

# **El análisis espacial: problema de la unidad espacial modificable**

Sobre cómo influye la unidad espacial escogida en el análisis del nivel educativo del jefe de hogar en la ciudad de Buenos Aires

# Objetivos del taller

- Comprender, manipular y transformar proyecciones
- Cargar y visualizarlos datos a software SIG
- Seleccionar por atributos o por ubicación
- Realizar uniones por atributos (joins) y uniones por ubicación (spatial joins)
- Comprender el funcionamiento e identificación de las diferentes unidades espaciales utilizadas por el Censo
- -Comprender cabalmente el impacto del problema de la unidad espacial modificable en el análisis espacial (escala y zonificación) y la falacia ecológica
- -Proponer los elementos fundamentales de la técnica de aporcionamiento de polígonos y sus variantes (en base a área o unidades)

# Organización del taller

## Primera parte:

- Breve repaso de Introducción a Software GIS
- Definición del Problema de la Unidad Espacial Modificable (PUEM)
- Ejemplo del aspecto de *escala* del PUEM: agregar datos para unidades espaciales macro a partir de unidades espaciales micro, perfectamente anidadas.

## Segunda parte:

- Ejemplo del aspecto de *zonificación* del PUEM: distribuir totales y proporciones entre unidades espaciales no anidadas

# El Problema de la Unidad Espacial Modificable (PUEM)

Cuando trabajamos con estadísticos, casi siempre los calculamos a lo largo de un índice  $i$ , y debemos reflexionar mucho sobre las diferencias entre unidades de análisis y unidades de información

$$\sum_{i=m}^n a_i = a_m + a_{m+1} + a_{m+2} + \cdots + a_n$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

# Las unidades espaciales y su recorte espacial

Cuando los humanos construimos conocimiento, los recortes que hacemos de nuestra realidad y objeto de estudio no son triviales.

Menos aún en el espacio. Cualquier espacio puede ser recortado de, literalmente, infinitas formas. La forma de recortar concreta que elijamos incide en nuestro análisis.

Veremos dos formas puntuales que conforman el PUEM:

- La escala
- La zonificación

# Mapas, escalas y el conocimiento humano

**Del Rigor en la Ciencia**

*Jorge Luis Borges*

En aquel Imperio, el Arte de la Cartografía logró tal Perfección que el mapa de una sola Provincia ocupaba toda una Ciudad, y el mapa del Imperio, toda una Provincia. Con el tiempo, estos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenía el tamaño del Imperio y coincidía puntualmente con él.

Menos Adictas al Estudio de la Cartografía, las Generaciones Siguiertes entendieron que ese dilatado Mapa era Inútil y no sin Impiedad lo entregaron a las Inclemencias del Sol y los Inviernos. En los desiertos del Oeste perduran despedazadas Ruinas del Mapa, habitadas por Animales y por Mendigos; en todo el País no hay otra reliquia de las Disciplinas Geográficas.

*Suárez Miranda, Viajes de Varones Prudentes, Libro Cuarto, Cap. XLV, Lérida, 1658.*

# Zonificación

Michel Foucault, Prefacio a Las palabras y las cosas (1966):

Este libro nació de un texto de Borges. De la risa que sacude, al leerlo, todo lo familiar al pensamiento al nuestro: al que tiene nuestra edad y nuestra geografía trastornando todas las superficies ordenadas y todos los planos que ajustan la abundancia de seres, provocando una larga vacilación e inquietud en nuestra práctica milenaria de lo Mismo y lo Otro. Este texto cita\* "cierta enciclopedia china" donde está escrito que "los animales se dividen en:

1. pertenecientes al Emperador,
2. embalsamados,
3. amaestrados,
4. lechones,
5. sirenas,
6. fabulosos,
7. perros sueltos,
8. incluidos en esta clasificación,
9. que se agitan como locos,
10. innumerables,
11. dibujados con un pincel finísimo de pelo de camello,
12. etcétera, m] que acaban de romper el jarrón,
13. que de lejos parecen moscas

En el asombro de esta taxonomía, lo que se ve de golpe, lo que, por medio del apólogo, se nos muestra como encanto exótico de otro pensamiento, es el límite del nuestro: la imposibilidad de pensar esto.

\*El idioma analítico de John Wilkins, Otras inquisiciones, Jorge Luis Borges Bs. As. 1960

# PUEM: efecto escala y zonificación

**El efecto de escala:** El análisis estadístico basado en datos agregados en áreas de diferentes tamaños producirá resultados diferentes.

Efecto escala

630	651	174	162	169	161
		150	144	161	160
		131	127	162	163
		142	148	158	158
548	641				

**El efecto de zonificación:** Dos zonas de igual área y forma, pueden producir diferentes resultados de acuerdo a su ubicación.

Efecto zonificación

968	972	965	
530		526	979



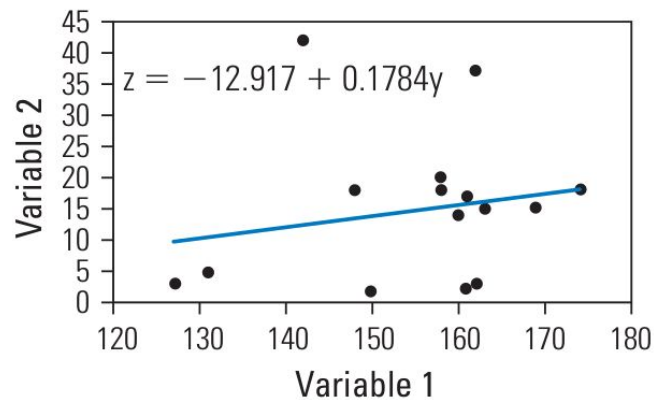
# Impacto del PUEM en un estadístico

Variable 1

174	162	169	161
150	144	161	160
131	127	162	163
142	148	158	158

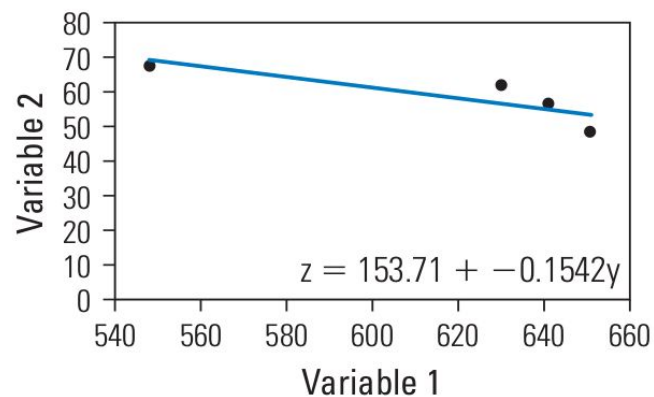
Variable 2

18	37	15	17
2	5	2	14
5	3	3	15
42	18	18	20

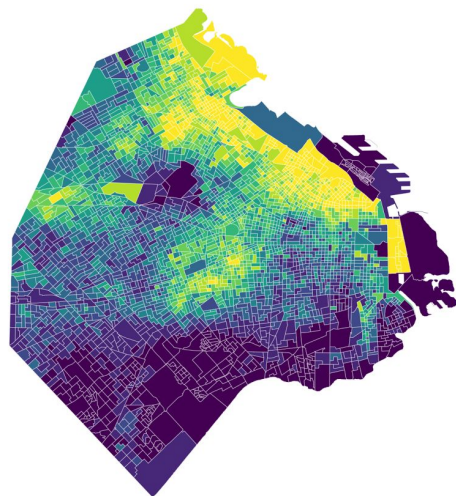
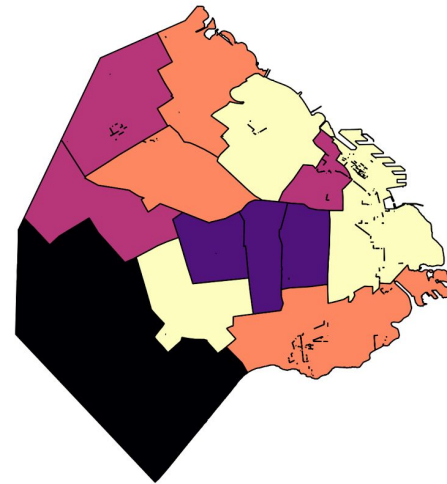
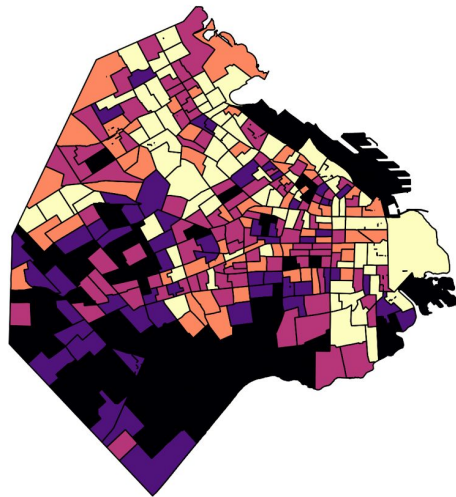


630	651
548	641

62	48
68	56

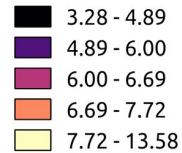


# Ejemplo :



Desvios estandar

comunas

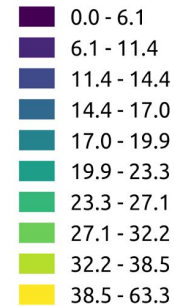


fracciones



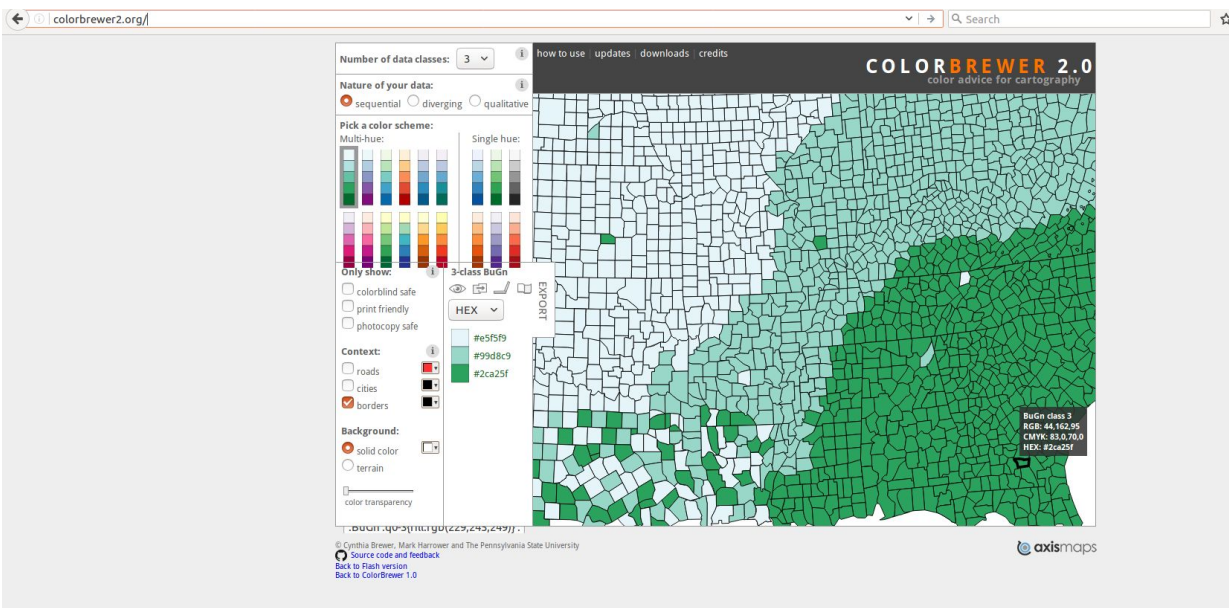
Jefe Universitario

radiosPosgar07-3



# Técnicas y estrategias de visualización

Qué colores (en base a cuantiles, rangos iguales, jenks) es una técnica en el marco de una estrategia de visualización en función del objetivo del mapa: resaltar agrupaciones, contrastes, ofrecer una lectura sencilla, mostrar rangos que tienen sentido propio (tarifas o porcentajes)



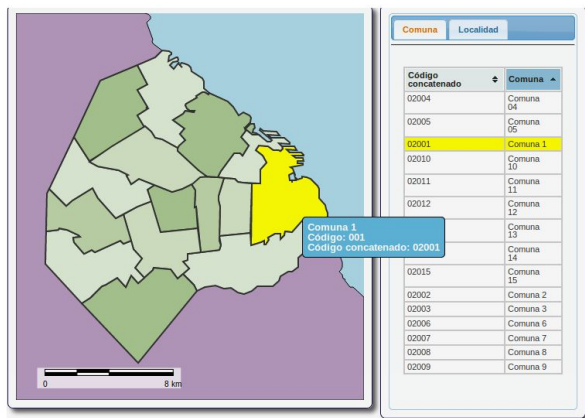
Colorbrewer2.org

# Primera parte: efecto escala

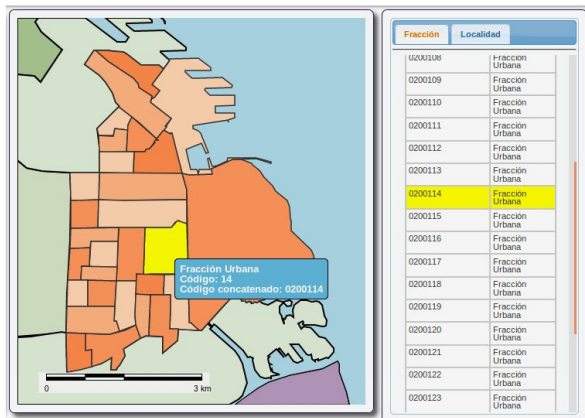
- Cargar geografía radios con datos
- Disolver en áreas macro (fracciones y departamentos)
- Calcular estadísticos resumen (promedio y desvío estándar)
- Realizar uniones de la tabla con estadísticos a la geografías disueltas

Disolver + Uniones

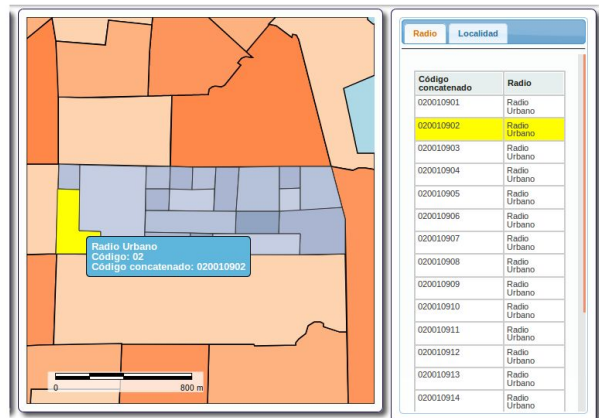
# Unidades censales



Departamento



Fracción



Radio

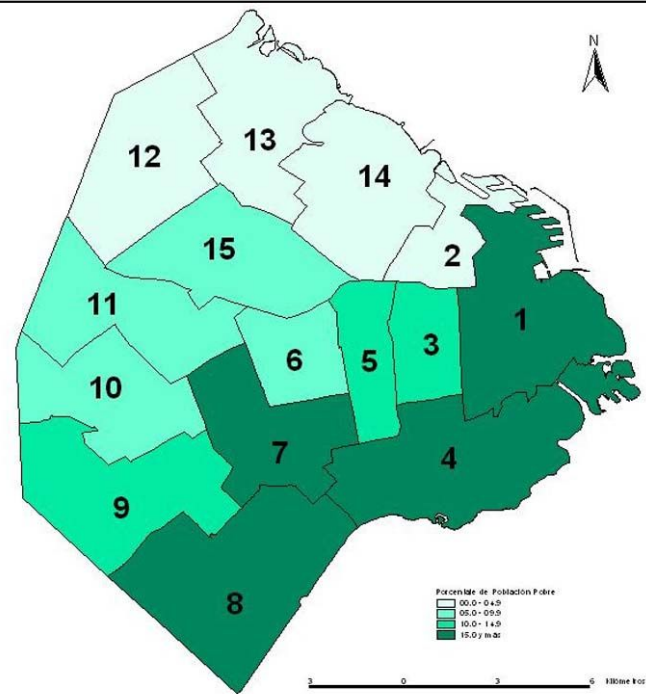
Estas son las unidades espaciales básicas del censo. Son identificadas con un campo comun: CO\_FRAC\_RA. Por ejemplo Comuna 1 Fracción 2 Radio 10 es 1\_2\_10

<http://geoservicios.indec.gov.ar/codgeo/index.php?pagina=mapas>

# Comparar con EAH (CABA)

Algunas veces tenemos que comparar con datos agregados de otras fuentes como la Encuesta Anual de Hogares (EAH CABA) cuya muestra es representativa sólo a nivel de comuna.

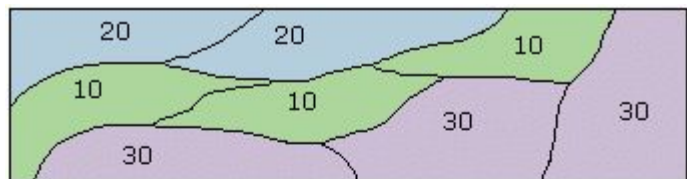
Porcentaje de población bajo la línea de pobreza (LP) según comuna.  
Ciudad de Buenos Aires. Año 2006



Fuente: Dirección General de Estadística y Censos (Ministerio de Hacienda - G.C.B.A.), EAH 2006

Nota: Se excluyen a los habitantes de las comunas 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15.

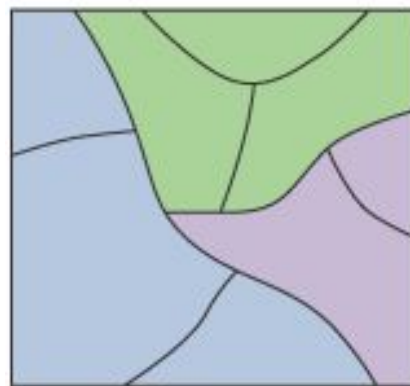
# Disolver



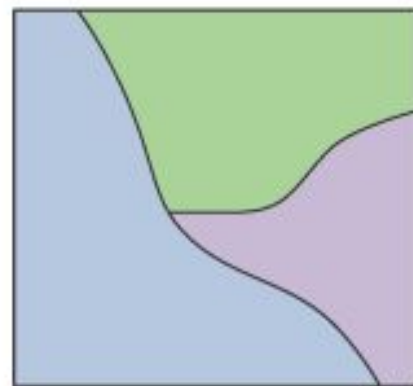
AREA	PERIMETER	POLY#	POLY-ID	CODE
107.09	18.49	2	2	30
85.36	16.36	3	3	10
138.85	17.97	4	4	30
73.09	16.27	5	5	10
144.29	16.46	6	6	30
116.03	16.40	7	7	20
93.74	16.68	8	8	10
158.06	23.01	9	9	20



**INPUT COVERAGE**



**OUTPUT COVERAGE**



# Uniones o Joins

Guardar datos en una tabla única, estilo Excel repite mucha información, en filas y columnas.

- Si se la materia, ya se el profesor y el aula.
- Si se el DNI, ya se el nombre del estudiante.

Materia	DNI	Nombre	Nota	Aula	Profesor
Matemática	1234	Juan	8	101	Gomez
Castellano	4567	Maria	9	202	Perez
Castellano	4567	Maria	10	202	Perez



# Uniones o Joins

Es una forma más eficiente de relacionar datos tomada de las bases de datos SQL

- Ahora memoria
- Es más rápido
- Maneja explícitamente los faltantes de información. Mejor que ordenar una columna y copiar y pegar como en Excel.

# Uniones o Joins

Estudiantes

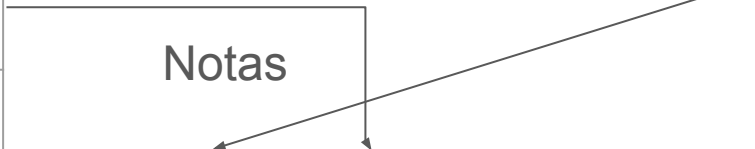
DNI	Nombre
1234	Juan
4567	Maria

Notas

Materia	DNI	Nota
1	1234	8
2	4567	9
2	4567	10

Materias

Codigo	Nombre	Aula	Profesor
1	Matemática	101	Gomez
2	Castellano	202	Perez



# Promedio simple o ponderado

10 %

$n = 2$

80 %

$n = 10$

$$P_s = (10\% + 80\%) / 2 = 10\% * \frac{1}{2} + 80\% * \frac{1}{2} = 45 \%$$

$$P_p = 10\% * \frac{2}{12} + 80\% * \frac{10}{12} = 68.3 \%$$

# Segunda parte: problema de la zonificación

Cómo distribuir información de unidades agregadas a otras unidades que no coinciden?

Ejemplo: circuitos electorales y radios censales.

Si queremos ver si los resultados por distrito electoral coinciden con el nivel educativo promedio del distrito, como pasamos de los radios a los distritos si no se anidan?

Efecto escala

630	651
548	641

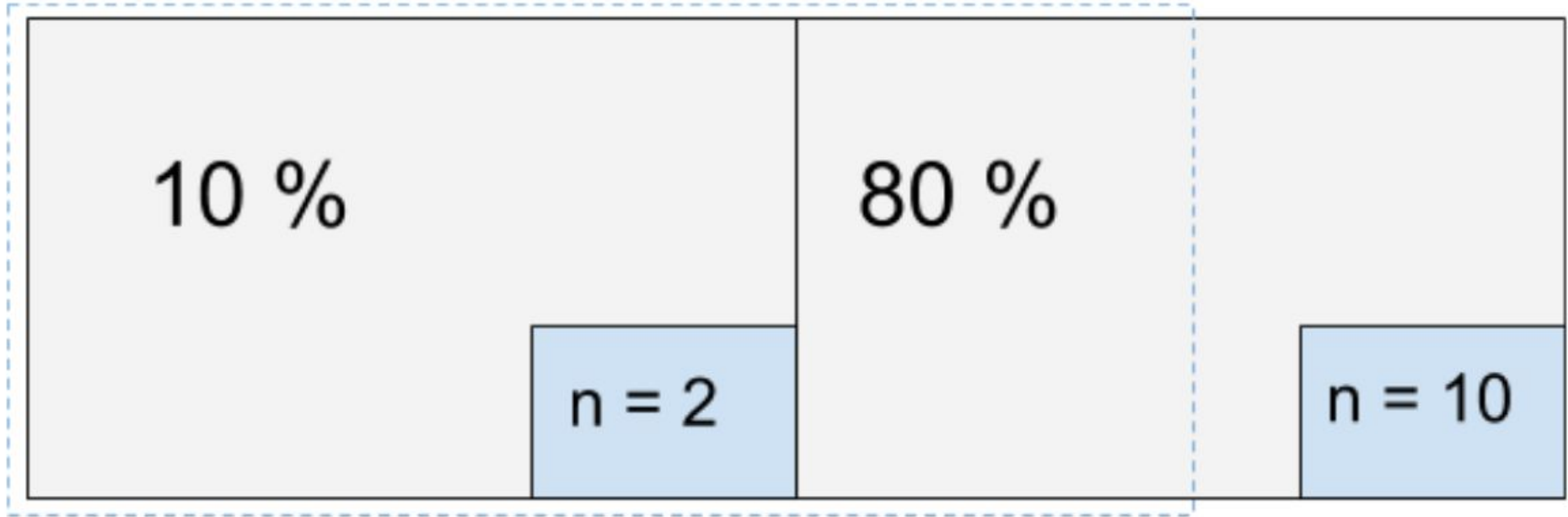
174	162	169	161
150	144	161	160
131	127	162	163
142	148	158	158

Efecto zonificación

968	972
530	

965	
526	979

# Unidades no anidadas



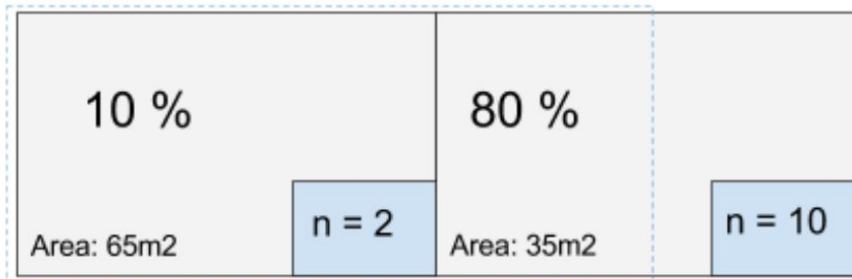
Cómo agregar para la zona macro datos de zonas micro que no coinciden?

Promedio ponderado... ponderado por?

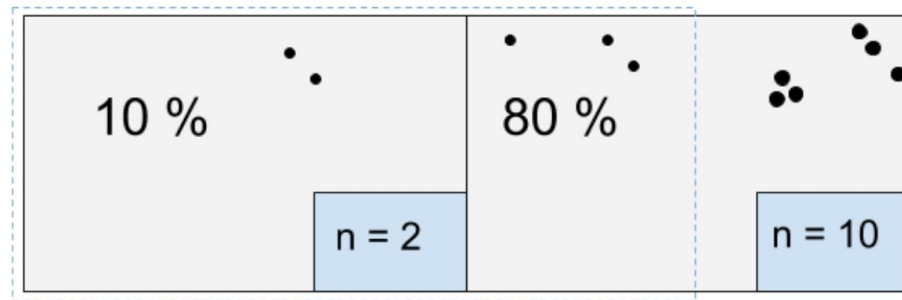


# Promedios ponderados

Area: 100 m<sup>2</sup>



$$P_p = 10\% * 65/100 + 80\% * 35/100 = 34.5 \%$$

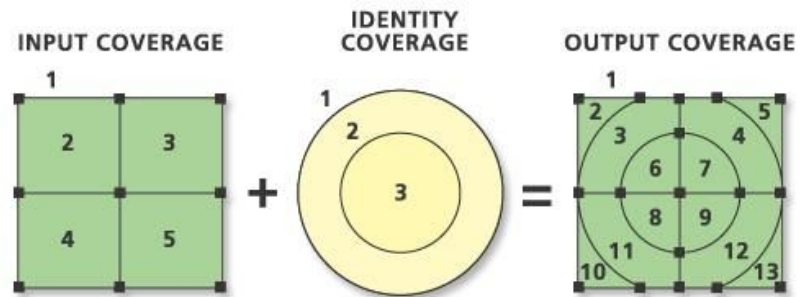


$$P_p = 10\% * \frac{2}{10} + 80\% * \frac{10}{10}$$

Si podemos observar unidades de vivienda utilizando data de catastro pueden servir de proxy para la población, suponiendo que el tamaño del hogar y las viviendas vacías se distribuyan consistentemente. Heroicamente, ignoraremos la falacia ecológica y supondremos que todos los hogares del radio comparten la misma la tasa de la variable bajo análisis.

# Identidad

El proceso de GIS que nos ayudará es el de Identidad: cruza dos geometrías y crea una única con todas las combinaciones posibles en el **espacio** y también en los **datos**.



INPUT COVERAGE

#	ATTRIBUTE
1	
2	A
3	B
4	C
5	D

IDENTITY COVERAGE

#	ATTRIBUTE
1	
2	102
3	103

OUTPUT COVERAGE

#
1
2
3
4
5
6
7
8
9
10
11
12
13

INPUT COVERAGE

#	ATTRIBUTE
1	
2	A
3	B
4	C
5	D
6	A
7	B
8	C
9	D
10	C
11	C
12	D
13	D

IDENTITY COVERAGE

#	ATTRIBUTE
1	
2	102
3	102
4	
5	
6	103
7	103
8	103
9	103
10	
11	102
12	102
13	

# Diferentes pesos según sean proporciones o totales

Conteo

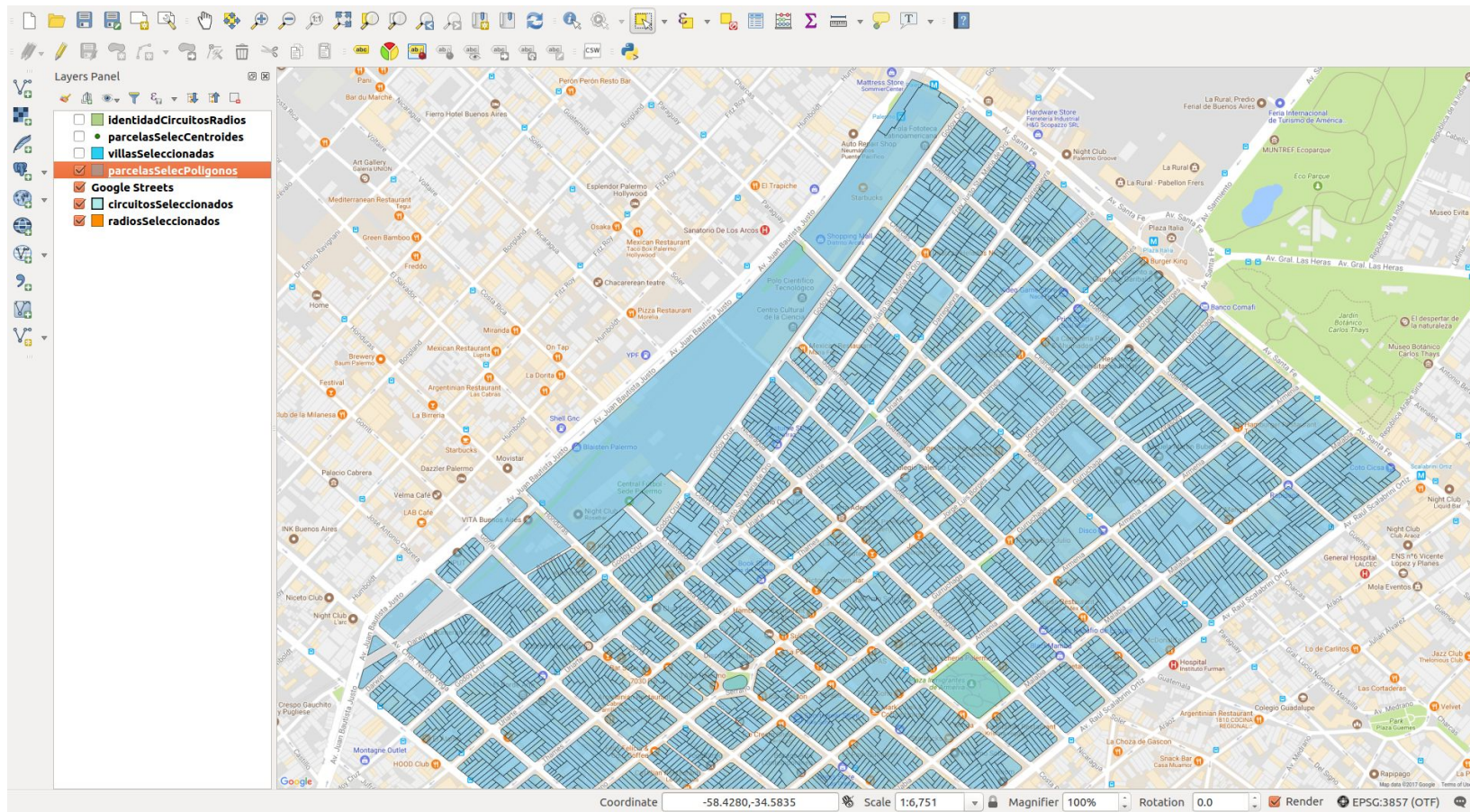
$$v_t = \sum_{i \in t} v_s \frac{h_i}{h_s}$$

Proporciones

$$v_t = \sum_{i \in t} v_s \frac{h_i}{h_t}$$

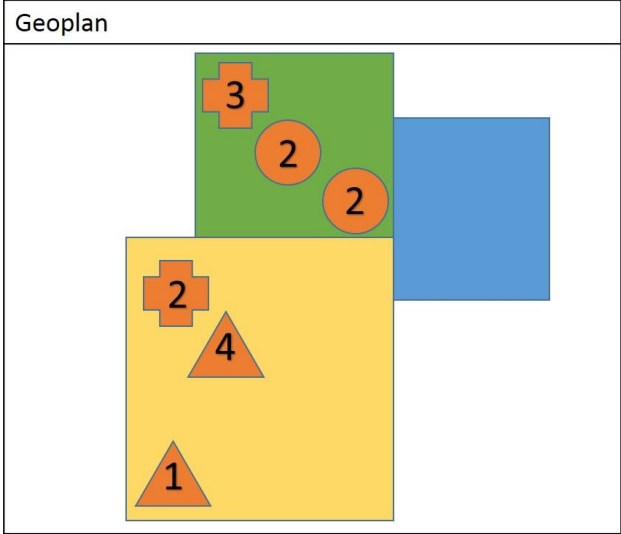


# No siempre las areas estan pobladas



# Uniones espaciales o Spatial Joins

## Mapa



## Tabla




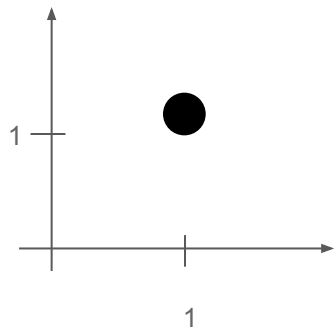
Table 1: Reported Incidents			
Date Reported	Blockage 	Flooding 	Odour 
01/02/03	3		4
02/02/03		2	
03/02/03	2	2	1

Table 2: No. of Incidents per City			
City Names	Blockage Total	Flooding Total	Odour Total
City 1	3	4	
City 2	2		5
City 3			

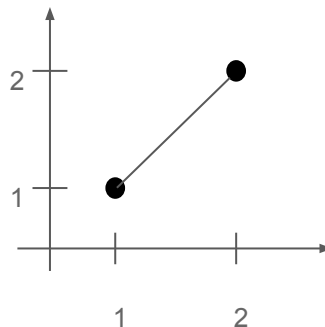
# Coordenadas

Punto



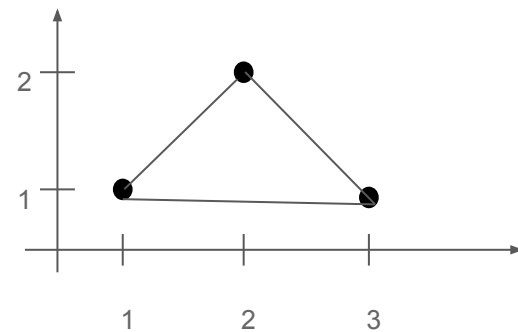
$[(1,1)]$

Linea



$[(1,1),(2,2)]$

Polígono



$[(1,1),(2,2),(3,1),(1,1)]$

# Proyección de Coordenadas

Elementos a considerar al elegir una proyección:

- Zona geográfica
- Variables a analizar
- Unidades de medida

Toda proyección deforma:

- Distancia
- Forma
- Area
- Dirección

# Proyección de Coordenadas

Selected transformation

Argentina - onshore and offshore.  
code 1210

Accuracy 0.2 m (default)  
3 parameters

Method: Geocentric translations (geog2D domain)

Remarks:

Information source: OGP

Revision date: 2008-06-24

Covered area



Center coordinates  
5234771.06 5557249.73

Projected bounds:  
4709004.19 3444524.01  
6263746.79 7574229.97

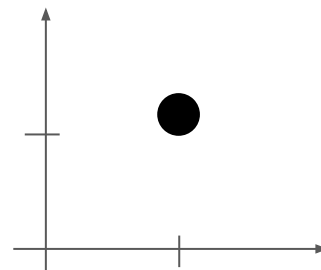
WGS84 bounds:  
-73.59 -58.41  
-52.63 -21.78

Argentina - onshore and offshore.

urn scale topographic mapping and engineering

POSGAR 94  
Argentina 5  
(metros)

2747781.45

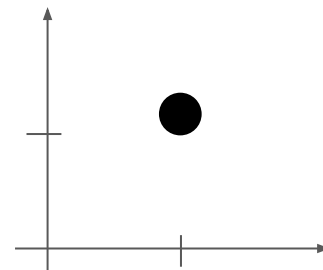


6083958



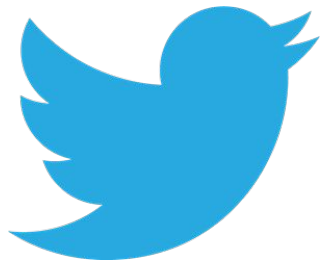
WGS 84  
(lat, long)

-34.61



-58.43

## Contacto



twitter: @lephero



github: /alephcero