

ANALISIS ESPACIAL AVANZADO

Análisis espacial avanzado: clusters espaciales y
hotspots con SIG

Docente: Felipe González

UNTREF

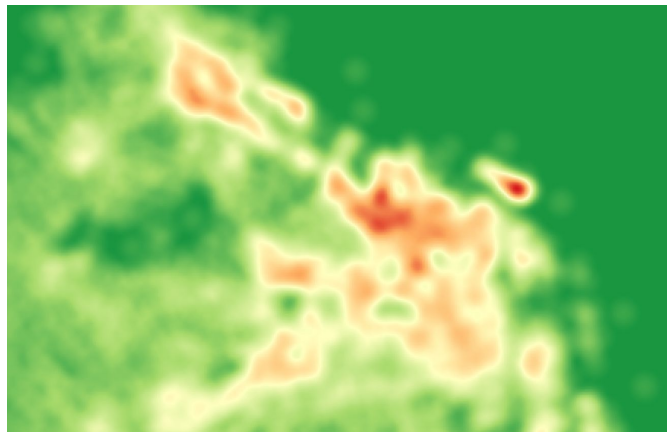
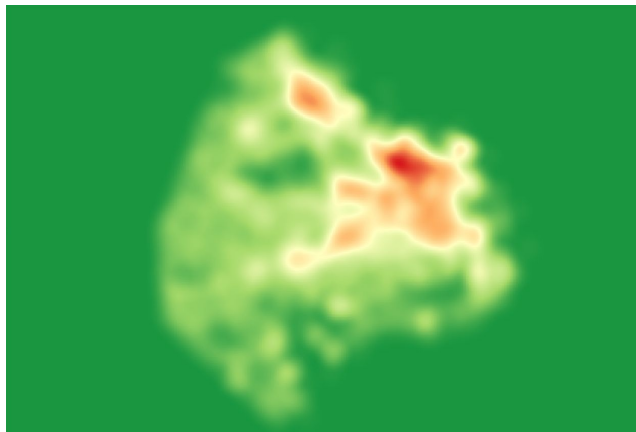
Objetivo general del taller

Poder analizar la distribución en el espacio de una variable pudiendo defender la significancia estadística de los resultados.

Ser crítico de las herramientas

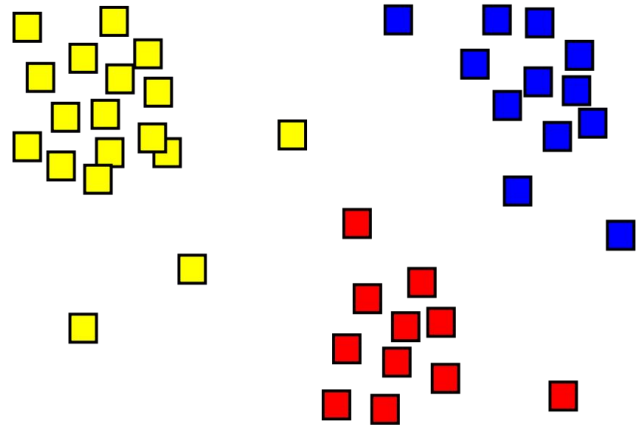
Normalmente los paquetes de software nos ofrecen salidas “sencillas” a expensas de tomar decisiones por nosotros. La propuesta de este taller es comprender esas decisiones y sus influencias en nuestros análisis de un modo crítico.

A qué nos referimos cuando decimos “clusters”, heatmap, hotspots?



Cluster

Esta dimensión refiere al agrupamiento (*clustering*) espacial del grupo x, es decir, al *grado en que las unidades espaciales donde residen los miembros de dicho grupo se encuentran contiguas en el espacio*.



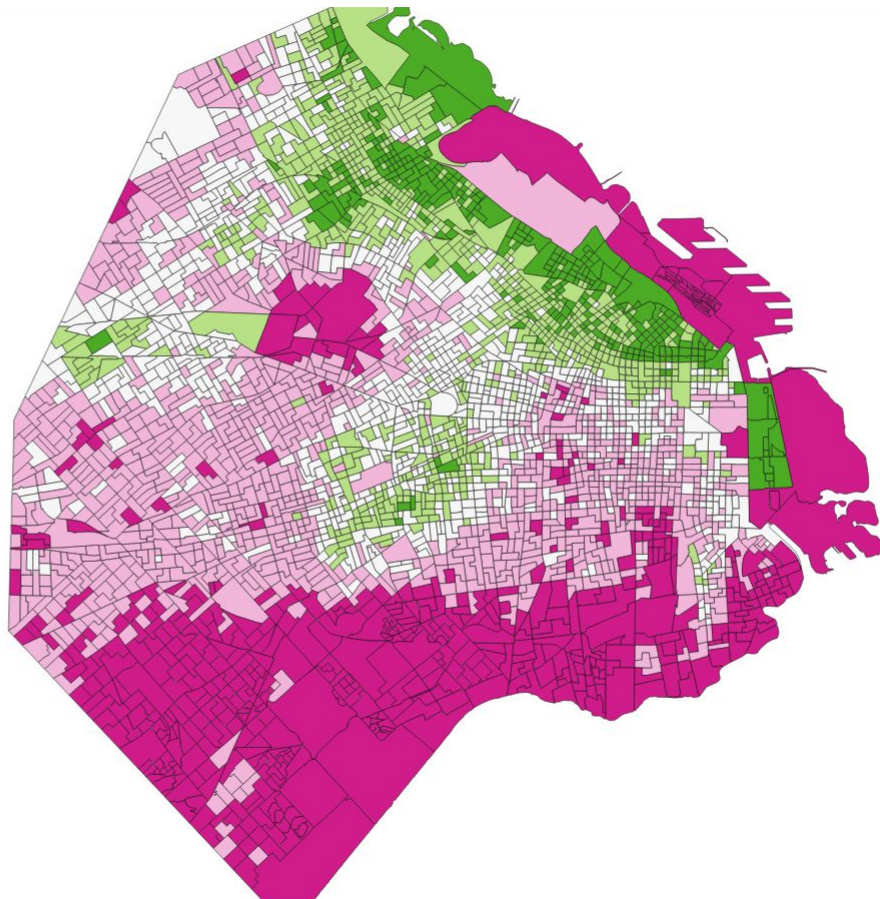
Objetivos particulares

Clase 1

- Introducción al software GIS
- Tipos de datos en software GIS
- Acceso, descarga y procesamiento de datos abiertos con GIS
- Manejo de proyecciones espaciales
- Técnicas y estrategias de visualización
- Visualización sencilla en coropleta

Clase 1

Coropleta sencilla con 5 quintiles
por radio censal



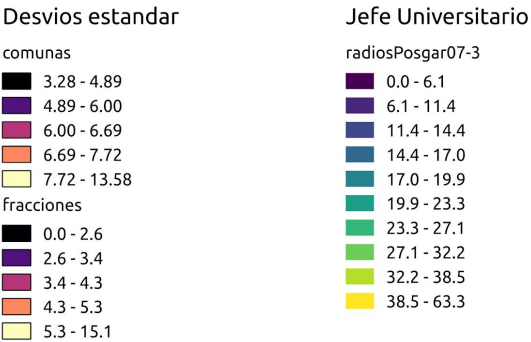
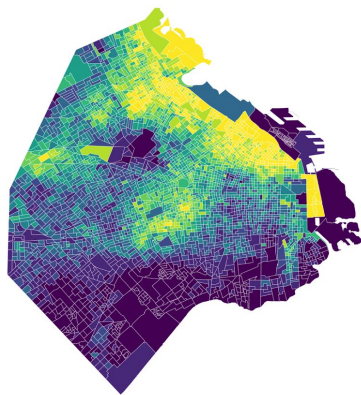
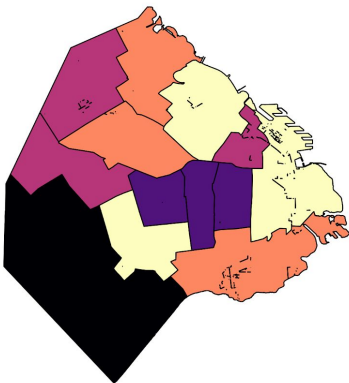
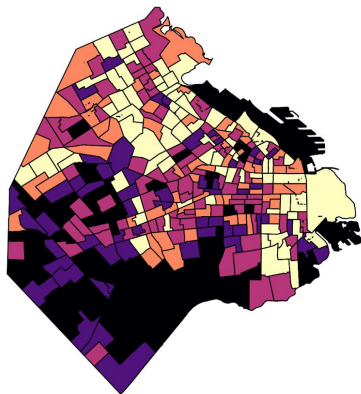
Objetivos particulares

Clase 2

- La dimensión de la escala en el problema de la unidad espacial modificable
- Uniones espaciales
- Agregación de datos por unidades geoestadísticas
- Disolución de unidades geoestadísticas

Clase 2

Distribución de la variable por
radio censal y su desvío
estándar por fracción y
comuna



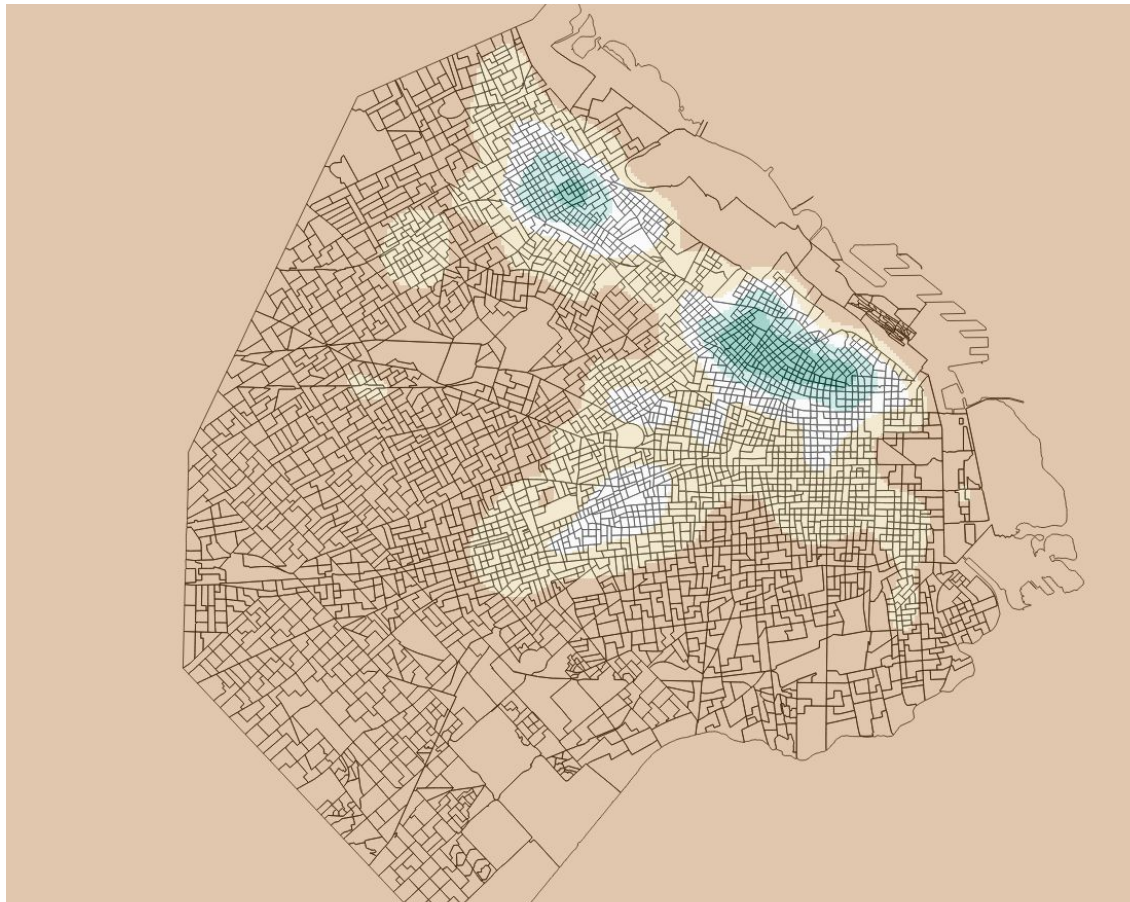
Objetivos particulares

Clase 3

- La dimensión de la zonificación en el problema de la unidad espacial modificable
- Creación de mallas de celdas uniformes y transferencia de datos
- Mapas de calor con densidades kernel

Clase 3

Mapa de calor utilizando
densidades kernel



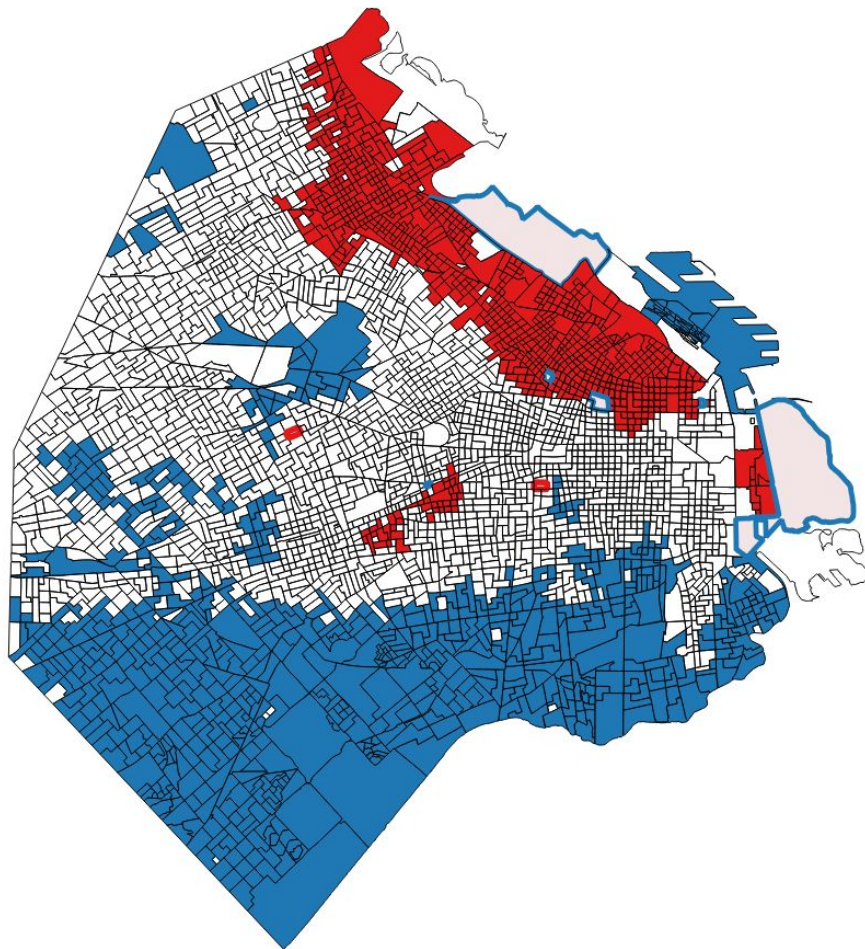
Objetivos generales

Clase 4

- Comprender el estadístico global I de Moran
- Comprender la versión local del estadístico I de Moran
- Comprender la lectura de Hots spots, Cold Spots
- Comprender la lectura de

Clase 4

I de Moran Local



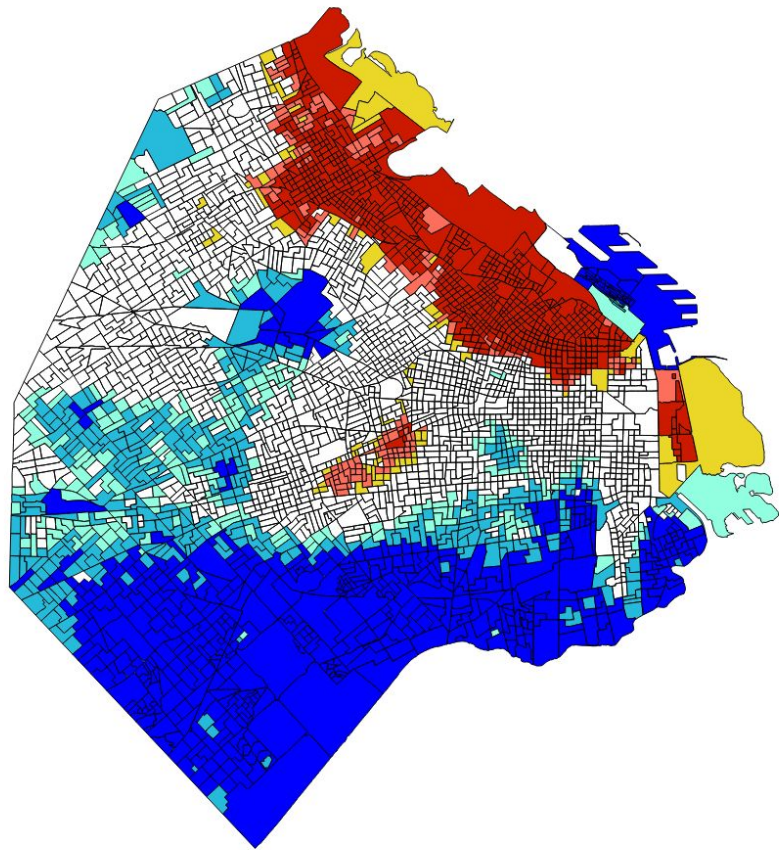
Objetivos particulares

Clase 5

- Comprensión del indicador G^* - Getis Ord
- Comprensión del significado del test de hipótesis utilizado

Clase 5

G* - Getis Ord



QGIS



Software Libre para sistemas de información
georeferenciada SIG

Los datos

Tipos de datos utilizados en software SIG

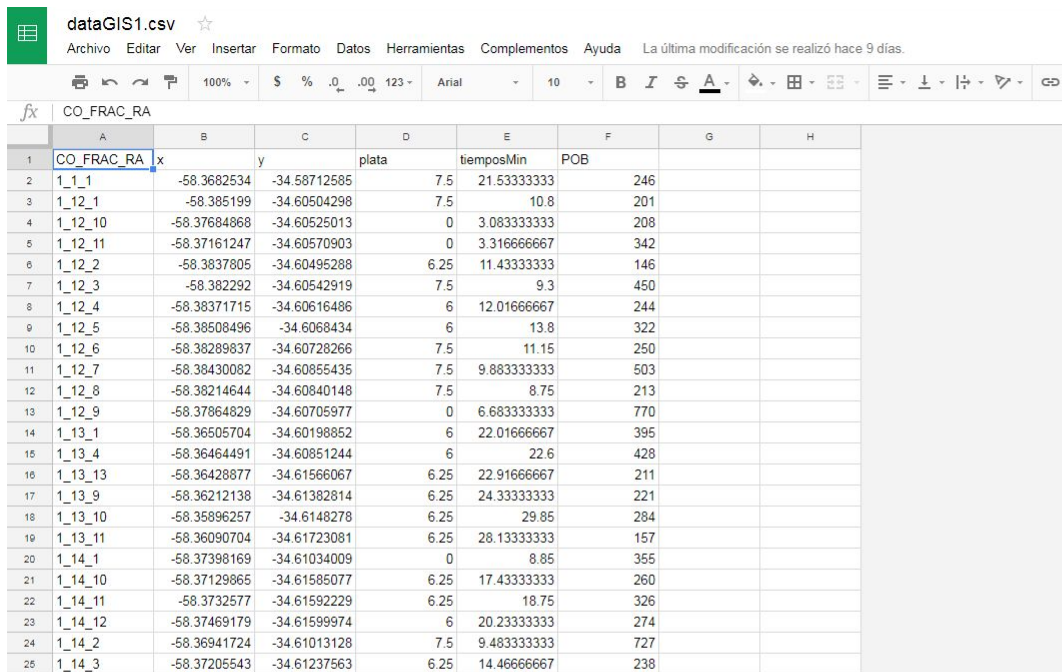
Tipos de datos

- Tablas de datos (xls, csv, DBF, etc)
- Raster
- Vectoriales (Shapefiles, GeoJSON, etc)

Tablas de datos

Estructurada en filas (casos) y columnas (variables)

Normalmente identificadas con un ID único



The screenshot shows a spreadsheet application window titled "dataGIS1.csv". The interface includes a menu bar with options like Archivo, Editar, Ver, Insertar, Formato, Datos, Herramientas, Complementos, and Ayuda. Below the menu is a toolbar with various icons for editing and formatting. The spreadsheet itself has a grid with columns labeled A through H. The first column (A) contains a unique identifier for each row, such as "1_1_1", "1_12_1", "1_12_10", etc. The second column (B) contains numerical values, likely coordinates. The third column (C) contains more numerical values. The fourth column (D) is labeled "plata" and contains numerical values. The fifth column (E) is labeled "tiemposMin" and contains numerical values. The sixth column (F) is labeled "POB" and contains numerical values. The seventh column (G) and eighth column (H) are empty. The data is organized into rows, with each row representing a specific case or location.

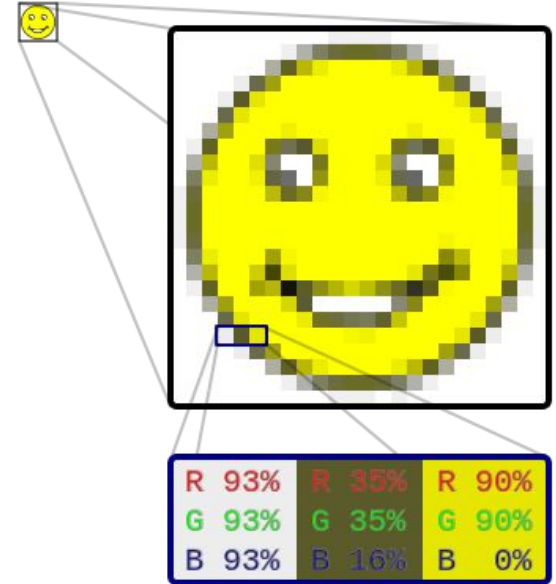
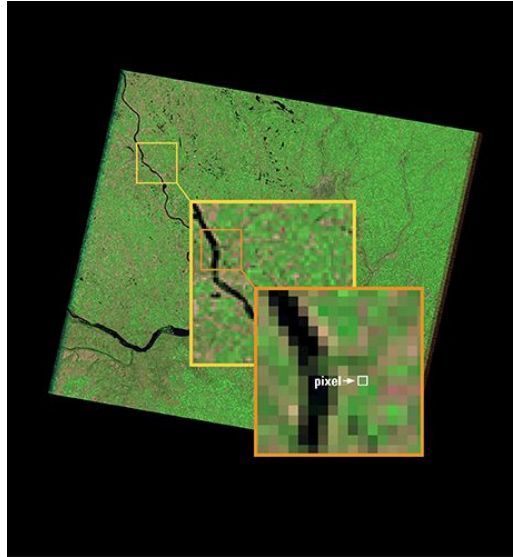
	A	B	C	D	E	F	G	H
1	CO_FRAC_RA	x	y	plata	tiemposMin	POB		
2	1_1_1	-58.3682534	-34.58712585	7.5	21.53333333	246		
3	1_12_1	-58.385199	-34.60504298	7.5	10.8	201		
4	1_12_10	-58.37684868	-34.60525013	0	3.083333333	208		
5	1_12_11	-58.37161247	-34.60570903	0	3.316666667	342		
6	1_12_2	-58.3837805	-34.60495288	6.25	11.43333333	146		
7	1_12_3	-58.382292	-34.60542919	7.5	9.3	450		
8	1_12_4	-58.38371715	-34.60616486	6	12.01666667	244		
9	1_12_5	-58.38508496	-34.6068434	6	13.8	322		
10	1_12_6	-58.38289837	-34.60728266	7.5	11.15	250		
11	1_12_7	-58.38430082	-34.60855435	7.5	9.883333333	503		
12	1_12_8	-58.38214644	-34.60840148	7.5	8.75	213		
13	1_12_9	-58.37964829	-34.60705977	0	6.683333333	770		
14	1_13_1	-58.36505704	-34.60198852	6	22.01666667	395		
15	1_13_4	-58.36464491	-34.60851244	6	22.6	428		
16	1_13_13	-58.36428877	-34.61566067	6.25	22.91666667	211		
17	1_13_9	-58.36212138	-34.61382814	6.25	24.33333333	221		
18	1_13_10	-58.35896257	-34.6148278	6.25	29.85	284		
19	1_13_11	-58.36090704	-34.61723081	6.25	28.13333333	157		
20	1_14_1	-58.37398169	-34.61034009	0	8.85	355		
21	1_14_10	-58.37129865	-34.61585077	6.25	17.43333333	260		
22	1_14_11	-58.3732577	-34.61592229	6.25	18.75	326		
23	1_14_12	-58.37469179	-34.61599974	6	20.23333333	274		
24	1_14_2	-58.36941724	-34.61013128	7.5	9.483333333	727		
25	1_14_3	-58.37205543	-34.61237563	6.25	14.46666667	238		

Raster

Matrices de datos de n (filas o alto) \times m (columnas o ancho) \times l (dimensiones)

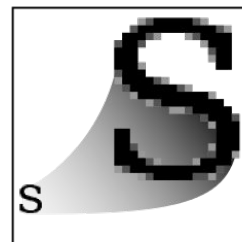
Una imagen en B&N tiene 1 dimensión, en color tiene 3 dimensiones (cantidades de Rojo, Verde y Azul)

Pero puede tener L dimensiones



Vectores

Descripciones geométricas de objetos que pueden renderizarse en cualquier tamaño de pantalla deseado sin problemas



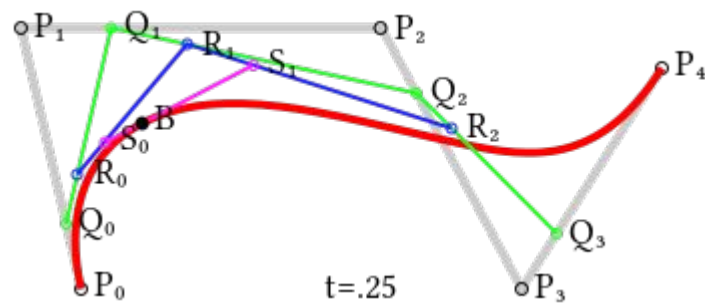
Raster

.jpeg .gif .png



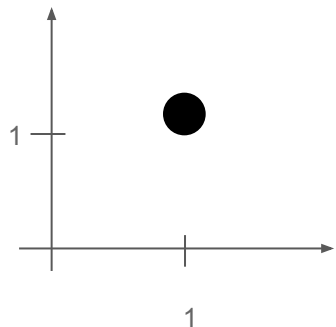
Vector

.svg



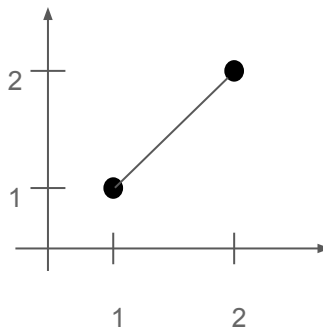
Tipos de datos espaciales vectoriales

Punto



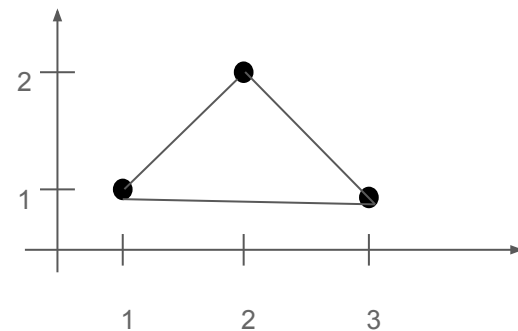
$[(1,1)]$

Línea



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

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

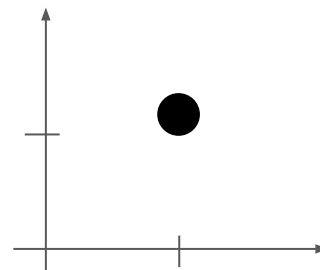
WGS84 bounds:
-73.59 -58.41
-52.63 -21.78

Argentina - onshore and offshore.

urn scale topographic mapping and engineering

POSGAR 94
Argentina 5
(metros)

2747781.45

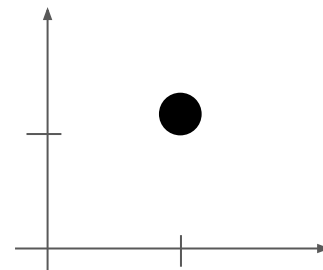


6083958



WGS 84
(lat, long)

-34.61



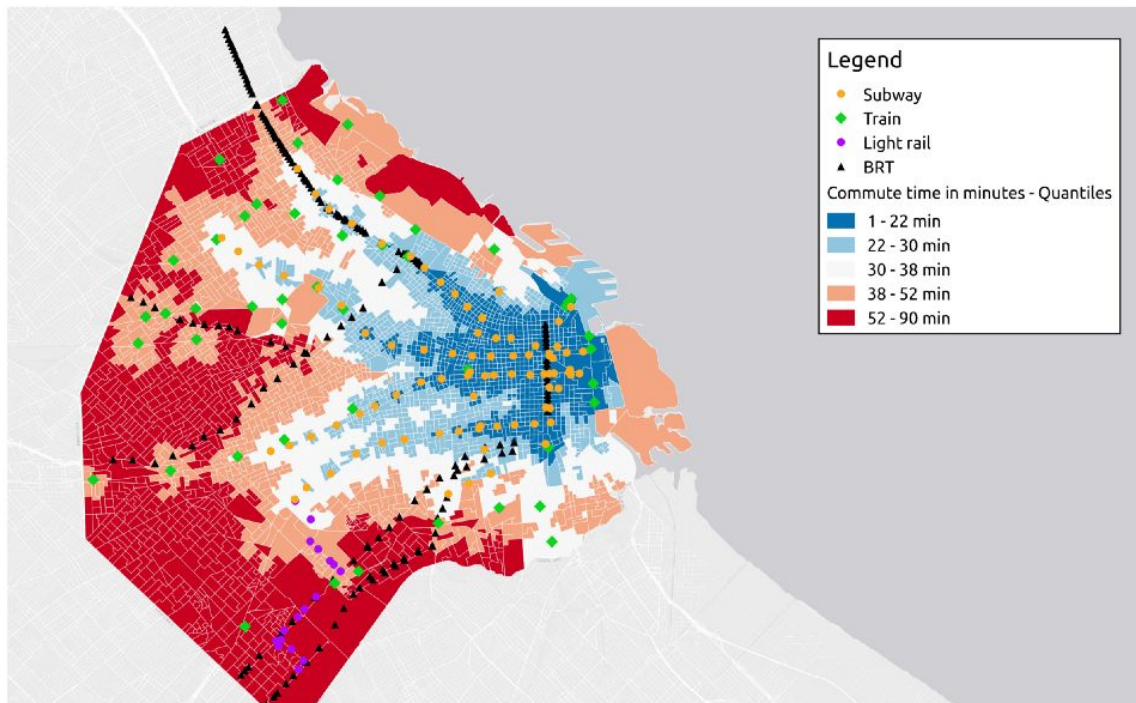
-58.43

¿Por qué usar GIS?

- El espacio es una gran forma de estructurar datos no estructurados
- Muchas veces, el espacio, la posición de las cosas, las relaciones entre ellas, son una variable a considerar en el análisis (en movilidad es fundamental)
- La primera ley de Tobler: *“todo está relacionado con todo lo demás, pero cosas cercanas están más relacionadas que cosas distantes”*

Visualización

Commute trip time in minutes for Buenos Aires City by public transit



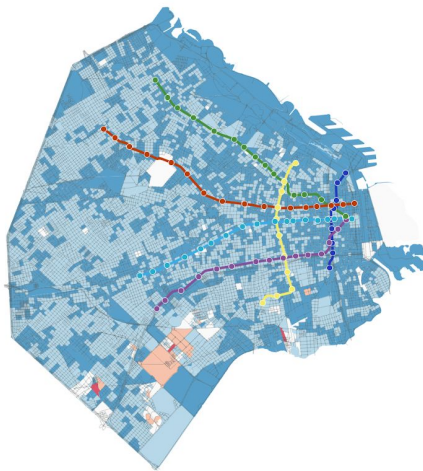
Subway and train have the highest impact in commute time to downtown. Light rail also has some influence, especially in the first part of the route. BRT doesn't appear to influence that much. Maybe the Cabildo Avenue's BRT in the north of the city has some influence. The San Martin Avenue BRT competes with the train and the South BRT doesn't appear to have much of an impact on commute trip

Radios censales (polígonos) coloreados en función del tiempo de viaje en minutos al centro (Quantiles).

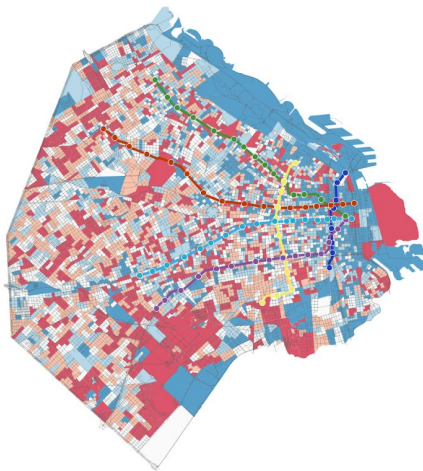
Qué colores (paleta) y cómo (quantiles, jenks, intervalos iguales) no es una decisión ingenua. Implica una estrategia de visualización.
CONSTRUYE SENTIDO

Visualización

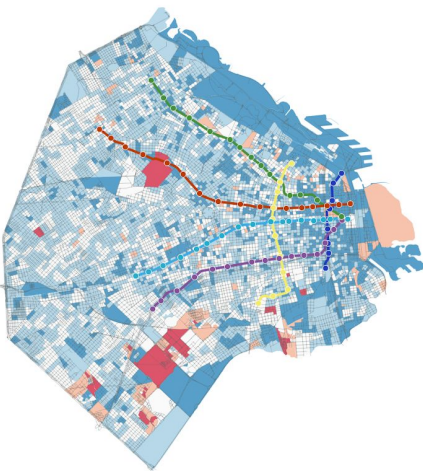
iguales



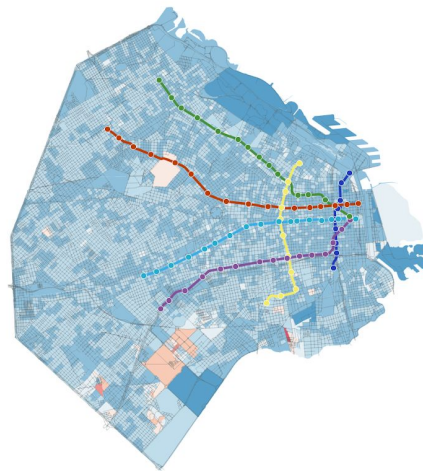
quintiles



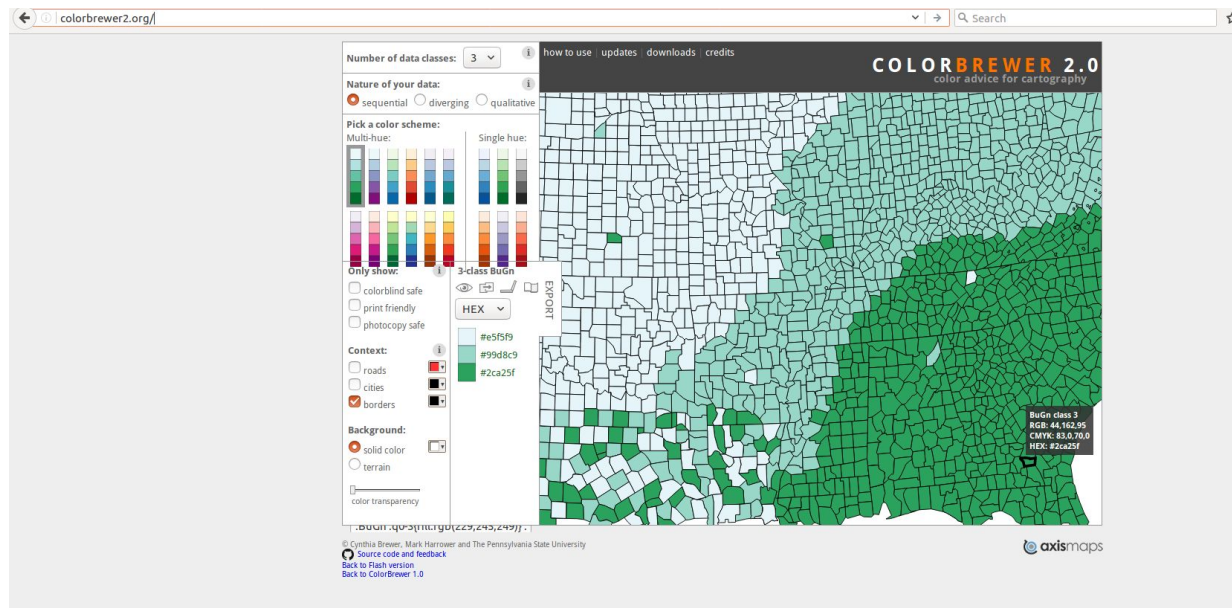
jenks



desvios



Colorbrewer2.org



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

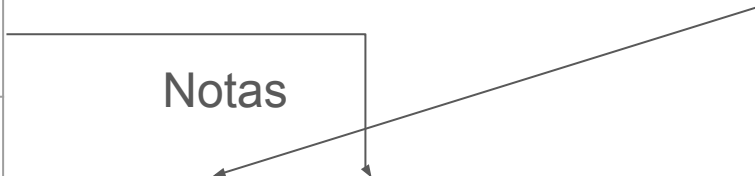
DNI	Nombre
1234	Juan
4567	Maria

Notas

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

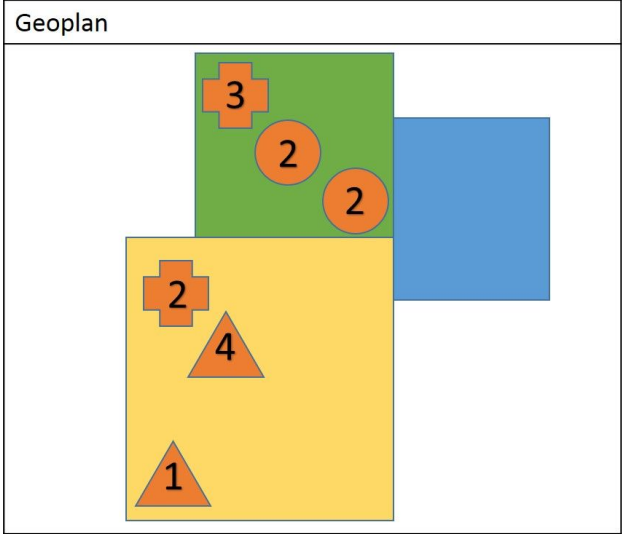
Materias

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



Spatial Joins

Mapa



Tabla




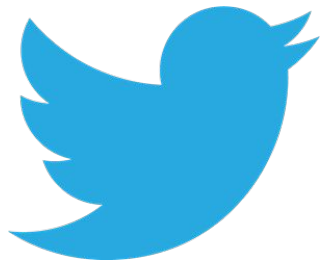
Table 1: Reported Incidents			
Date Reported	Blockage 	Flooding 	Odour 
01/02/03	3		4
02/02/03		2	
03/02/03	2	2	1

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			

Contacto



twitter: @lephero



github: /alephcero