PRINCALS. A continuación, se presentan sus respectivas distribuciones de frecuencia.

La variable “Distancia en metros a comisarías” tiene media de 1.277,10; lo que indica que, ese corresponde a la distancia promedio en metros entre el centroide de la manzana y la comisaría más cercana; por otro lado, también se evidencia que la distribución de esta variable presenta asimetría positiva, y cuenta con presencia de algunos valores atípicos asociados a la naturaleza de la variable, ver gráfico 1.

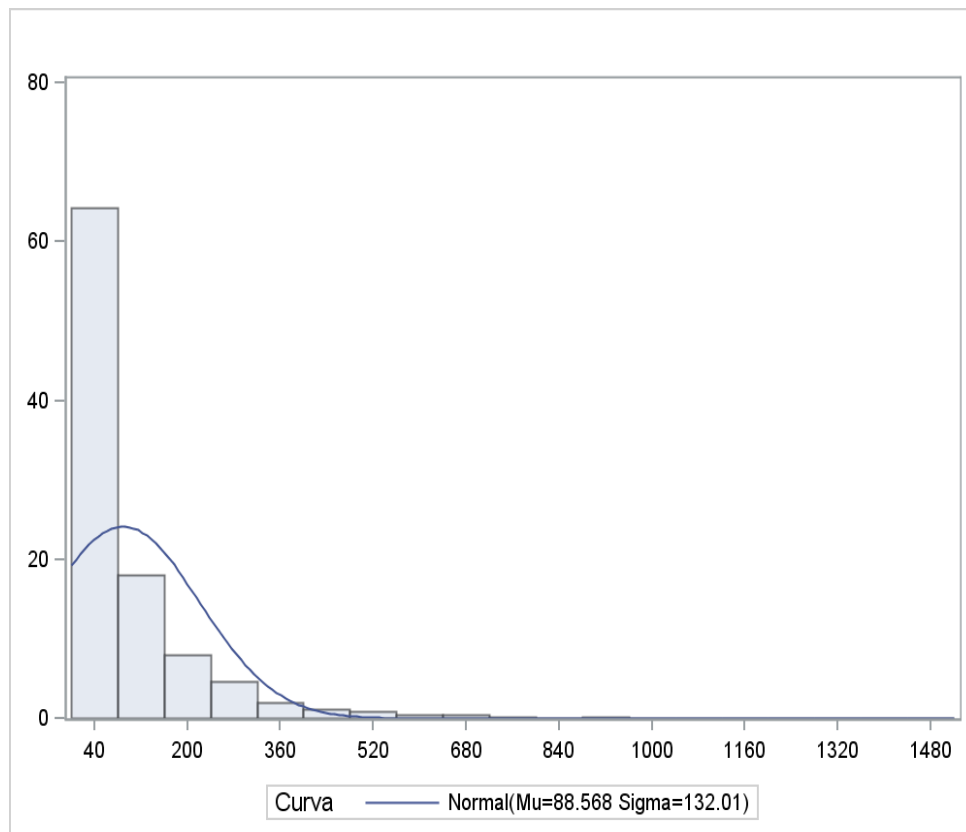
Gráfico 1. Distribución de frecuencia de variable Distancia en metros a comisaría.



Elaborado por el Instituto Nacional de Estadísticas.

En cuanto a la variable “Distancia en metros a áreas verdes”, se obtiene promedio de 88,57, haciendo referencia ésta variable a la distancia en metros mínima entre áreas verdes y el centroide de la manzana. Igual que en el caso anterior, la distribución de la variable no se ajusta a una normal, en cambio presenta asimétrica positiva y con presencia de valores atípicos propios de la naturaleza de la variable, ver gráfico 2.

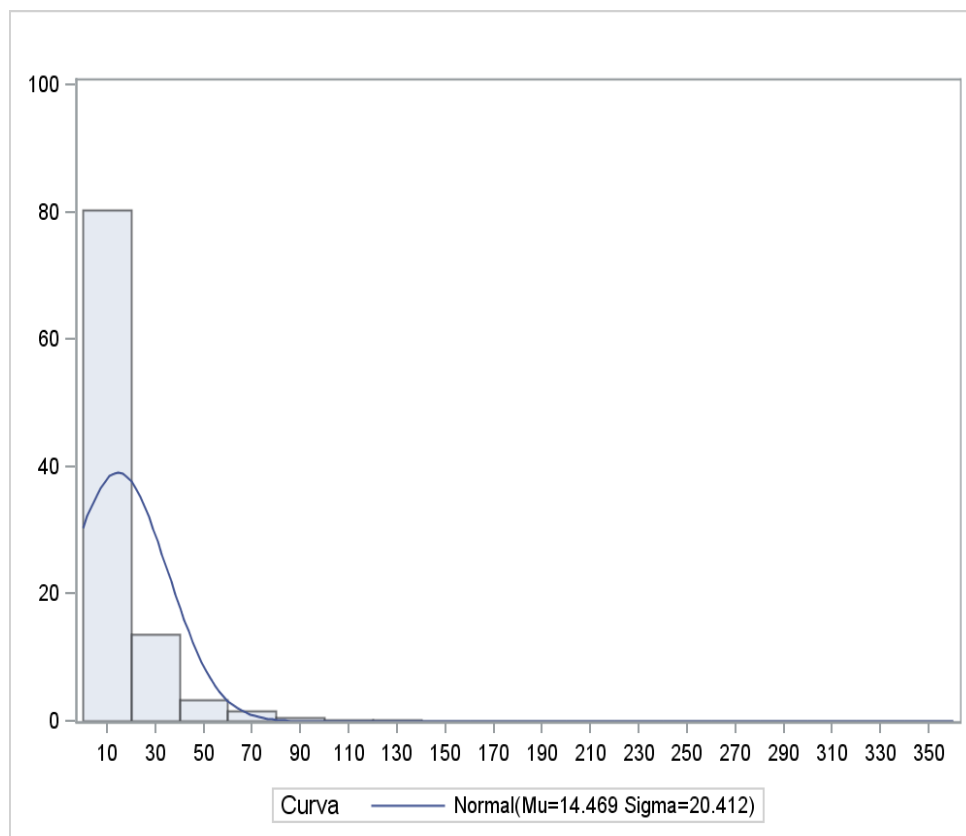
Gráfico 2. Distribución de frecuencia de variable Distancia en metros a áreas verdes.



Elaborado por el Instituto Nacional de Estadísticas.

Respecto al avalúo, esta variable presenta un promedio de aproximadamente 14,47 millones, e igualmente que en los casos previos la distribución no se ajusta a una normal, sino que posee asimetría positiva y con presencia de valores atípicos, también propios de la naturaleza de la variable, ver gráfico 3.

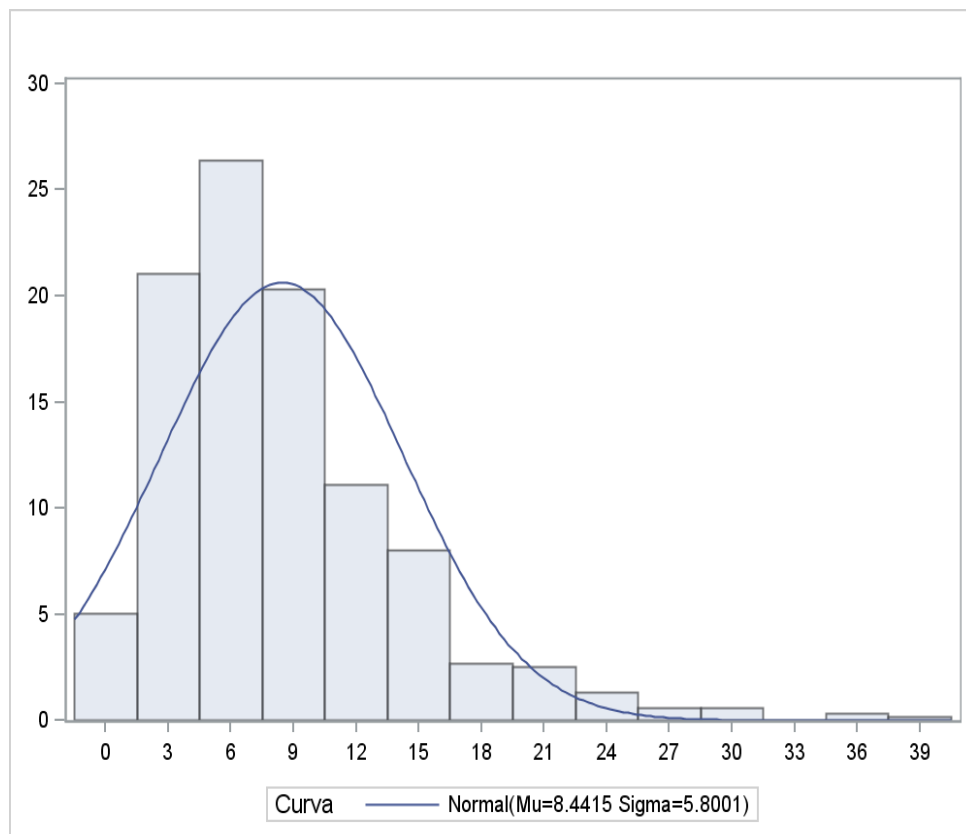
Gráfico 3. Distribución de frecuencia de variable Avalúo por manzana (en millones de pesos).



Elaborado por el Instituto Nacional de Estadísticas.

En cuanto al número de establecimientos educativos por distrito censal, se presenta un valor promedio de 8,44 y valor modal de 6,00; presentando un máximo de 40 instituciones educativas por Distrito. Cabe mencionar que la distribución de esta variable se asemeja un poco más a la normal con respecto a las demás evaluadas.

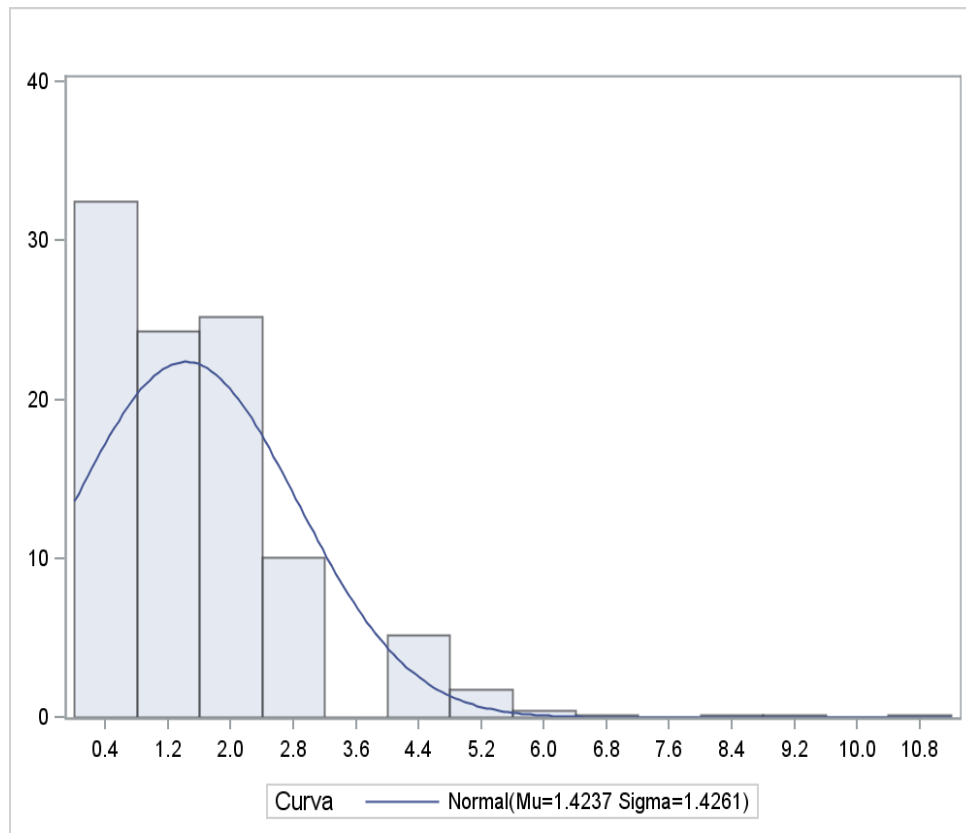
Gráfico 4. Distribución de frecuencia de Número de establecimientos educativos.



Elaborado por el Instituto Nacional de Estadísticas.

Al examinar la distribución de la variable “Número de establecimientos de salud”, se obtiene una media de 1,42 y mediana de 1, presentando un valor máximo de 11 establecimientos de salud por Distrito censal.

Gráfico 5. Distribución de frecuencia de Número de establecimientos de Salud.



Elaborado por el Instituto Nacional de Estadísticas.

IV. Factor de Expansión para post estratificación

Para la aplicación del método de post estratificación PRINCALS, se hace necesario generar un factor de expansión que dé cuenta del total de la población, y no solo de la población objetivo del estudio; ya que, también se emplean variables asociadas a cada uno de los miembros del hogar, lo que permite estimar más acertadamente el nivel socioeconómico.

Para el desarrollo de éste nuevo factor de expansión, se siguen igualmente los mismos lineamientos empleados para construir el factor que permite expandir los indicadores de la Encuesta Nacional de Participación Cultural 2017⁷, salvo que, en los dos últimos ajustes se originan variantes, las cuales se detallan a continuación. En este sentido, se enumeran las ponderaciones y ajustes aplicados para obtener este nuevo factor de expansión:

1. Ponderador de selección de unidades de primera etapa.
2. Ajuste por omisión de conglomerados.
3. Ponderador de selección de unidades de segunda etapa.
4. Ajuste por elegibilidad de las viviendas (elegibilidad desconocida y no elegibilidad).
5. Suavizamiento del ponderador ajustado por elegibilidad.
6. Ajuste por no respuesta.
7. Ponderador de selección de unidades de tercera etapa.
8. Suavizamiento ponderador de selección de personas.
9. Ponderador a nivel de personas para post estratificación.
10. Ajuste para calibrar al total de población según proyecciones por región, sexo y tramo de edad, incluyendo menores de 15 años.

Con respecto a las variantes incorporadas, para este nuevo factor de expansión, a continuación, se detallan los pasos 9 y 10:

IV.1. Ponderador a nivel de personas para post estratificación

Según la enumeración anterior, hasta el punto 8 se tiene un factor asociado al informante, denotado “Suavizamiento ponderador de selección de personas”; sin embargo, para efecto de la post estratificación se requiere tomar en cuenta a todos los integrantes de la vivienda incluyendo también a

⁷ Para mayor detalle consultar Informe metodológico factor de expansión de la Encuesta Nacional de Participación Cultural 2017.

los menores a 15 años. Es por ello que dicho factor es asociado a cada uno de los integrantes de la vivienda y luego dividido entre el total de personas que la conforman.

Con el paso previo, se asigna un factor a cada integrante de la vivienda, haciendo la abstracción, como si en vez de seleccionar al informante, se hubiese seleccionando a todos los integrantes del hogar, incluyendo a los menores de 15 años, con lo que se busca que la estimación del total de personas apunte al total poblacional. Es por ello que en el paso subsiguiente se requiere el ajuste al total poblacional según proyecciones.

Así, el ponderador a nivel de personas para post estratificación queda definido por:

$$w_{rgij}^{pe} = w_{rgij}^{t**} / p_{rgij}$$

Donde:

- r : índice de región
- g : Índice de grupo tamaño
- i : Índice de manzana
- j : Índice de vivienda
- w_{rgij}^{pe} : Ponderador a nivel de personas para post estratificación en la vivienda j de la manzana i de grupo tamaño g dentro de la región r .
- w_{rgij}^{t**} : Ponderador que considera el suavizamiento al ponderador de selección de personas de la vivienda j en la manzana i de grupo de tamaño g en la región r .
- p_{rgij} : Total de personas en la vivienda j de la manzana i de grupo tamaño g dentro de la región r .

IV.2. Ajuste para calibrar al total poblacional por región, sexo y tramo etario

Hasta este punto el factor de expansión utilizado da cuenta del total de personas de la población objetivo, sin embargo, para afectos de la post estratificación se requiere que el factor empleado dé cuenta del total de personas; es por ello que se implementa un nuevo ajuste que permite calibrar las

estimaciones de la muestra a los stocks poblacionales, según proyecciones⁸, para el área urbana por región, sexo y tramo etario.

La tabla 7 muestra las proyecciones de población por sexo según región.

Tabla 7. Proyecciones de Población por sexo y región.

Región	Total	Hombres	Mujeres
Total	15.843.567	7.756.281	8.087.286
I Tarapacá	337.800	169.269	168.531
II de Antofagasta	604.113	310.257	293.856
III de Atacama	266.057	134.062	131.995
IV de Coquimbo	647.988	316.925	331.063
V de Valparaíso	1.728.917	846.235	882.682
VI de O'Higgins	672.628	335.514	337.114
VII del Maule	720.596	351.887	368.709
VIII del Biobío	1.467.572	715.540	752.032
IX de La Araucanía	698.632	334.612	364.020
X de Los Lagos	650.234	325.614	324.620
XI de Aisén	96.818	49.375	47.443
XII de Magallanes y La Antártica	151.236	76.437	74.799
XIII Metropolitana	7.060.881	3.437.169	3.623.712
XIV de Los Ríos	264.472	127.658	136.814
XV de Arica y Parinacota	149.850	68.685	81.165
XVI de Ñuble	325.773	157.042	168.731

Elaborado por el Instituto Nacional de Estadísticas.

Para el cálculo del ajuste, se requiere obtener las estimación del total de personas por región, sexo y tramo etario, mediante el factor de personas para post estratificación w_{rgij}^{pe} . Para luego obtener la razón entre las proyecciones de población y la estimación de población, tal como se expresa a continuación:

$$\hat{R}_{rst} = \frac{P_{rst}}{\hat{P}_{rst}}$$

Donde:

⁸ Proyecciones de población al 30 de octubre de 2017, que consideran cinco tramos etarios para las 15 regiones existentes al momento del levantamiento de la encuesta (14 y menos, 15-29, 30-44, 45-59, 60 y más); y dos tramos etarios para la nueva Región de Ñuble (14 y menos, 15 y más).

P_{rst} : Proyección de población según región r , sexo s y tramo etario t ⁹.
 \hat{P}_{rst} : Estimación de población según región r , sexo s y tramo etario t .

con

$$\hat{P}_{rst} = \sum_s \sum_t \sum_{rgij} w_{rgij}^{pe}$$

Siendo:

s : Índice asociado al sexo
 t : Índice asociado al tramo etario
 w_{rgij}^{pe} : Ponderador nivel de personas para post estratificación.

Luego del cálculo de la razón, el factor de expansión calibrado a las proyecciones de población para la post estratificación se expresa como:

$$w_{rgij}^{cpe} = w_{rgij}^{pe} \cdot \hat{R}_{rst}$$

Se advierte que, el factor de expansión a las proyecciones de población para la post estratificación, w_{rgij}^{cpe} , se construye única y exclusivamente para ser utilizado en la aplicación del método PRINCALS, por lo que se sugiere no utilizarlo para la expansión de indicadores o para cualquier otro tipo de análisis con la información recabada en la encuesta.

⁹ Se consideran cinco tramos etarios para las 15 regiones existentes al momento del levantamiento de la encuesta (14 y menos, 15-29, 30-44, 45-59, 60 y más); y dos tramos etarios para la nueva Región de Ñuble (14 y menos, 15 y más).

V. Configuración de método de PRINCALS

V.1. Conjunto de variables consideradas en escenarios de post estratificación

Para la configuración del procedimiento PRINCALS se consideran, como ya fue explicado, una serie de variables relacionadas con el nivel socioeconómico del hogar y altamente correlacionadas con el ingreso. En tabla 8, se presenta el conjunto de variables¹⁰ que componen algunos de los escenarios de post estratificación, donde se verifica que el escenario que tiene más variables es el 1, con 11 variables, mientras que el que tiene menos es el escenario 5, con 6 variables.

Tabla 8. Variables empleadas en los escenarios de post estratificación.

N°	Variable	Escenario				
		1	2	3	4	5
1	Grupo ocupacional CIUO	✓	✓	✓	✓	✓
2	Categoría ocupacional CISE	✓	✓	✓	✓	✓
3	Nivel Educacional	✓	✓	✓	✓	✓
4	Comuna	✓	✓	✓	✓	✓
5	Tipo de vivienda	✓	✓	✓	✓	✓
6	Número de vehículos	✓	✓	✓	✓	
7	Avalúo promedio por manzana	✓	✓	✓	✓	✓
8	Distancia en metros a comisarias	✓	✓	✓		
9	Distancia en metros a áreas verdes	✓	✓			
10	Número establecimientos educacionales	✓				
11	Número establecimientos salud	✓				

Elaborado por el Instituto Nacional de Estadísticas.

V.2. Bondad del ajuste de los escenarios evaluados de post estratificación

Entre los criterios considerados para conformar los escenarios de post estratificación se encuentran: capacidad explicativa de la varianza, consistencia de los signos y representatividad; todo esto para cada variable en la primera dimensión. En cuanto al segundo criterio, se busca garantizar la consistencia entre nivel socioeconómico de los hogares y los signos de los coeficientes asociados a las categorías de las variables, es decir, la correspondencia entre nivel socioeconómico alto con categorías asociadas a este

¹⁰ Se consideraron otras variables, sin embargo, a efectos de resumen se presentan solo 11 de ellas.

nivel, y viceversa. Mientras que, respecto a la representatividad, se cuida que cada variable tenga un aporte significativo en cuanto a la conformación de la primera dimensión.

Igualmente, también se considera como criterio adicional el indicador Alfa de Cronbach, que permite medir la fiabilidad de las variables empleadas para medir un mismo constructo o dimensión teórica, en este caso el nivel socioeconómico de los hogares. A continuación, se presentan algunos estadísticos de bondad de ajuste de los diferentes escenarios evaluados.

Tabla 9. Estadísticos de bondad de escenarios de post estratificación.

Escenario	Alfa de Cronbach	Varianza contabilizada (autovalor)	Varianza total explicada (%)
1	0,76	3,26	29,67
2	0,79	3,33	37,05
3	0,80	3,35	41,88
4	0,80	3,16	45,13
5	0,79	2,89	48,20

Elaborado por el Instituto Nacional de Estadísticas.

V.3. Selección de escenario definitivo de post estratificación

Una vez evaluados los escenarios de post estratificación, bajo los criterios ya mencionados, se opta por elegir el escenario 5, ya que además de ser el más parsimonioso, es el que presenta mayor coherencia entre las variables empleadas y la clasificación socioeconómica que genera. También se puede destacar que logra explicar 48,2% de la variabilidad total (porcentaje que logra acumular el autovalor asociado a la primera dimensión evaluada) y que el índice Alfa de Cronbach se ubica en 0,79, lo que indica que las variables consideradas miden un mismo constructo y están altamente correlacionadas, caracterizando este escenario en cuanto a fiabilidad como bueno, ver tabla 9.

En consecuencia, el escenario definitivo de post estratificación considera seis variables, a saber:

- Grupo ocupacional CIUO modificado
- Categoría ocupacional CISE modificado
- Nivel educacional modificado

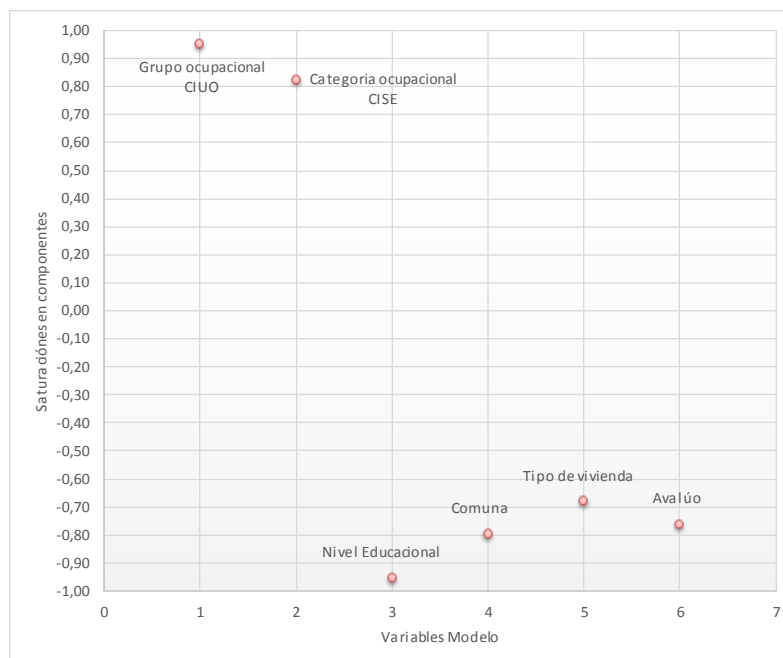
- Comuna
- Tipo de vivienda modificado
- Avalúo

V.4. Importancia de las variables en el escenario de post estratificación

En cuanto a nivel de importancia asociado a cada variable que participa en el escenario, según los valores que toman las saturaciones en componentes, se puede observar que todas las variables explican significativamente a la primera dimensión, separando las variables de Categoría ocupacional CISE y Grupo ocupacional CIUO, con valores positivos, de las demás variables, con valores negativos (ver gráfico 6).

Al considerar el valor absoluto de las saturaciones en componentes se puede precisar el siguiente orden de importancia (de menor a mayor): Tipo de vivienda (0,68), Avalúo (0,76), Comuna (0,80), Categoría ocupacional CISE (0,82), y, Nivel educativo y Grupo ocupacional CIUO, como las variables más importantes, ambas con igual valor de 0,95.

Gráfico 6 . Saturaciones en componentes.



Elaborado por el Instituto Nacional de Estadísticas.

VI. Utilización de los factores de expansión

Al Consejo Nacional de la Cultura y las Artes, se enviaron dos bases de datos en SPSS, denominadas “Base Entrevistados Postestratificación.sav” y “Base Hogar Postestratificación.sav”. La primera, está a nivel de registro de hogares o entrevistados y es la base que debiera utilizarse para realizar las expansiones de las variables de la Encuesta Nacional de Participación Cultural 2017. La segunda base está a nivel de personas y es sólo de uso interno, utilizada para el cálculo de los factores **de persona y hogar** en la aplicación del método PRINCALS, permitiendo obtener clasificación socioeconómica de los hogares usando características de las personas.

En la siguiente tabla se muestra el nombre de la base, el nivel de los registros y el factor en ella.

Tabla 10 .Descripción de nivel de registro y factor de expansión de bases de datos.

Nombre de base de datos entregadas	Nivel de registro	Factor de expansión
Base Entrevistados Postestratificación.sav	Informante	W_hij_nr_a
Base Hogar Postestratificación.sav	Personas	Factor_Hogar_Postestratificación

Elaborado por el Instituto Nacional de Estadísticas.

Para la correcta utilización del factor de expansión se debe considerar el tipo de variable que se va a expandir (referente a la persona u hogar) así como el nivel de registro de la base (a nivel de persona u hogar).

VI.1. Utilización de los factores de expansión

En la base a nivel de registro de hogares o informante “Base Entrevistados Postestratificación.sav”, el factor de expansión “W_hij_nr_a” es un expansor calibrado por sexo y tramo de edad¹¹ que expande a la población objetivo, es decir, a la población de personas mayores (de 15 o más años).

Ponderando la base (a nivel del entrevistado) por este factor, se puede constatar que éste expande a la población objetivo, es decir a las personas mayores, según las proyecciones de población, ver Tabla 11.

Tabla 11. Distribución poblacional según quintil socioeconómico.

Quintil	Frecuencia	Porcentaje
Total	12.571.939	100
1	2.899.174	23,1
2	2.671.550	21,3
3	2.582.523	20,5
4	2.475.636	19,7
5	1.943.056	15,5

Elaborado por el Instituto Nacional de Estadísticas.

También se puede observar la distribución de estas personas mayores¹² por nivel educacional y quintil socioeconómico, en términos de frecuencias absolutas, dadas en la tabla 12. Al examinar las casillas que conforman el cruce de las variables, se puede precisar que se dificulta la confiabilidad¹³ de los resultados en cada casilla (por ejemplo, quintil socioeconómico 2, 4 y 5 para el nivel educacional 3 “Pre kínder/Kinder”).

¹¹ Se calibra según proyecciones de población al 30 de octubre de 2017, que consideran cuatro tramos etarios para las 15 regiones existentes al momento del levantamiento de la encuesta (15-29, 30-44, 45-59, 60 y más); y un tramo etario para la nueva Región de Ñuble (15 años y más).

¹² Es importante destacar que el informante, representa a todas las personas mayores del hogar y se asume que, si es mujer, representa a todas las mujeres mayores, como también, si es hombre, representa a todos los hombres mayores.

¹³ Tener presente que estos son datos expandidos y los de menor cuantía pueden corresponder a 1 o a lo más 2 casos muestrales.

Tabla 12. Distribución de personas mayores según nivel educacional y quintil.

Nivel educacional	Quintil					
	Total	1	2	3	4	5
Total	12.571.934	2.899.172	2.671.550	2.582.523	2.475.635	1.943.054
1 Nunca asistió	133.582	56.567	54.712	18.476	3.359	468
3 Pre kinder/Kinder	1.715	752	0	963	0	0
4 Educación especial (diferencial)	23.104	11.176	9.775	334	1.541	278
5 Básica	1.553.446	679.716	466.587	265.446	126.422	15.275
6 Primaria/Preparatoria (S. Antiguo)	473.547	153.408	153.505	107.819	45.257	13.558
7 Media Científico Humanista	3.706.507	979.606	977.744	925.096	650.405	173.656
8 Humanidades (S. Antiguo)	544.194	45.972	97.037	186.051	157.176	57.958
9 Media Técnico-Profesional	1.561.615	511.021	368.785	328.224	259.267	94.318
10 Técnica (S. Antiguo)	186.955	24.546	33.539	59.142	59.311	10.417
11 Técnica Nivel Superior	1.173.010	237.955	200.173	236.131	335.486	163.265
12 Profesional	2.874.395	196.471	301.543	442.747	787.207	1.146.427
13 Post-título	126.520	168	3.789	3.954	20.213	98.396
14 Magíster	170.888	1.814	4.361	4.394	20.740	139.579
15 Doctorado	42.456	0	0	3.746	9.251	29.459

Elaborado por el Instituto Nacional de Estadísticas.

Es importante destacar que no es conveniente hacer análisis descriptivo de frecuencias considerando variables con muchas categorías, debido a que la información en las casillas se hace menos precisa a medida que se consideran más desagregaciones, disminuyendo la confiabilidad aún más, cuando se realizan cruces de variables.

Adicionalmente, se puede corroborar que, al comparar los totales entre las tablas 11 y 12 no coinciden; esto es debido a que la suma de las frecuencias absolutas marginales en una tabla cruzada no coincide exactamente con las sumas en cada variable en tablas de frecuencias simples, cuando el factor de expansión no es un número entero, debido a que el software realiza una aproximación de decimales internamente para la construcción de estas tablas de frecuencias. Para aminorar este problema, se recomienda utilizar estructuras porcentuales para los análisis y colocar nota al pie de página explicando que si la suma de las marginales no coincide exactamente al total (100%), se debe exclusivamente a ajustes de redondeo de decimales.

A partir de los resultados de tabla 12, al colapsar algunas categorías de nivel educacional, son mostrados en la tabla 13, 14 y 15, respectivamente.

Tabla 13. Distribución de personas mayores según nivel educacional y quintil.

Nivel educacional	Quintil					
	Total	1	2	3	4	5
Total	12.571.941	2.899.173	2.671.551	2.582.524	2.475.636	1.943.057
1 Menos de enseñanza básica	158.402	68.496	64.487	19.773	4.900	746
2 Enseñanza básica	2.026.994	833.124	620.092	373.265	171.679	28.834
3 Enseñanza media	5.999.275	1.561.145	1.477.106	1.498.514	1.126.160	336.350
4 Técnico superior o profesional	4.047.407	434.426	501.717	678.878	1.122.693	1.309.693
5 Postítulo, magíster o doctorado	339.863	1.982	8.149	12.094	50.204	267.434

Elaborado por el Instituto Nacional de Estadísticas.

Igualmente, al comparar las tablas 12 y 13 se puede observar que el total de personas mayores obtenidos no coinciden (12.571.934 versus 12.571.941) tal como se explicó anteriormente, debido al redondeo o aproximación de decimales.

Tabla 14. Distribución porcentual de personas mayores según nivel educacional y quintil.

Nivel educacional	Quintil					
	Total	1	2	3	4	5
Total	100	23,1	21,3	20,5	19,7	15,5
1 Menos de enseñanza básica	100	43,2	40,7	12,5	3,1	0,5
2 Enseñanza básica	100	41,1	30,6	18,4	8,5	1,4
3 Enseñanza media	100	26	24,6	25	18,8	5,6
4 Técnico superior o profesional	100	10,7	12,4	16,8	27,7	32,4
5 Postítulo, magíster o doctorado	100	0,6	2,4	3,6	14,8	78,7

Elaborado por el Instituto Nacional de Estadísticas.

Tabla 15 Distribución porcentual de personas mayores según nivel educacional y quintil.

Nivel educacional	Quintil					
	Total	1	2	3	4	5
Total	100	100	100	100	100	100
1 Menos de enseñanza básica	1,3	2,4	2,4	0,8	0,2	0
2 Enseñanza básica	16,1	28,7	23,2	14,5	6,9	1,5
3 Enseñanza media	47,7	53,8	55,3	58	45,5	17,3
4 Técnico superior o profesional	32,2	15	18,8	26,3	45,3	67,4
5 Postítulo, magíster o doctorado	2,7	0,1	0,3	0,5	2	13,8

Elaborado por el Instituto Nacional de Estadísticas.

En términos porcentuales, si se toma en forma literal estos porcentajes expresados con un decimal, se podría dar el caso que, la suma no coincidiera exactamente con 100%, lo cual se debe al redondeo, ya que, si se expresara haciendo uso de todos los decimales, la suma coincidiría con 100%.

VI.2. Base a nivel de registro de personas

En la base a nivel de personas “Base Hogar Postestratificación.sav”, desde la cual se construyó el factor PRINCALS, el ponderador “Factor_Hogar_Postestratificación” es el ponderador del hogar y da cuenta de los hogares en la población que representa un hogar en la muestra. Este ponderador aplica a todos los integrantes del hogar, ya sea mayores o menores de edad y se construyó exclusivamente para la aplicación del método PRINCALS en la ponderación de los hogares para construir los quintiles de nivel socioeconómico.

Si se pondera la base de datos por este factor¹⁴ y se construye una tabla de frecuencias cruzada de la variable nivel educacional colapsado, con Quintil socioeconómico, se llega al resultado presentado en la Tabla 15 (15.843.567 personas).

¹⁴ Al multiplicar este factor por el número de personas de la vivienda, se construye un factor de personas que no presenta calibración por sexo y tramo de edad, a diferencia del utilizado en el método PRINCALS en la clasificación socioeconómica de las personas. Al sumarlo sobre todas las personas de todos los hogares, se llega al total de personas, mayores más menores. Se recalca que es de uso interno, exclusivo para la post estratificación de los hogares.

Tabla 16. Distribución de total personas según nivel educacional y quintil socioeconómico.

Nivel educacional	Quintil					
	Total	1	2	3	4	5
Total	15.843.567	4.146.553	3.346.506	3.104.373	2.909.216	2.336.919
1 Menos de enseñanza básica	1.208.059	671.842	253.235	113.027	68.981	100.974
2 Enseñanza básica	3.531.872	1.592.167	999.903	535.531	282.456	121.815
3 Enseñanza media	6.443.485	1.510.005	1.579.020	1.692.691	1.259.486	402.283
4 Técnico de nivel superior o profesional	4.331.348	370.576	509.574	753.700	1.255.680	1.441.818
5 Postítulo, magíster o doctorado	328.803	1.963	4.774	9.424	42.613	270.029

Elaborado por el Instituto Nacional de Estadísticas.

Nótese que el total de personas expandidas de la tabla 16 es distinto a la presentada en la Tabla 13; mientras en esta última se llega al total de personas (15.843.567, mayores más menores) en la primera se llega al total de personas mayores (12.571.941).

Cabe mencionar que para lograr la expansión de la tabla 16, se construye un factor adicional de persona (ver nota 14), multiplicando el “Factor_Hogar_Postestratificación” por el número de personas total de la vivienda (“numper”) que es de uso exclusivo para la post estratificación de los hogares; no corresponde la calibración por sexo ni tramo de edad (a todos los integrantes del hogar le corresponde el mismo ponderador del hogar).

También se observa la distribución distinta de la población de la tabla 16 con respecto a la tabla 13, a través de los quintiles. Por ejemplo, la distribución de la categoría 5: “Postítulo, magíster o doctorado”, que sólo involucra a las personas mayores, es distinta en ambas tablas, debido a que un factor está calibrado a las personas mayores por sexo y tramo etario y el otro no lo está, porque es un ponderador de hogar, ver tabla 17.

Tabla 17. Comparación de distribución de personas mayores según quintil

Nombre Tabla	Nivel educacional	Quintil					
		Total	1	2	3	4	5
Tabla 13	5 Postítulo, magíster o doctorado	339.863	1.982	8.149	12.094	50.204	267.434
Tabla 16	5 Postítulo, magíster o doctorado	328.803	1.963	4.774	9.424	42.613	270.029

Elaborado por el Instituto Nacional de Estadísticas.

Si se filtra esta base de registro a nivel de personas por el informante, se obtiene una base a nivel de registro de hogares y, por tanto, al expandir por el factor “Factor_Hogar_Postestratificación”, se puede observar la distribución de hogares según quintil socioeconómico, tal como se muestra en la tabla 18. Este factor de hogar tiene sentido en la expansión de variables relacionadas con características del hogar o vivienda asociada al hogar, no así sobre las variables relacionadas con las características de las personas.

Tabla 18 Distribución de hogares expandidos según quintil socioeconómico.

Quintil	Frecuencia	Porcentaje
Total	5.542.129	100
1	1.108.243	20
2	1.108.541	20
3	1.107.053	20
4	1.109.568	20
5	1.108.724	20

Elaborado por el Instituto Nacional de Estadísticas.

Si en la base a nivel de registro del entrevistado se adosara este factor (Factor_Hogar_Postestratificación), se tendría una base con ambos factores (el anterior y W_hij_nr_a, éste último que corresponde al factor de personas mayores) en donde se utilizaría el factor de personas mayores para expandir las variables relacionadas con las personas, mientras tanto, se podría utilizar el factor de hogar para expandir las variables relacionadas con el hogar.

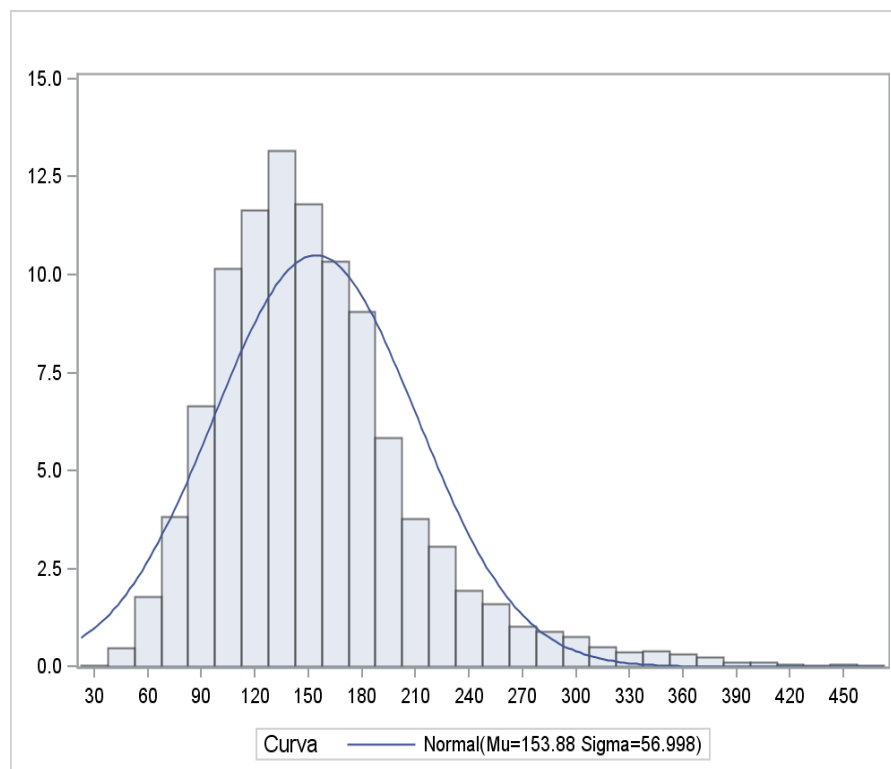
VII.Resultados

Tras aplicación del método de PRINCALS a la Encuesta Nacional de Participación Cultural 2017, se logran obtener puntajes asociados a cada persona, que son después promediados a nivel de cada hogar, para con ello realizar los respectivos cálculos de quintiles socioeconómicos.

En cuanto al factor de expansión, se emplea el obtenido en el procedimiento desarrollado en el punto IV para la aplicación del método PRINCALS, y luego para el cálculo de los quintiles, ese mismo factor de expansión es agregado a nivel de cada hogar y dividido entre el número de personas que los componen.

En el gráfico 7, se muestra la distribución de los puntajes asignado a los hogares, donde se verifica que los puntajes se distribuyen aproximadamente normal con media 153,88 y desviación estándar 56,99, presentando una pequeña asimetría positiva, lo que significa que hay levemente más valores a la izquierda de la media que puntajes a la derecha de esta.

Gráfico 7. Distribución de frecuencia de los puntajes PRINCALS asociados a los hogares.



La distribución por quintil según las variables empleadas en el método PRINCALS se muestran a continuación.

Al examinar los datos de la variable Grupo ocupacional CIUO modificado por quintil, destaca que las categorías que presentan un orden creciente al pasar del quintil 1 al 5 son: “Miembros del poder ejecutivo y de los cuerpos legislativo y personal” y “Profesionales científicos e intelectuales”; siendo la categoría de “Profesionales científicos e intelectuales” la que acumula más personas en el quintil 5, con 54,5%. Por otro lado, las categorías que muestra orden decreciente son: “Agricultores y trabajadores calificados agropecuarios y pesqueros”, “Operadores de instalaciones máquinas y montadores” y “Trabajadores no calificados”; siendo las categorías “Fuerzas Armadas” y “agricultores y trabajadores calificados agropecuarios y pesqueros”, las que presentan mayor porcentaje de personas en el quintil 1, con 44,4% y 41,3% respectivamente, ver tabla 19.

Tabla 19. Distribución porcentual de personas por quintil según grupo ocupacional CIUO modificado.

Grupo ocupacional CIUO modificado	Quintil					
	Total	1	2	3	4	5
Total	100	26,2	21,1	19,6	18,4	14,7
Miembros del poder ejecutivo y de los cuerpos legislativo y personal	100	15,3	7,9	14,3	21,4	41,2
Profesionales científicos e intelectuales	100	5,2	8,4	9,8	22,1	54,5
Técnicos y profesionales de nivel medio	100	18,5	18,1	16,3	26,7	20,5
Empleados de oficina	100	25,6	23,4	21,7	20,2	9,1
Trabajadores de los servicios y vendedores de comercios y mercados	100	30,0	17,6	21,6	22,1	8,7
Agricultores y trabajadores calificados agropecuarios y pesqueros	100	41,3	38,0	12,6	6,9	1,1
Oficiales, operarios y artesanos de artes mecánicas y de otros oficios	100	33,0	24,6	20,5	18,7	3,2
Operadores de instalaciones y máquinas y montadores	100	37,7	24,3	19,9	14,5	3,6
Trabajadores no calificados	100	35,1	27,3	20,3	13,5	3,7
Fuerzas armadas	100	44,4	23,7	24,8	7,1	0,0
Ocupados sin declaración de actividad	100	19,7	17,6	17,9	30,5	14,4
Desocupados	100	23,5	25,7	19,7	16,5	14,7
Inactivos	100	28,8	23,8	22,1	15,0	10,3

Elaborado por el Instituto Nacional de Estadísticas.

Al examinar los niveles educativos asociados a los informantes clasificados por quintil, se detecta un patrón decreciente, al pasar del quintil 1 al 5 en las categorías: “Sin enseñanza básica” y “Enseñanza básica”; en este mismo orden de idea, las categorías que presentan patrón creciente son: “Técnico de nivel superior o profesional” y “Postítulo, magíster o doctorado”. También se puede destacar que las categorías “Sin enseñanza básica” y “Enseñanza Básica”, son las que presentan mayor concentración de personas en el quintil 1, ambas con 43,4%; mientras que la categoría “Postítulo, magíster o doctorado” es la que presenta mayor concentración de personas para el quintil 5, con 77%, ver tabla 20.

Tabla 20. Distribución porcentual de personas por quintil según nivel educativo modificado.

Variable 3: Nivel educativo	Quintil					
	Total	1	2	3	4	5
Total	100	26,2	21,1	19,6	18,4	14,7
Sin enseñanza básica	100	43,4	40,4	12,9	2,8	0,4
Enseñanza básica	100	43,4	29,1	17,9	8,3	1,3
Enseñanza media	100	30,1	23,9	23,3	17,4	5,3
Técnico de nivel superior o profesional	100	12,9	13,6	16,4	25,8	31,3
Postítulo, magíster o doctorado	100	0,8	2,9	4,1	15,2	77,0

Elaborado por el Instituto Nacional de Estadísticas.

En el anexo 1, se muestra la distribución de los hogares por quintil según las comunas consideradas en este estudio, y aunque se advierte que a nivel comunal no se garantiza representatividad¹⁵, sin embargo se puede destacar lo siguiente:

- Las comunas que concentran más de 70% de los hogares en el quintil 5 son: Santiago, La Reina, Las Condes, Lo Barnechea, Ñuñoa, Providencia, Vitacura y San Miguel.
- Por el contrario, las comunas que concentran más de 70% de los hogares en el quintil 1 son: Chañaral, La Ligua, Nogales, Lota, Quellón, Lo Espejo y Mariquina.

¹⁵ Según informe metodológico del diseño muestral la IV Encuesta Nacional de Participación Cultural 2017, se garantiza representatividad para el área urbana a nivel nacional y para para cada una de las 16 regiones del país, que pertenecen a la población objetivo (personas de 15 años y más residentes en zonas urbanas).

Al examinar la distribución de los hogares por quintil según tipo de vivienda, se aprecia que las categorías, “Rancho, choza o ruca” y “Vivienda precaria de materiales reutilizados”, concentran 100% de los hogares en el quintil 1; mientras que la categoría que presenta mayor concentración en el quintil 5 corresponde a “Departamento en edificio con ascensor”, con 95,8%, ver tabla 21.

Tabla 21. Distribución porcentual de hogares por quintil según tipo de vivienda.

Tipo de vivienda	Quintil					
	Total	1	2	3	4	5
Total	100	20,0	20,0	20,0	20,0	20,0
Casa aislada (no pareada)	100	17,5	20,6	23,5	25,9	12,5
Casa pareada por un lado	100	26,9	25,8	22,6	18,8	5,9
Casa pareada por ambos lados	100	22,4	21,3	23,1	25,2	8,1
Departamento en edificio con ascensor	100	0,0	0,1	0,8	3,2	95,8
Departamento en edificio sin ascensor	100	17,7	18,5	18,3	20,0	25,5
Pieza en casa contigua o conventillo	100	15,1	13,6	21,0	38,9	11,3
Mediagua o mejora	100	42,6	17,6	28,7	11,1	0,0
Rancho, choza o ruca	100	100,0	0,0	0,0	0,0	0,0
Vivienda precaria de materiales reutilizados	100	100,0	0,0	0,0	0,0	0,0

Elaborado por el Instituto Nacional de Estadísticas.

En la tabla 22, se aprecia la distribución de los hogares por quintil según tramo de avalúo, donde destaca que los dos últimos tramos de avalúo más altos concentran el mayor número de hogares en el quintil 5, con 74,4% y 82,3%, respectivamente; mientras que el primer tramo (Menor a \$8.547.300) muestra un patrón decreciente al pasar del quintil 1 a 5, pasando de 41,6% a 0,7%.

Tabla 22. Distribución porcentual de hogares por quintil según tramo de avalúo promedio.

Tramo de avalúo promedio (pesos)	Quintil					
	Total	1	2	3	4	5
Total	100	20,1	20,1	20,0	20,0	19,9
Menor a 8.547.300	100	41,6	30,0	20,1	7,5	0,7
8.561.855 - 17.515.513	100	12,4	23,9	30,8	28,2	4,8
17.537.953 - 26.945.933	100	3,5	11,5	20,6	49,5	14,9
27.208.893 - 38.674.864	100	0,6	2,1	6,6	16,2	74,4
Mayor a 38.935.681	100	0,0	0,7	2,5	14,5	82,3

Elaborado por el Instituto Nacional de Estadísticas.

VIII. Bibliografía

B. Visauta Vinacua: Análisis estadístico con SPSS para Windows. Análisis Multivariante vol.2 (1998). Editorial McGraw-Hill / Interamericana de España.

CEPAL (2013): La medición multidimensional de la pobreza. Documento de la Duodécima reunión del Comité Ejecutivo de la Conferencia Estadística de las Américas de la Comisión Económica para América Latina y el Caribe. División de estadísticas, CEPAL

De Leeuw, J., Young, F. W., and Takane, Y. 1976. Additive structure in qualitative data: an alternating least squares method with optimal scaling features. *Psychometrika*, 31: 33–42.

De Leeuw, J., and Van Rijckevorsel, J. 1980. HOMALS and PRINCALS—Some generalizations of principal components analysis. In: *Data Analysis and Informatics*, E. Diday et al, eds. Amsterdam: North-Holland.

Gifi, A. 1985. PRINCALS. Leiden: Department of Data Theory, Internal Report UG-85–02.

INE (2004): “Metodología para la caracterización socioeconómica”. En *Cómo ha cambiado la vida de los chilenos. Análisis comparativo de las condiciones de vida de los hogares con menor bienestar socioeconómico (Censos 1992-2002)*. Instituto Nacional de Estadísticas, Santiago de Chile. (pp. 11-28).

INE (2011): *Estratificación Socioeconómica en Encuestas de Hogares*. Instituto Nacional de Estadísticas, Departamento de Estudios Sociales. Santiago de Chile, diciembre 2011.

León A. y Martínez J.: *La estratificación social chilena hacia finales del siglo XX*. Serie 52, CEPAL-ECLAC. Santiago de Chile, Agosto 2001.

McFadden, Daniel.: “Modeling the choice of residential location”, *Spatial Interaction Theory and Planning Models*. New York. North-Holland, 1978. pp. 75-96 (v.3).

Mendez, Roberto: "Segmentación socioeconómica: el Nuevo perfil del consumidor chileno", Mimeo, Junio 1999.

Meulman, J. J., A. J. Van der Kooij, and W. J. Heiser. 2004. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data. In: Handbook of Quantitative Methodology for the Social Sciences, D. Kaplan, eds. Thousand Oaks, Calif.: Sage Publications, Inc.

Ministerio de Desarrollo Social (2015a): Nueva Metodología de Medición de la Pobreza por Ingresos y Multidimensional. Gobierno de Chile, Ministerio de Desarrollo Social: Serie Documentos Metodológicos Nº28, División Observatorio Social. Versión revisada: 26 de Enero de 2015.

Navarro Céspedes, J.M. (2008) Análisis de Componentes Principales y Análisis de Regresión para datos categóricos. Aplicación en HTA. Tesis de Maestría, Universidad Central de las Villas, Santa Clara, Cuba.

Rodríguez Vignoli, Jorge: Segregación Residencial Socioeconómica: ¿Qué es?, ¿Cómo se mide?, ¿Qué está pasando?, ¿Importa? . Serie 16, CEPAL-ECLAC. Santiago de Chile, agosto 2001.

Sanhueza, C., & Larrañaga, O. (2007). Residential segregation effects on poor's opportunities in Chile. Documentos de Trabajo, 259.

Young, F. W., Y. Takane, and J. De Leeuw. 1978. The principal components of mixed measurement level multivariate data: An alternating least squares method with optimal scaling features. Psychometrika, 43, 279-281.

IX. Anexo

Anexo 1. Distribución porcentual de hogares por quintil según comuna.

Comuna	Quintil					
	Total	1	2	3	4	5
Iquique	100	5,0	10,9	18,5	39,4	26,2
Alto Hospicio	100	47,1	36,1	14,2	2,3	0,3
Pozo Almonte	100	63,0	19,8	17,2	0,0	0,0
Antofagasta	100	9,4	11,9	25,2	42,0	11,5
Calama	100	13,5	12,4	13,5	45,0	15,6
Tocopilla	100	41,2	52,7	4,3	1,8	0,0
Copiapó	100	18,6	19,1	24,3	26,4	11,7
Chañaral	100	80,7	17,5	1,9	0,0	0,0
Vallenar	100	31,4	34,3	18,7	14,0	1,6
La Serena	100	22,4	22,2	29,8	21,6	4,0
Coquimbo	100	36,5	26,5	20,8	14,0	2,1
Illapel	100	82,2	17,8	0,0	0,0	0,0
Ovalle	100	43,0	37,7	13,4	5,0	1,0
Valparaíso	100	7,0	18,8	25,3	35,5	13,3
Concón	100	3,8	1,4	23,4	53,4	18,0
Viña del Mar	100	2,4	6,8	19,1	38,0	33,7
Los Andes	100	17,3	28,9	25,4	24,2	4,3
La Ligua	100	83,3	11,8	4,9	0,0	0,0
Quillota	100	22,0	31,5	21,6	23,3	1,6
Calera	100	31,1	43,4	20,6	4,9	0,0
Nogales	100	79,4	20,6	0,0	0,0	0,0
San Antonio	100	28,4	30,1	24,4	14,3	2,9
Cartagena	100	44,1	34,4	16,9	4,6	0,0
San Felipe	100	16,3	25,5	31,0	20,3	6,9
Quilpué	100	33,4	20,6	31,8	12,3	1,9
Limache	100	58,5	25,6	11,7	4,2	0,0
Villa Alemana	100	13,9	21,6	36,1	24,5	3,9
Rancagua	100	10,3	16,2	28,4	38,2	6,9
Doñihue	100	9,1	50,7	3,9	16,7	19,7
Graneros	100	28,2	24,7	39,4	7,7	0,0
Machalí	100	0,0	9,1	26,0	47,7	17,1
Mostazal	100	59,3	40,7	0,0	0,0	0,0
Rengo	100	53,0	29,8	14,8	2,3	0,0
San Vicente	100	45,8	38,6	0,0	15,6	0,0

Anexo 1. Distribución porcentual de hogares por quintil según comuna (continuación).

Comuna	Quintil					
	Total	1	2	3	4	5
Pichilemu	100	55,2	32,3	8,7	0,0	3,9
San Fernando	100	38,1	24,9	26,3	9,2	1,6
Chimbarongo	100	45,4	32,4	22,2	0,0	0,0
Santa Cruz	100	6,3	21,9	36,6	31,2	4,0
Talca	100	22,4	24,2	24,4	21,1	7,9
Constitución	100	18,4	30,7	27,6	22,1	1,3
Cauquenes	100	30,2	10,7	31,1	25,4	2,6
Curicó	100	38,7	24,8	21,5	10,0	4,9
Molina	100	31,2	38,4	22,5	6,1	1,8
Teno	100	50,4	20,7	28,9	0,0	0,0
Linares	100	24,8	29,8	24,3	20,0	1,2
Parral	100	36,3	38,4	22,3	3,0	0,0
San Javier	100	15,4	32,2	31,4	21,1	0,0
Concepción	100	2,2	8,6	14,7	40,5	34,1
Coronel	100	39,2	26,1	19,7	14,9	0,0
Chiguayante	100	31,0	35,2	18,9	11,7	3,2
Lota	100	71,9	14,7	13,4	0,0	0,0
Penco	100	25,3	24,7	29,9	20,1	0,0
San Pedro de la Paz	100	18,7	19,7	25,8	24,5	11,3
Talcahuano	100	26,0	25,6	33,9	13,4	1,2
Tomé	100	44,8	36,0	17,1	2,0	0,0
Hualpén	100	22,4	24,4	26,5	23,6	3,1
Lebu	100	61,9	13,6	24,5	0,0	0,0
Arauco	100	0,0	0,0	0,0	100,0	0,0
Curanilahue	100	21,6	36,2	26,3	12,4	3,4
Los Angeles	100	35,4	30,9	15,7	14,8	3,3
Cabrero	100	45,9	23,4	23,8	6,9	0,0
Mulchén	100	55,9	31,2	8,2	4,6	0,0
Chillán	100	19,3	31,4	24,3	20,3	4,7
Chillán Viejo	100	36,3	39,1	16,1	8,4	0,0
San Carlos	100	53,4	35,5	11,1	0,0	0,0
Yungay	100	68,8	26,8	4,4	0,0	0,0
Temuco	100	9,0	19,2	24,5	29,0	18,3
Freire	100	55,8	30,2	0,0	14,0	0,0

Anexo 1. Distribución porcentual de hogares por quintil según comuna (continuación).

Comuna	Quintil					
	Total	1	2	3	4	5
Lautaro	100	44,8	34,3	14,0	7,0	0,0
Loncoche	100	42,2	40,0	17,9	0,0	0,0
Nueva Imperial	100	62,7	33,7	0,0	3,5	0,0
Padre las Casas	100	42,3	29,2	11,9	11,8	4,8
Pitrufquén	100	72,7	25,7	1,6	0,0	0,0
Pucón	100	14,1	8,6	20,6	42,1	14,6
Villarrica	100	25,6	20,8	31,9	18,8	3,0
Angol	100	26,7	22,9	18,5	21,9	10,0
Collipulli	100	60,4	30,6	9,0	0,0	0,0
Traiguén	100	39,5	33,4	15,3	11,8	0,0
Victoria	100	41,7	39,4	15,9	3,0	0,0
Puerto Montt	100	6,0	15,6	20,9	32,5	25,0
Frutillar	100	88,3	11,7	0,0	0,0	0,0
Llanquihue	100	48,2	51,8	0,0	0,0	0,0
Puerto Varas	100	16,2	14,4	36,0	24,8	8,7
Castro	100	23,1	37,6	15,7	17,3	6,2
Ancud	100	39,2	22,5	33,9	4,4	0,0
Quellón	100	74,6	21,7	0,0	3,7	0,0
Osorno	100	28,9	22,9	23,7	19,8	4,8
Coihaique	100	17,3	28,5	20,6	23,2	10,4
Aisén	100	32,8	27,9	29,1	9,0	1,1
Punta Arenas	100	10,3	28,9	31,5	23,7	5,6
Natales	100	13,0	27,5	36,4	22,0	1,2
Santiago	100	0,0	0,4	1,6	8,9	89,1
Cerrillos	100	50,1	29,0	16,7	4,1	0,0
Cerro Navia	100	43,9	46,0	8,6	1,5	0,0
Conchalí	100	43,3	40,5	9,0	7,2	0,0
El Bosque	100	29,9	39,9	21,7	8,5	0,0
Estación Central	100	8,0	22,9	39,3	24,3	5,4
Huechuraba	100	31,0	23,9	27,6	15,3	2,3
Independencia	100	21,6	26,0	34,0	12,3	6,1
La Cisterna	100	0,0	6,6	6,6	69,2	17,6
La Florida	100	6,3	22,5	28,0	33,2	10,0
La Granja	100	16,3	19,6	42,4	21,7	0,0

Anexo 1. Distribución porcentual de hogares por quintil según comuna (continuación).

Comuna	Quintil					
	Total	1	2	3	4	5
La Pintana	100	60,2	23,9	14,8	1,1	0,0
La Reina	100	0,0	0,0	17,1	8,3	74,6
Las Condes	100	0,0	0,0	0,0	0,8	99,2
Lo Barnechea	100	0,0	0,0	3,9	11,3	84,7
Lo Espejo	100	79,3	17,5	3,2	0,0	0,0
Lo Prado	100	17,1	16,7	27,7	33,4	5,0
Macul	100	21,8	37,7	22,8	17,7	0,0
Maipú	100	8,5	27,0	28,7	33,0	2,9
Ñuñoa	100	0,0	0,0	2,7	8,4	88,9
Pedro Aguirre Cerda	100	56,4	30,8	11,6	1,2	0,0
Peñalolén	100	8,1	22,1	22,6	26,5	20,7
Providencia	100	0,0	0,0	0,0	1,3	98,7
Pudahuel	100	23,2	25,6	37,0	11,6	2,7
Quilicura	100	41,2	34,8	19,2	4,8	0,0
Quinta Normal	100	2,2	20,1	20,0	48,5	9,2
Recoleta	100	6,7	18,4	32,5	40,0	2,4
Renca	100	63,8	23,8	8,7	3,7	0,0
San Joaquín	100	8,6	12,7	35,8	39,3	3,6
San Miguel	100	0,0	0,0	0,0	28,3	71,7
San Ramón	100	24,1	28,7	37,6	9,5	0,0
Vitacura	100	0,0	0,0	0,0	0,0	100,0
Puente Alto	100	35,9	29,8	23,7	9,4	1,2
Colina	100	46,3	36,0	6,8	10,9	0,0
San Bernardo	100	28,4	18,9	26,5	10,2	15,9
Melipilla	100	54,9	13,0	23,0	9,1	0,0
Talagante	100	31,6	40,4	18,2	9,8	0,0
Valdivia	100	7,2	11,0	17,2	32,6	32,0
Mariquina	100	92,8	0,0	7,2	0,0	0,0
Panguipulli	100	38,8	27,6	20,6	8,0	5,0
La Unión	100	44,9	31,5	18,7	4,9	0,0
Futrono	100	0,0	27,4	51,1	21,5	0,0
Río Bueno	100	58,3	29,1	4,3	8,2	0,0
Arica	100	25,2	28,1	27,5	16,1	3,1

Elaborado por el Instituto Nacional de Estadísticas.

Anexo 2. Sintaxis utilizada para aplicar el método PRINCALS.

El método PRINCALS se aplica usando el software SPSS mediante el procedimiento Análisis de Componentes Principales Categóricos (CATPCA, con sus siglas en inglés), comando que permite, una vez definidas las variables (nominal, ordinal o numérica), imputar las observaciones perdidas (en este caso con la moda) y aplicando normalización simétrica. Como resultado se generan puntajes asociados a las personas, que corresponden a las coordenadas de la primera componente principal. Luego dichos puntajes son promediados a nivel de hogar, para luego crear quintiles o deciles, lo que permite obtener una clasificación socioeconómica de los hogares. A continuación, se presenta la sintaxis empleada:

Recuadro 1: Sintaxis empleada para aplicar el PRINCALS.

SPSS

```
CATPCA VARIABLES=var1SE_CIUO var2SE_CISE var3SE_Niv_Educ var4SE_Comuna var7SE_Tipo_Viv AVALUO
/ANALYSIS=var1SE_CIUO(WEIGHT=1,LEVEL=NOMI) var2SE_CISE(WEIGHT=1,LEVEL=NOMI)
var3SE_Niv_Educ(WEIGHT=1,LEVEL=ORDI) var4SE_Comuna(WEIGHT=1,LEVEL=NOMI)
var7SE_Tipo_Viv(WEIGHT=1,LEVEL=NOMI) AVALUO(WEIGHT=1,LEVEL=NUME)
/MISSING = var1SE_CIUO(PASSIVE,MODEIMPU) var2SE_CISE(PASSIVE,MODEIMPU)
var3SE_Niv_Educ(PASSIVE,MODEIMPU) var4SE_Comuna(PASSIVE,MODEIMPU)
var7SE_Tipo_Viv(PASSIVE,MODEIMPU) AVALUO(PASSIVE,MODEIMPU)
/DIMENSION=1
/NORMALIZATION=SYMMETRICAL
/MAXITER=100
/CRITITER=.00001
/PRINT=CORR LOADING OCORR QUANT(var1SE_CIUO var2SE_CISE var3SE_Niv_Educ var4SE_Comuna var7SE_Tipo_Viv AVALUO ) VAF
/PLOT=LOADING(20)
/SAVE=OBJECT.
```

Elaborado por el Instituto Nacional de Estadísticas.