



PONTIFICIA  
UNIVERSIDAD  
CATÓLICA  
DE CHILE

**Pontificia Universidad Católica de Chile**

Facultad de Matemáticas

## **Proyecto Final**

**Entrega 1: Análisis de CASEN 2022**

## EYP2417 - Muestreo

Segundo Semestre 2025

Profesor: Guillermo Marshall Rivera

### Grupo 4

#### Integrantes:

Francisca Sepúlveda

Esteban Román

Alexander Pinto

Julian Vargas

Entrega: 24 de octubre de 2025



## Índice

<b>1</b>	<b>Objetivo General .....</b>	<b>2</b>
<b>2</b>	<b>Encuesta CASEN 2022: Descripción y Accesibilidad .....</b>	<b>2</b>
2.1	Nombre e Institución Responsable.....	2
2.2	Año(s) Disponibles y Cobertura.....	3
2.3	Variables de Interés .....	3
2.4	Variables Auxiliares .....	4
2.5	Base de Datos y Factibilidad de Acceso .....	4
2.6	Tabla Resumen de Variables .....	5
<b>3</b>	<b>Objetivos Específicos .....</b>	<b>5</b>
3.1	Eje 1: Distribución de la Pobreza.....	5
3.2	Eje 2: Brecha Salarial de Género .....	6
3.3	Preguntas Transversales.....	6
<b>4</b>	<b>Revisión de Antecedentes .....</b>	<b>6</b>
4.1	Estudios sobre Pobreza en Chile .....	7
4.2	Estudios sobre Brecha Salarial de Género .....	7
4.3	Metodología de Muestreo Aplicada .....	8
4.4	CASEN como Fuente de Datos.....	9
4.5	Validez y Pertinencia de Este Análisis .....	9



## 1 Objetivo General

La desigualdad socioeconómica es uno de los desafíos estructurales más relevantes de Chile. Este proyecto busca **contribuir al diagnóstico de la desigualdad chilena** mediante el análisis riguroso de datos de la Encuesta de Caracterización Socioeconómica Nacional (CASEN) 2022.

### Dos ejes temáticos complementarios:

#### **1. Distribución de la pobreza**

Caracterizar geográficamente (por región y zona urbana/rural) y demográficamente la incidencia de pobreza en el país.

#### **2. Brecha salarial de género**

Analizar diferencias en ingresos del trabajo entre jefes de hogar hombres y mujeres.

El análisis se fundamenta en **teoría de muestreo** (Thompson, 2012) considerando el diseño muestral complejo de CASEN. Aplicamos correctamente el **factor de expansión (expr)** en todas las estimaciones, proporcionando evidencia para políticas públicas orientadas a reducir inequidades.

## 2 Encuesta CASEN 2022: Descripción y Accesibilidad

### **2.1 Nombre e Institución Responsable**

La **Encuesta de Caracterización Socioeconómica Nacional (CASEN 2022)** es el principal instrumento de caracterización socioeconómica de la población en Chile. Es levantada por el **Ministerio de Desarrollo Social y Familia**, con el apoyo técnico del Banco Interamericano de Desarrollo (BID).



## 2.2 Año(s) Disponibles y Cobertura

La encuesta CASEN se realiza periódicamente desde 1987. La base de datos utilizada en este proyecto corresponde a:

- **Año de levantamiento:** 2022
- **Fecha de disponibilidad:** 18 de marzo de 2024
- **Cobertura:** Nacional, representativa a nivel de región, zona (urbano/rural) y comuna
- **Muestra:** Aproximadamente 202.000 personas en 70.000 hogares

## 2.3 Variables de Interés

Las variables principales utilizadas en este análisis se clasifican en dos categorías según el eje temático:

### Eje 1: Distribución de la Pobreza

- **pobreza:** Condición de pobreza (1 = No pobre, 2 = Pobre, 3 = Extremo pobre)
- **ytotalcorh:** Ingreso total corregido del hogar (pesos chilenos)
- **tot\_per\_h:** Total de personas en el hogar
- **esc:** Años de educación formal completados del jefe de hogar

### Eje 2: Brecha Salarial de Género

- **sexo:** Sexo del jefe de hogar (1 = Hombre, 2 = Mujer)
- **yoprinc:** Ingreso del trabajo principal (pesos chilenos)
- **esc:** Años de educación formal del jefe
- **edad:** Edad del jefe de hogar (años)
- **ocup:** Ocupación (códigos CIUO)



## 2.4 Variables Auxiliares

- **expr:** Factor de expansión (ponderador de muestreo) para ajustar por diseño complejo
- **region:** Región de residencia (1-16, según división política actual)
- **area:** Zona geográfica (1 = Urbano, 2 = Rural)
- **edad:** Edad de la persona encuestada
- **hogar:** Identificador del hogar

## 2.5 Base de Datos y Factibilidad de Acceso

### Formato y Ubicación

- **Nombre del archivo:** Base de datos Casen 2022 STATA\_18 marzo 2024.dta
- **Formato:** STATA (.dta), versión 18
- **Tamaño:** Aproximadamente 501 MB
- **Ubicación en proyecto:** 01\_Datos/

### Acceso y Disponibilidad

La encuesta CASEN 2022 está disponible públicamente en:

- **Sitio oficial:** Ministerio de Desarrollo Social y Familia
- **Base de datos:** <https://www.ministeriodesarrollosocial.gob.cl/>
- **Modalidad de acceso:** Descarga directa sin restricciones de acceso
- **Documentación:** Manual técnico, cuestionario y guía de variables disponibles en el mismo sitio

El acceso es **sin restricciones y de carácter público**, permitiendo replicabilidad completa del análisis. La base de datos es implementada en este proyecto y todos los análisis son reproducibles en R, Python o STATA.



## 2.6 Tabla Resumen de Variables

**Cuadro 1:** Resumen de variables principales utilizadas en el análisis

Variable	Código	Descripción y Tipo
<b>• Variables de Estratificación</b>		
Región	<code>region</code>	Región de residencia (1-16, numérica)
Zona geográfica	<code>area</code>	Urbano (1) / Rural (2)
Factor expansión	<code>expr</code>	Ponderador para representatividad (numérica)
<b>▽ Eje 1: Pobreza</b>		
Condición pobreza	<code>pobreza</code>	1=No pobre, 2=Pobre, 3=Extremo pobre
Ingreso hogar	<code>ytotalcorh</code>	Ingreso total anual corregido (pesos \$)
Escolaridad	<code>esc</code>	Años de educación formal del jefe
Tamaño hogar	<code>tot_per_h</code>	Número de personas en el hogar
<b>≡ Eje 2: Brecha Salarial de Género</b>		
Sexo jefe hogar	<code>sexo</code>	1=Hombre, 2=Mujer
Ingreso principal	<code>yoprinc</code>	Ingreso del trabajo principal (pesos \$)
Edad	<code>edad</code>	Edad del jefe de hogar (años)
Ocupación	<code>ocup</code>	Código CIUO de ocupación

Nota: Todas las variables están ponderadas usando el factor de expansión (`expr`) para asegurar representatividad a nivel nacional y regional.

## 3 Objetivos Específicos

Este proyecto busca responder a las siguientes preguntas de investigación concretas:

### 3.1 Eje 1: Distribución de la Pobreza

1. ¿Cuál es la incidencia de pobreza por región en Chile? Se busca caracterizar la distribución geográfica de la pobreza en 2022, identificando qué regiones presentan



mayores tasas de pobreza e incidencia de pobreza extrema.

2. **¿Existen diferencias significativas entre zonas urbanas y rurales?** Se comparará la incidencia de pobreza en áreas urbanas versus rurales, analizando si el factor geográfico (urbanización) es relevante en la explicación de la pobreza.
3. **¿Cuáles son los factores asociados a la pobreza?** Se analizarán variables como tamaño del hogar, educación del jefe, edad y composición demográfica para identificar patrones asociados a mayor incidencia de pobreza.

### 3.2 Eje 2: Brecha Salarial de Género

1. **¿Existe brecha salarial entre jefes de hogar hombres y mujeres?** Se comparará el ingreso del trabajo principal entre hombres y mujeres que son jefes de hogar, cuantificando la magnitud de la diferencia.
2. **¿La brecha varía según nivel educativo?** Se estratificará el análisis por años de educación para determinar si la brecha persiste o se amplía conforme aumenta la escolaridad.
3. **¿Qué variables ayudan a explicar la brecha salarial?** Se analizará cómo factores como edad, región, ocupación y educación influyen en la persistencia de la brecha de ingresos entre géneros.

### 3.3 Preguntas Transversales

1. **¿Los estimadores ponderados difieren significativamente de los no ponderados?** Se validará la importancia de considerar el diseño muestral complejo de CASEN mediante comparación de resultados.
2. **¿Cómo la metodología de Thompson (2012) se aplica correctamente?** Se documentará la implementación rigorosa de estimadores para diseños estratificados polietápicos, asegurando validez nacional y regional.

## 4 Revisión de Antecedentes

---



## 4.1 Estudios sobre Pobreza en Chile

La pobreza en Chile ha sido objeto de múltiples estudios. A continuación se resumen investigaciones previas relevantes:

### Reportes Oficiales CASEN

El Ministerio de Desarrollo Social y Familia publica anualmente reportes estadísticos basados en CASEN. Los reportes más recientes (2020-2022) documentan:

- Evolución de tasas de pobreza por región y zona (urbano/rural)
- Descomposición de pobreza por características demográficas
- Análisis de vulnerabilidad y pobreza extrema
- Comparaciones temporales desde 2015

Estos reportes sirven como línea base para validación de nuestros estimadores y como contexto para interpretar resultados.

### Literatura Académica

Trabajos recientes en Chile incluyen:

- **Contreras y Núñez (2017):** Análisis de movilidad social y persistencia de pobreza intergeneracional usando datos CASEN
- **Agostini y Brown (2007):** Estimación de elasticidad ingreso-pobreza en zonas urbanas y rurales
- **CEPAL (2021):** Perspectivas socioeconómicas en América Latina y el Caribe, incluida caracterización de pobreza en Chile post-COVID

## 4.2 Estudios sobre Brecha Salarial de Género

La brecha de género en ingresos es ampliamente documentada en literatura internacional y nacional:



## Reportes Nacionales

- **Instituto Nacional de Estadísticas (INE):** Publica cifras sobre brechas salariales a partir de encuestas de empleo
- **Servicio Nacional de la Mujer (SERNAMEG):** Reportes anuales sobre equidad de género, incluyendo análisis salarial
- **Ministerio del Trabajo:** Diagnósticos sobre igualdad laboral y paridad de género

## Literatura Académica

Estudios aplicados a Chile:

- **Contreras et al. (1999):** Descomposición de la brecha salarial por género usando encuestas de empleo
- **Aldén y Hammarstedt (2014):** Análisis de brecha salarial en el contexto latinoamericano con datos CASEN
- **Fuentes y Palma (2014):** Estimación de discriminación salarial por género en Chile controlando por educación y experiencia

## 4.3 Metodología de Muestreo Aplicada

La correcta especificación del diseño muestral es fundamental. Literatura de referencia:

- **Thompson (2012):** “Sampling” - Referencia principal para diseños estratificados polietápicos y estimadores insesgados
- **Lumley (2010, 2011):** Implementación computacional de survey statistics en R
- **Särndal et al. (1992):** “Model Assisted Survey Sampling” - Fundamentos teóricos de inference bajo diseños complejos
- **Lohr (2019, 3<sup>a</sup> ed.):** “Sampling: Design and Analysis” - Aplicaciones prácticas y calibración de ponderadores



## 4.4 CASEN como Fuente de Datos

CASEN es ampliamente utilizada en investigación social chilena. Estudios que utilizan CASEN incluyen:

- Análisis de desigualdad y distribución de ingresos (Gini, polarización)
- Estudios de política social (evaluación de programas, focalización)
- Investigaciones sobre educación, empleo y características demográficas
- Comparaciones regionales y zonales

La base de datos de CASEN 2022 es reciente y pertinente para diagnosticar la situación post-pandemia en Chile.

## 4.5 Validez y Pertinencia de Este Análisis

Este proyecto se justifica por:

1. **Novedad temporal:** CASEN 2022 es reciente y aún no ha sido extensivamente analizada en comparación con rondas anteriores
2. **Diseño metodológico riguroso:** Aplicación correcta de Thompson (2012) asegura validez de inferencia
3. **Multidimensionalidad:** Análisis simultáneo de pobreza y brecha salarial permite identificar interconexiones
4. **Relevancia política:** Contribuye a debate sobre desigualdad socioeconómica y políticas públicas en Chile

