



INSTITUTO DE ESTUDIOS URBANOS  
Y TERRITORIALES  
FACULTAD DE ARQUITECTURA, DISEÑO Y ESTUDIOS URBANOS

# Regionalización frente a Identidad Territorial: Unidades Primarias de Muestreo basadas en la Regionalización para delimitar Zonas Urbanas Funcionales.

Ricardo Truffello, Sebastián Valdivia, Martín Rosas, Isidro Puig.

Noviembre de 2023

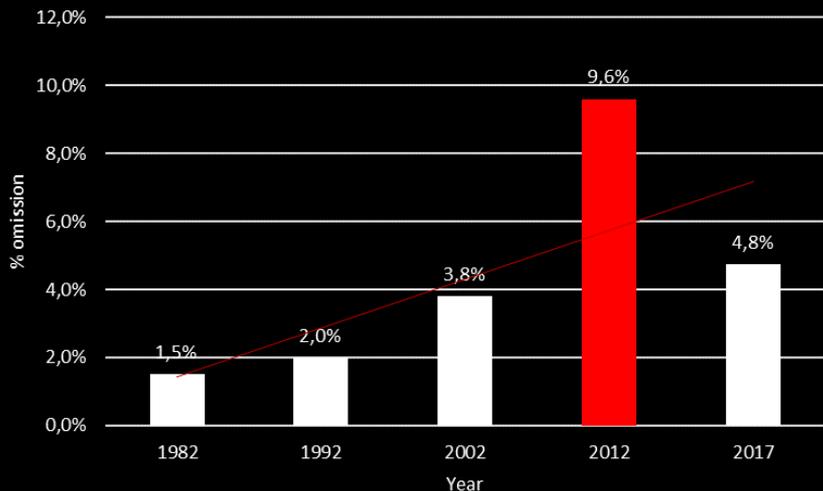
Sección

# Antecedentes del problema de investigación

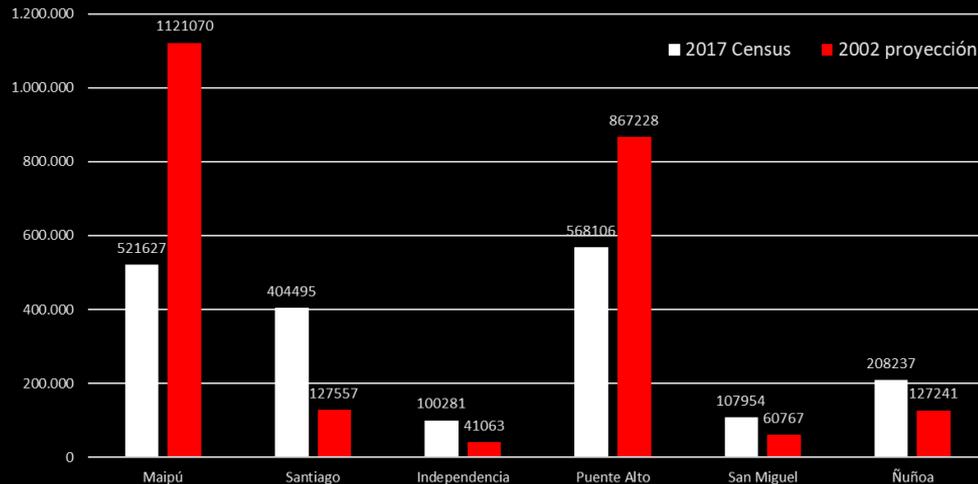
# Representatividad territorial

Datos Censales y Encuesta CASEN

Debido a la (1) invalidación del censo de población 2012 (Bravo, Larrañaga, Millán, Ruiz, & Zamorano, 2013), (2) el censo 2017 abreviado y (3) los problemas de representatividad comunal de la Encuesta de Caracterización Socioeconómica (CASEN), se hace necesario plantear nuevas metodologías que permitan generar una alternativa competitiva a las de muestreo tradicional.



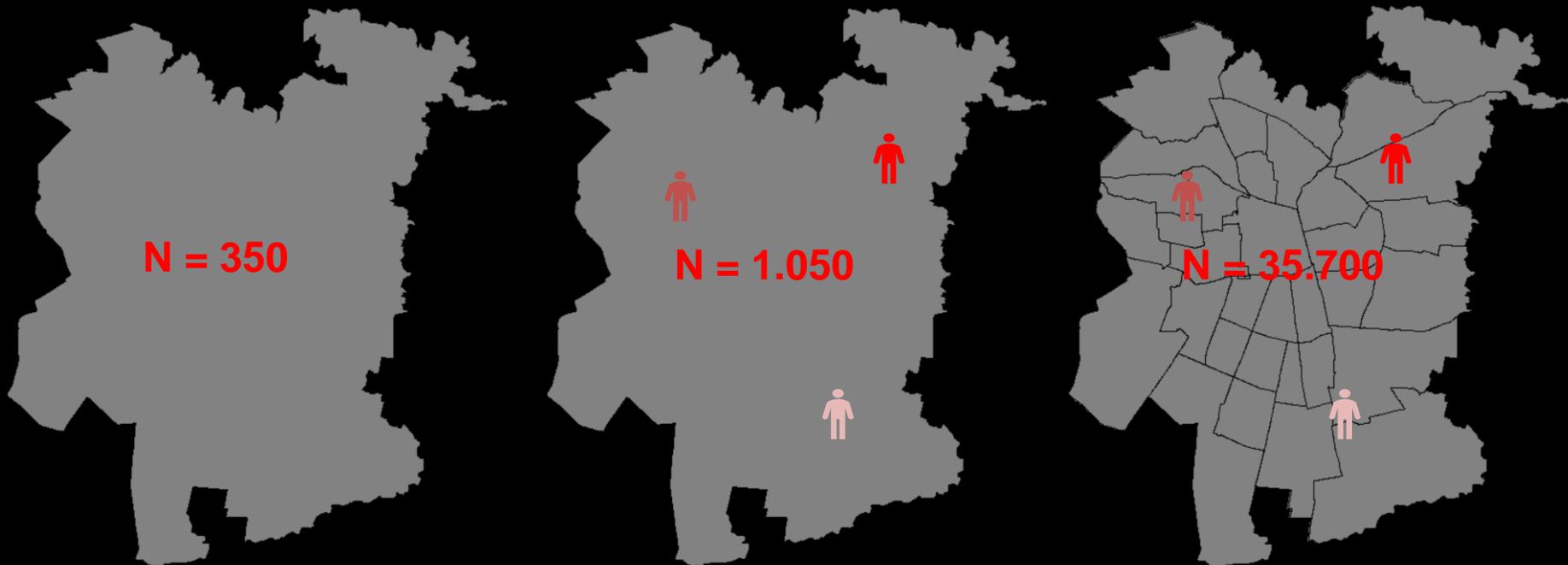
Fuente. Elaboración propia sobre la base de: (Villalón & Vera, 2011); (Bravo, Larrañaga, Millán, Ruiz, & Zamorano, 2013).



Fuente. Elaboración propia sobre la base de: Censo 2017, INE; CASEN 2017, MDS.

# Tipos de Muestreo

Del muestreo aleatorio simple al muestreo espacializado



# Tipos de Muestreo

Del muestreo aleatorio simple al muestreo espacializado

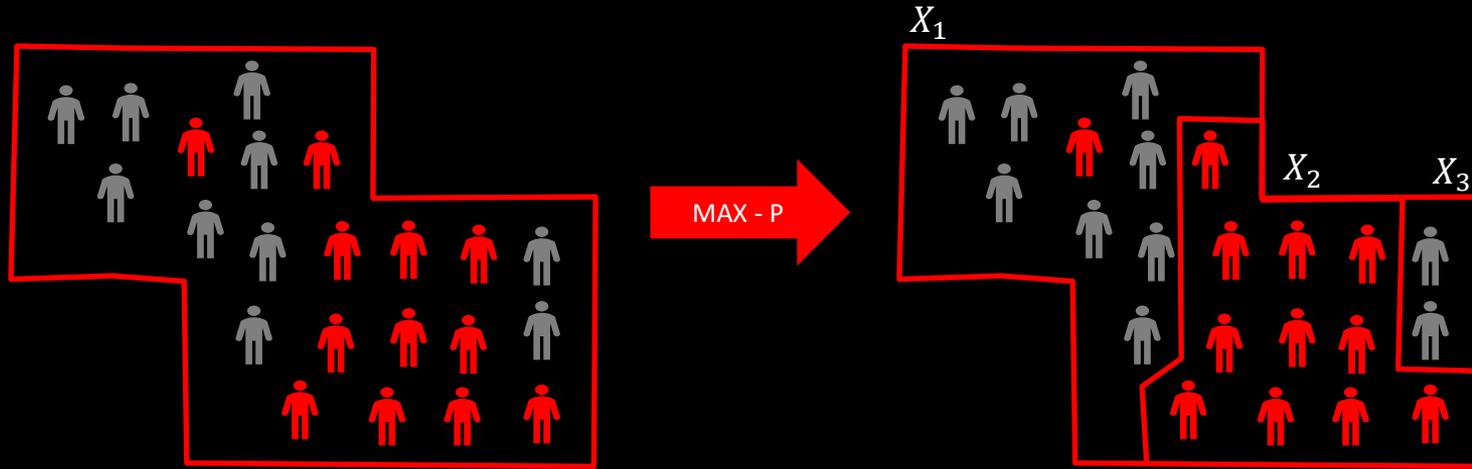


La representatividad en las encuestas está dada por la aleatoriedad, bajo la premisa fundamental de que las observaciones son **“independientes e idénticamente distribuidas”** (Brus & De Grujter, 1997; Stock, Ward, Thorson, Jannot, & Semmens, 2019), lo que no es necesariamente cierto en la distribución espacial de los datos.

Sin embargo, en la Ciudad las variables no tienen un comportamiento aleatorio, sino que en el mayor parte de los casos este **comportamiento es aglomerado**, situación que se puede comprobar fácilmente aplicando el índice de Moran

# Tipos de Muestreo

Del muestreo aleatorio simple al muestreo espacializado

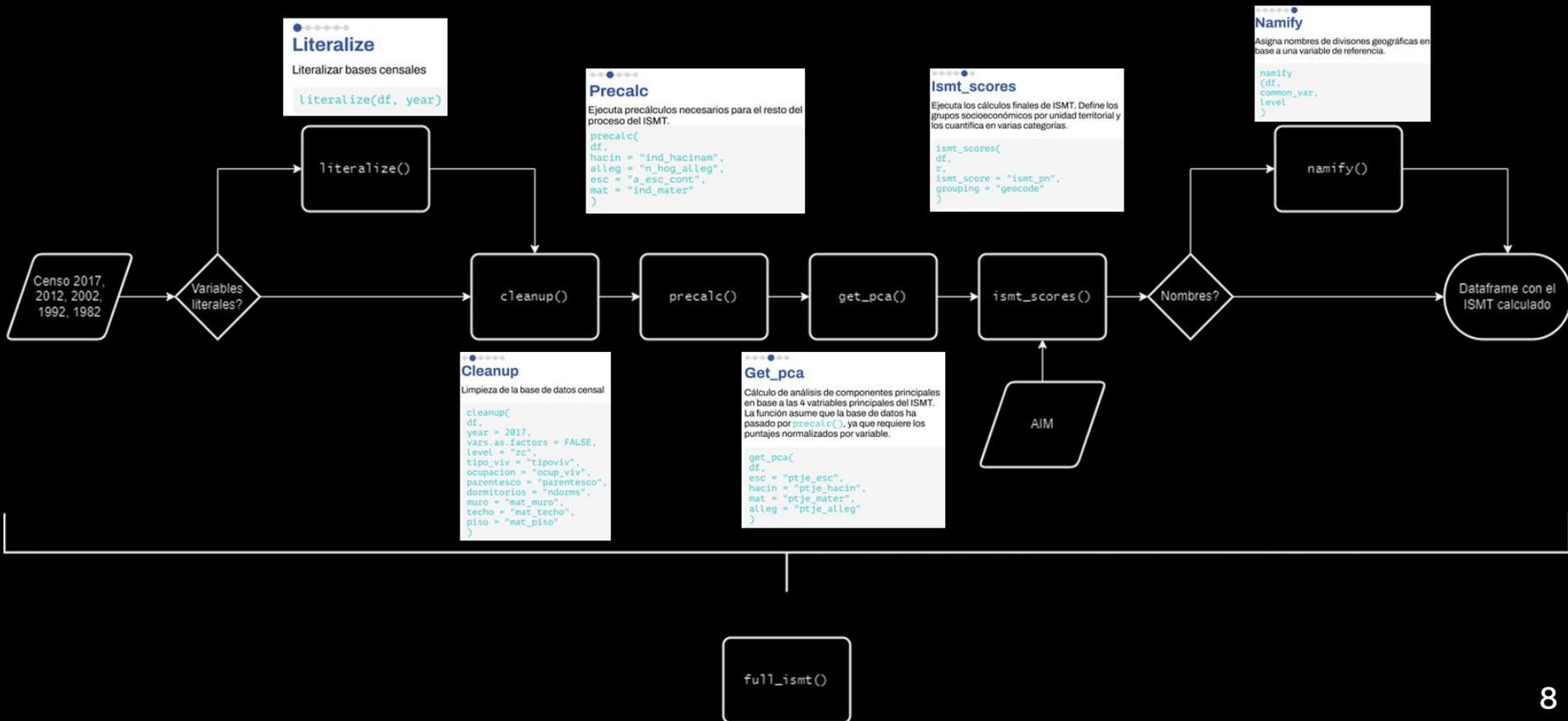


La idea del muestreo espacializado minimizar la redundancia de información en donde exista poca varianza, esto permite reducir el error de muestreo o en su defecto disminuir el tamaño de las muestras.

La zonificación de la derecha, en teoría requiere sólo una persona por cada zona.

# Variable Estratificación

## Índice Socio-Material Territorial (ISMT)



# Variable Estratificación

## Índice Socio-Material Territorial (ISMT)



### R package `ismtchile`:

Inicio Metodología >... Instalación Flujo de trabajo Limitaciones Resultados Metadata Autoría

#### Flujo de trabajo

El flujo de trabajo del paquete está pensado para satisfacer la mayor cantidad de necesidades del usuario, desde obtener la data y los shapefiles de origen hasta realizar el cálculo del ISMT paso por paso, dando también la opción de omitir el paso a paso y calcular el índice directamente con 1 sola función. La mejor descripción del flujo se encuentra en la *cheatsheet* del paquete, disponible para descarga [aquí](#). Este documento, además del flujo, contiene la descripción detallada de los parámetros de cada una de las funciones.

ismtchile-cheatsheet.pdf

1 / 2

67%



#### Paquete RStudio Índice Socio Material Territorial V2.1.6

ESP  
Paquete creado con el fin de facilitar el cálculo y distribución del Índice Socio-Material Territorial.

#### Literalize

Literalizar bases censales

```
literalize(df, year)
```

#### Arguments

- df** /objeto  
data.frame correspondiente a la base de datos original de los censos 2017, 2012, 2002, 1992 o 1982.
- year** /integer  
No provee Default para forzar la explicación del año de la base.

#### Value

objeto `data.frame` conteniendo solo las variables necesarias para los cálculos siguientes.

#### vars.as.factors /boolean

Si las variables están como factores (como en las bases censales originales), cambiar a TRUE. Default es FALSE, ya que se asume que la base pasó por la función `literalize()`.

#### level /string

Nivel de agrupación de los datos finales. Acepta valores `z` (zona censal) y `m` (manzana). El nivel de manzana no está disponible para el 2017 debido al secreto estadístico de la base de datos de origen. Default es `z`.

#### tipo\_viv /string

Nombre del campo de tipo de vivienda. Default es `tipoviv`.

#### ocupacion /string

Nombre del campo de ocupación de la vivienda. Default es `ocup_viv`.

#### parentesco /string

Nombre del campo de parentesco. Default es `parentesco`.

#### dormitorios /string

Nombre del campo con el número de dormitorios del hogar. Default es `admas`.

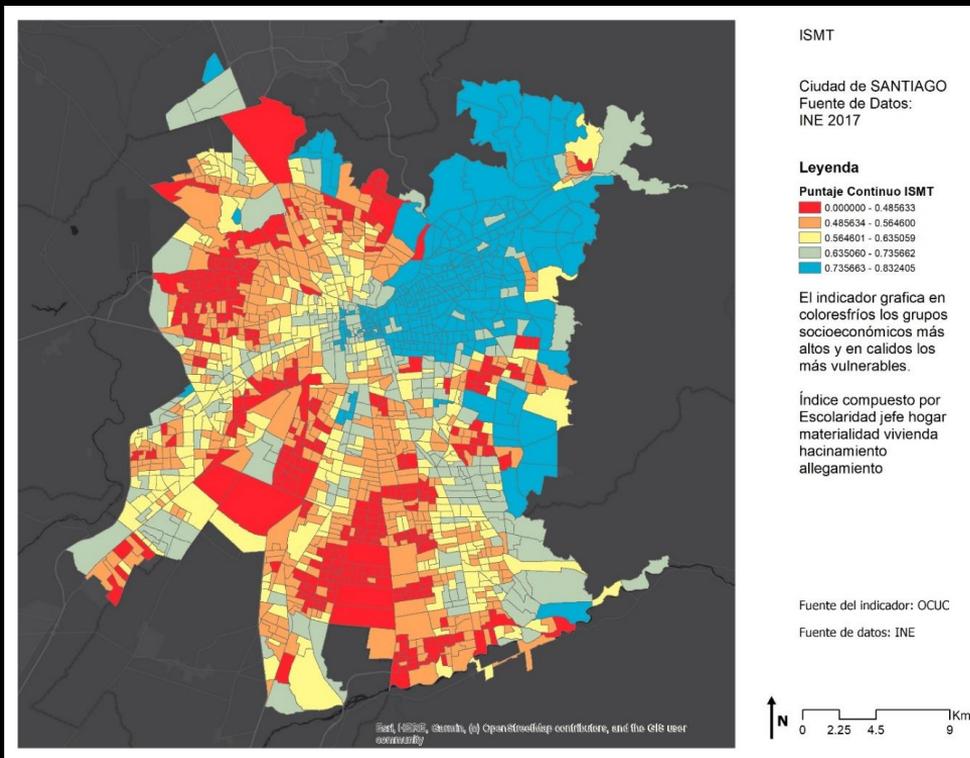


<https://ismtchile.mra-portal.dev/home>

`install.packages("ismtchile")`

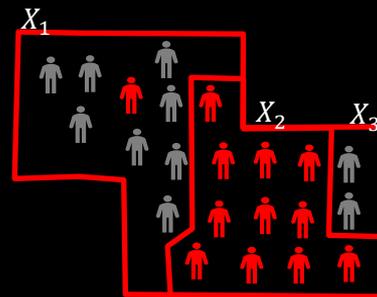
# Regionalización

## Algoritmos de Regionalización y Muestreo Espacializado



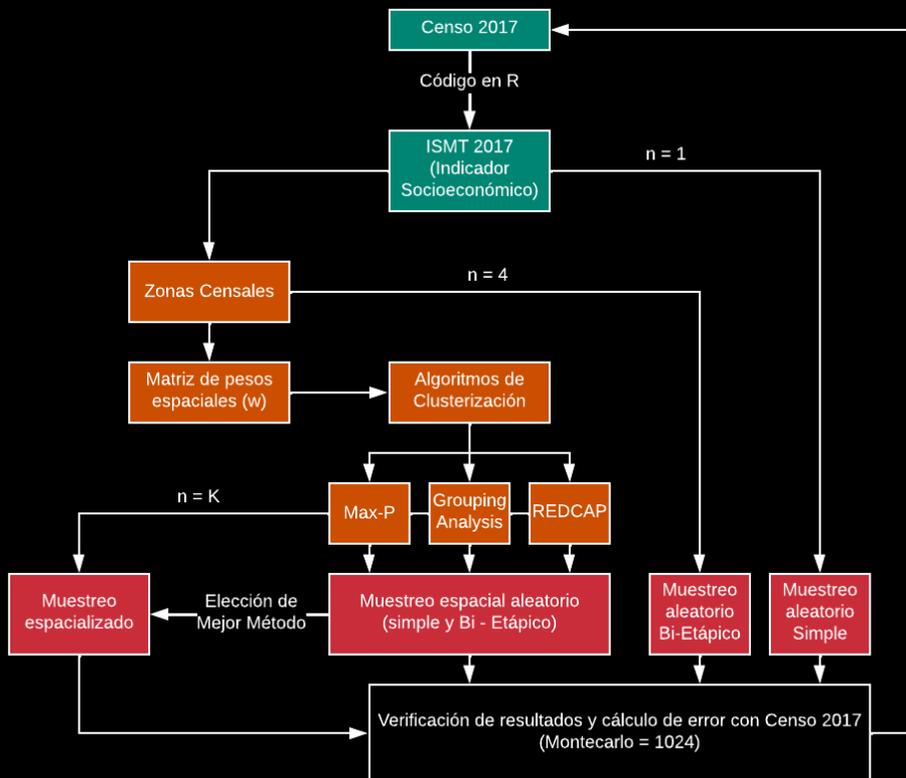
**Max-P:** algoritmo de partición propuesto por Duque en (Duque et al., 2012). Trabaja con una matriz de pesos espaciales, y puede segmentar por un piso de población, lo que da la posibilidad de dar un peso regular a las zonas. Posee métodos heurísticos (Simulated Annealing y búsqueda tabú) para optimizar los procesos de combinatoria y evitar caer en óptimos locales.

$$z = \left( - \sum_{k=1}^n \sum_{i=1}^n x_i^{k0} \right) * 10^k + \sum_i \sum_{j|j>i} d_{ij} t_{ij}$$



# Regionalización

## Muestreo Espacializado



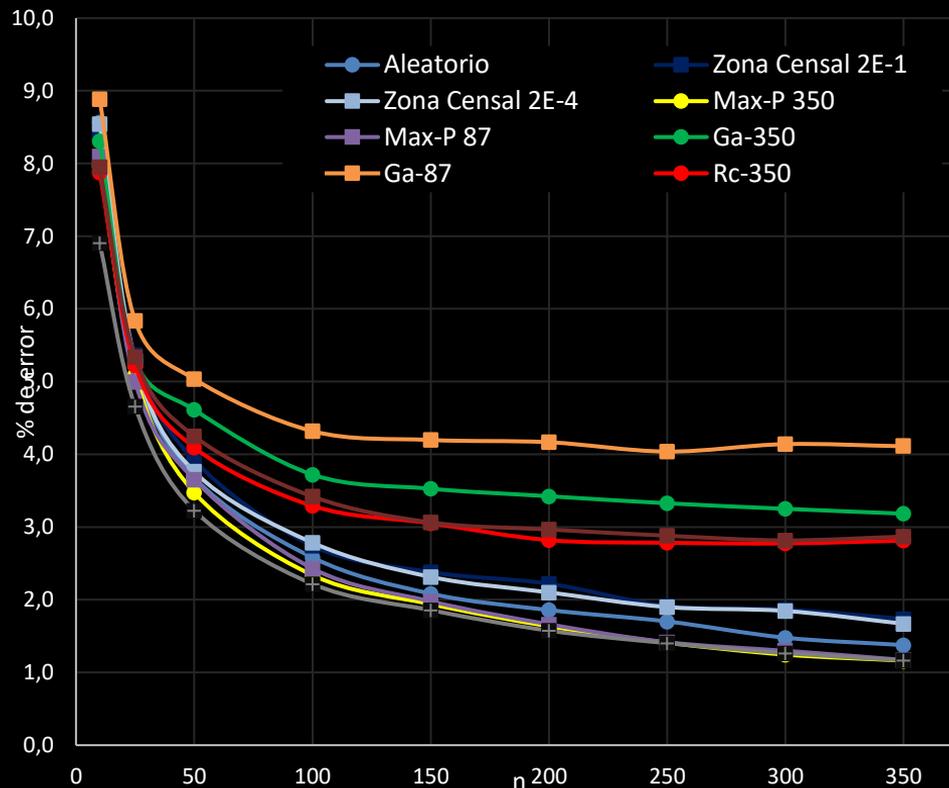
Nombre	Tipo de muestreo	N	k (número de zonas)	Progresión del muestreo
Aleatorio	se escoge aleatoriamente un hogar sobre base de 1,8MM	350	no aplica	
Zona Censal 2E-1	aleatorio en dos etapas; se escoge aleatoriamente una zona y luego un hogar por zona	350	1665	
Zona Censal 2E-4	aleatorio en dos etapas; se escoge aleatoriamente una zona y luego cuatro hogares por zona	87*4	1665	
Max-P 350	aleatorio espacializado. Se parcela el territorio en 350 zonas y se escoge un hogar por zona	87*4	350	
Max-P 87	aleatorio espacializado. Se parcela el territorio en 350 zonas y se escogen cuatro hogares por zona	87*4	87	10; 25; 50; 100; 150;
Ga-350	aleatorio espacializado. Se parcela el territorio en 350 zonas y se escoge un hogar por zona	350	350	200; 250; 300; 350
Ga-87	aleatorio espacializado. Se parcela el territorio en 350 zonas y se escogen cuatro hogares por zona	87*4	87	
Rc-350	aleatorio espacializado. Se parcela el territorio en 350 zonas y se escoge un hogar por zona	350	350	
Rc-87	aleatorio espacializado. Se parcela el territorio en 350 zonas y se escogen cuatro hogares por zona	87*4	87	
E-Max-P 350	muestreo espacializado. Se parcela el territorio en n= k y se escoge un hogar por persona.	350	n = k (350)	

## Antecedentes Previos



Algoritmos de Regionalización y Muestreo Espacializado

Se reduce el error, en el tramo entre 250 y 350 observaciones, entre un 15 y 17%, y en promedio para todos los  $n$ , en un 8%. La explicación se debe, en parte, a la optimización heurística, clave sobre todo en las particiones con mayor número de zonas, por la evidente mayor probabilidad de combinatorias presentes para el algoritmo.



## Objetivos



Fondecyt DAMES (Diseño Muestreo Espacializado)

Generar **regionalización** a partir de **unidades** de detalle como las **manzanas** permitirán **mejorar** la **homogeneidad** de las **Unidades Primarias de Muestreo**, disminuyendo significativamente el error muestral

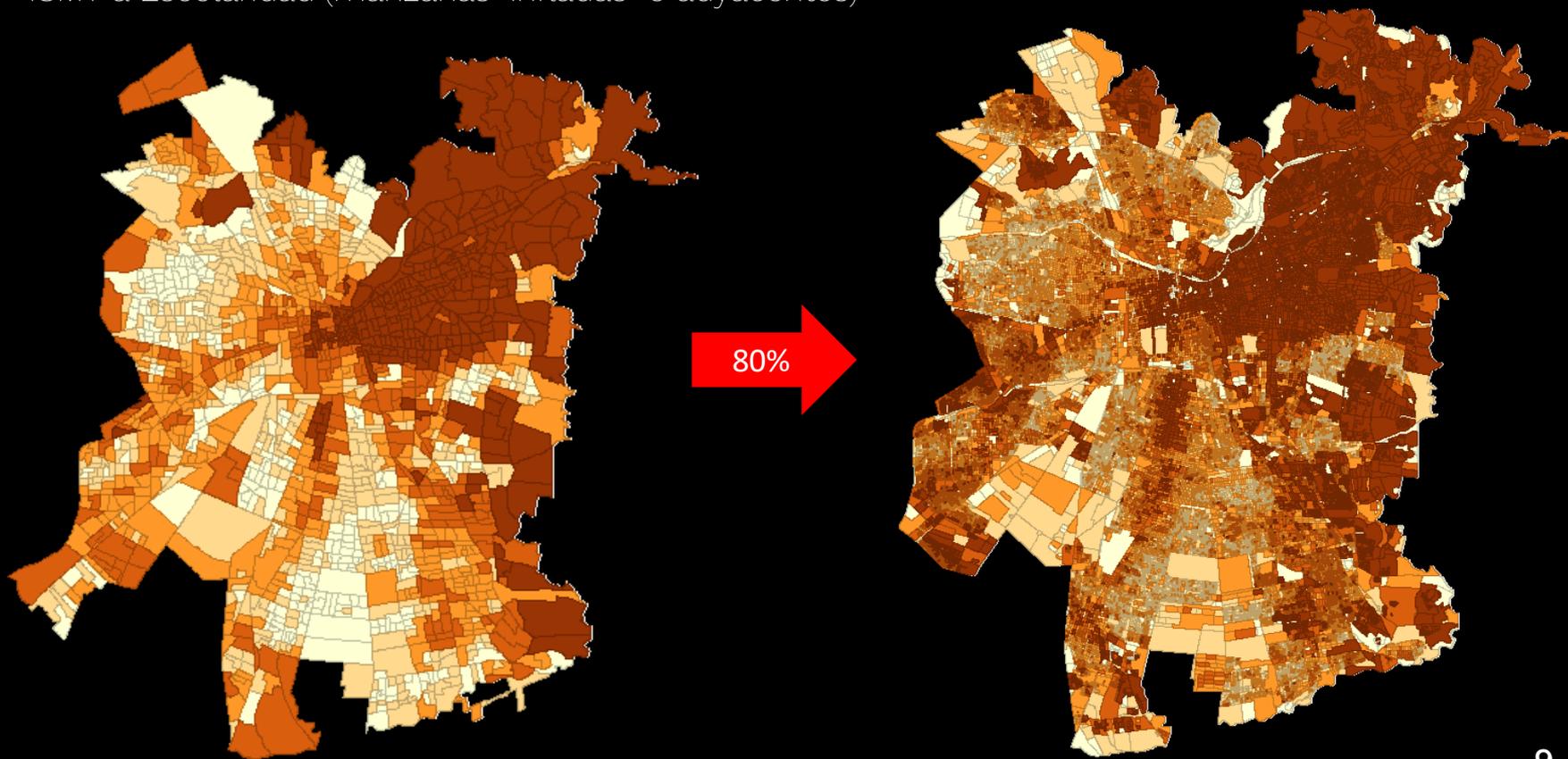
Es necesario contar con **estandarizaciones espaciales y temporales** en el uso de variables de estratificación (caso del ISMT)

Obtener detalles en la medición por zona, permite orientar **estrategias diferenciadas** al momento de realizar muestreo espacializado

Verificar si unidades espaciales con **identidad territorial** como los barrios o las unidades vecinales son una alternativa competitiva a la regionalización

## Variable Estratificación

ISMT a Escolaridad (manzanas “infladas” o adyacentes)



Sección

# Resultados

# Resultados

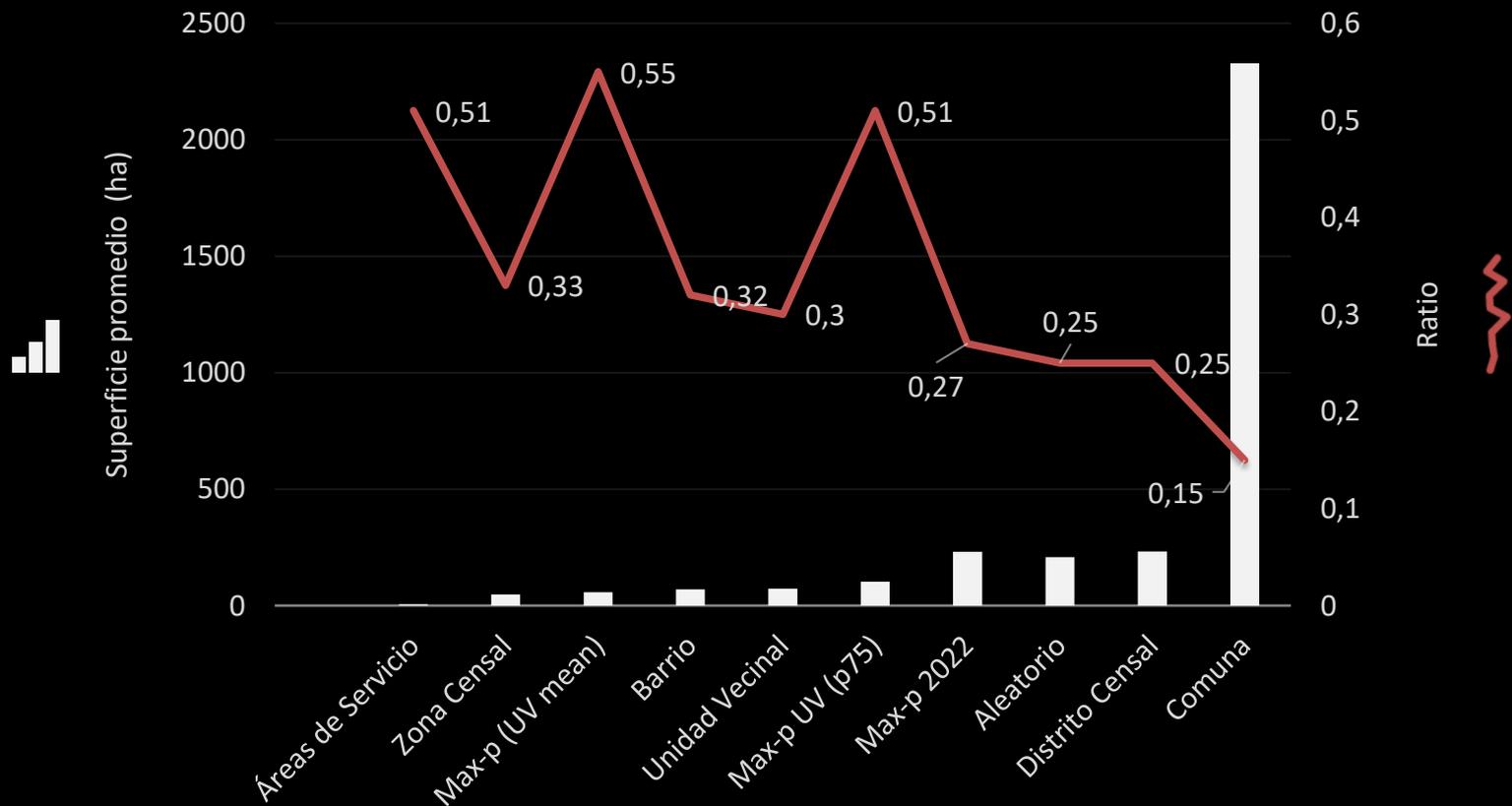


Aplicación Max-P

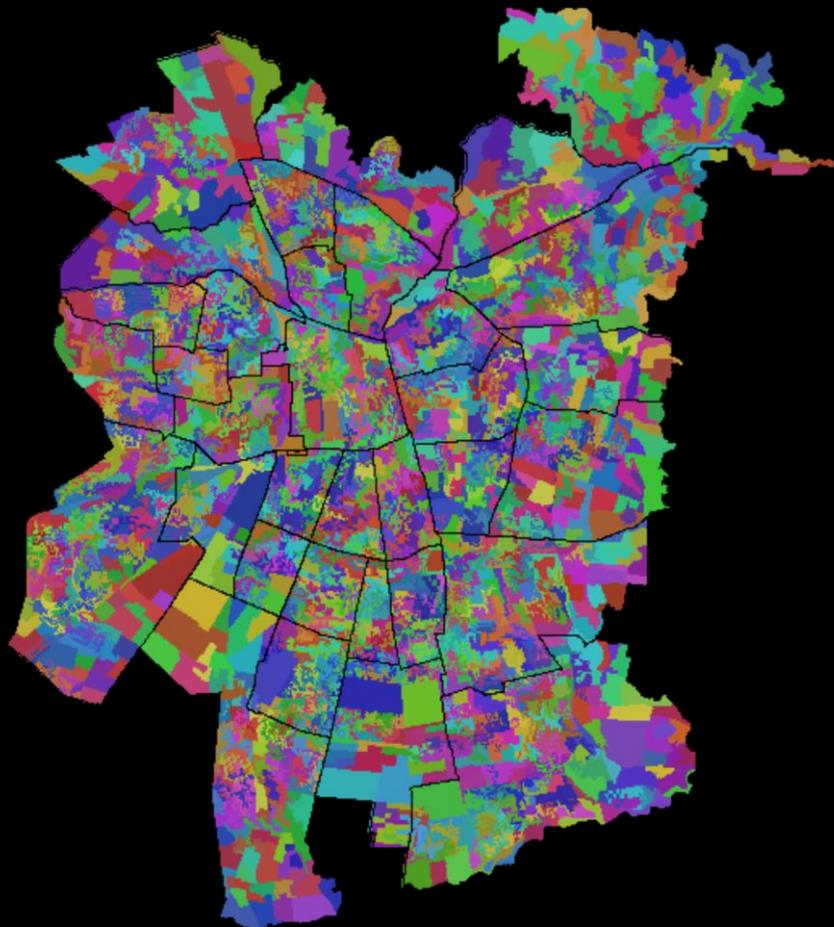
Escala análisis	Promedio superficie (ha)	D. Estándar superficie (ha)	Mínimo superficie (ha)	Máxima superficie (ha)	ratio (WCSS / TSS)
Áreas de Servicio	5,7	240,6	0,0	28228,9	0,51
Zona Censal	48,4	87,8	1,4	1572,4	0,33
<b>Max-p (UV mean)</b>	<b>58,6</b>	<b>39,9</b>	<b>38,0</b>	<b>421,6</b>	<b>0,55</b>
Barrio	70,7	633,2	0,5	21077,6	<b>0,32</b>
Unidad Vecinal	73,6	144,1	0,7	1978,2	<b>0,30</b>
<b>Max-p UV (p75)</b>	<b>103,5</b>	<b>45,7</b>	<b>73,6</b>	<b>428,9</b>	<b>0,51</b>
Max-p 2022	230,8	271,0	10,3	1779,2	0,27
Aleatorio	208,3	72,6	159,9	659,0	0,25
Distrito Censal	232,8	320,5	37,3	3071,0	0,25
Comuna	2328,0	1774,9	630,5	7665,0	0,15

# Resultados

Aplicación Max-P



## Resultados



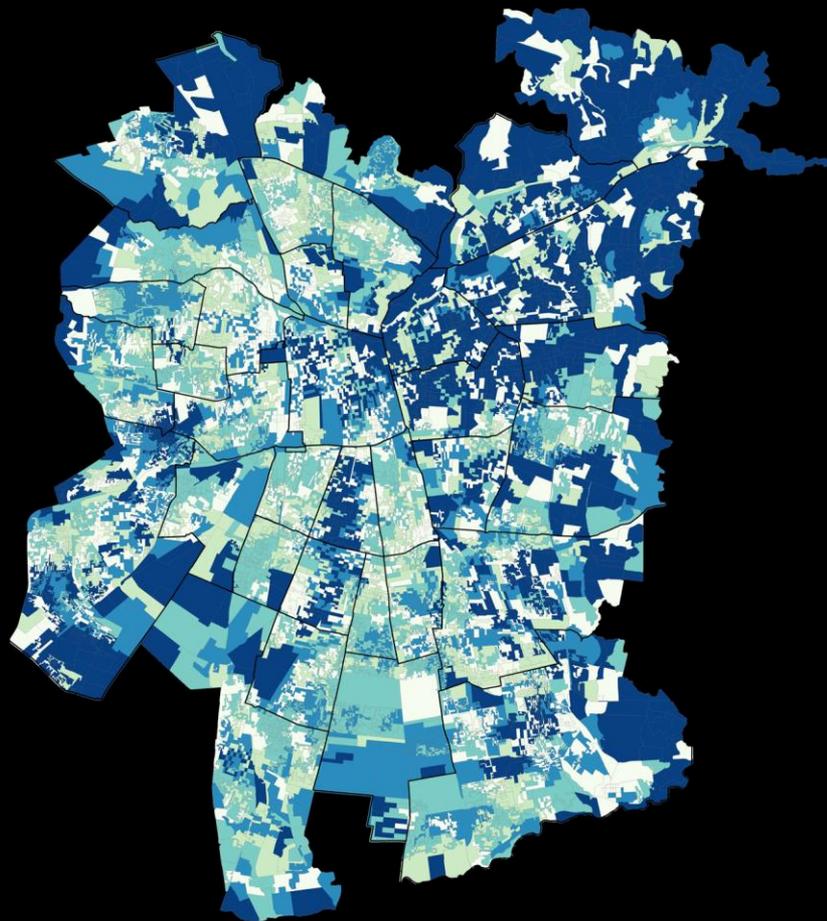
1351 clusters o "regiones"

Esc. Jefatura hogar (v° segmentación) –  
manzanas censales adyacentes

MPE de tipo "queen" de primer orden

Ratio general: 0.54

## Resultados



1351 clusters o "regiones"

Esc. Jefatura hogar (v° segmentación) –  
manzanas censales adyacentes

MPE de tipo "queen" de primer orden

Ratio general: 0.54

*Ratio por cluster*



0.1

1

## Conclusiones



Regionalización e Identidad Territorial

Se confirma que el **uso** de procesos de **regionalización** a partir de **manzanas censales mejora** en un **89%** los resultados de **homogeneidad**. Esto se podría traducir es disminuciones del marco muestral de entre un 30 y un 50% dependiendo de las ciudades en que se aplique.

Las **Unidades Vecinales** destacan como zonas fiables para su utilización como unidades primarias de muestreo, gracias a su apropiación por parte de la población, municipios y otros estamentos públicos (MIDESO); a pesar de estos **no son competitivas contra** proceso de **regionalización**.

Las **Comunas** a pesar de tener pequeños ratios de regionalización, por su escala son zonas altamente representativas, lo que desde el punto de vista numérico justifica su utilización para su uso en la generalización y cálculo de indicadores

## Conclusiones



Regionalización e Identidad Territorial

La **regionalización** como resultado para la generación de **zonas urbanas funcionales**, es desde el punto de vista cuantitativo la mejor opción, ya que aporta regiones homogéneas consistentes y dinámicas en el tiempo

Sin embargo, se debe trabajar en aspectos vinculados a su **compacidad cartográfica** y en una manera de **incorporarlas en el imaginario funcional y operativo** (como ha pasado sin querer con las zonas censales).

Para esto se debe avanzar en una regionalización operativa a nivel nacional que puedan ocupar instrumentos como el Censo o la CASEN de forma regular, así como también baterías de indicadores como el SIEDU o el SICVIR.

¡Gracias!

