Spatial Analysis – Introductory Concepts and Overview

Prabhakar Alok Verma
Scientist/ Engineer – 'SD'
Geoinformatics Department
prabhakar@iirs.gov.in

Learning Objective

- To understand concepts of spatial data retrieval, classification & measurement operations

Learning Outcome

At the end of the lecture you should be able to answer the following…..

- What is spatial data analysis?
- Importance of spatial data analysis
- Components of spatial data analysis
- Types of spatial data operations
- Steps to perform spatial data analysis
- Tools to perform spatial data analysis

Brain Storming

- Distance of nearest hospital?
- Route to the nearest hospital?
- Nearest school which is 100 meters away from major road?
- How to make landslide hazard zones?
- Suitable places to establish a textile industry?
- *Find out the suitable sites to construct a High School in Dehradun City based on the following criteria.*
  - At a distance of more than 1.5 Kilometres from existing schools
  - At a distance between 500 and 1000 metres of the main roads (to minimize traffic noise pollution, but still to have proper access)
  - Not in use by
    - Businesses (business areas)
    - Forest (forest areas)

What is spatial data analysis?

- Spatial analysis is the process of extracting information from spatial data, based on requirement
- This analysis could be:
  - Performing spatial queries
  - Network analysis
  - Creation of surface from point data
  - Making predictions

Importance of spatial data analysis

- To get meaningful information from spatial data

Components of spatial data analysis

- Spatial data
- Computer with GIS capabilities
- Requirement
- Algorithm/ tool able to perform desired analysis
- Visualization

Types of spatial operations
(Analytical Functions of GIS)

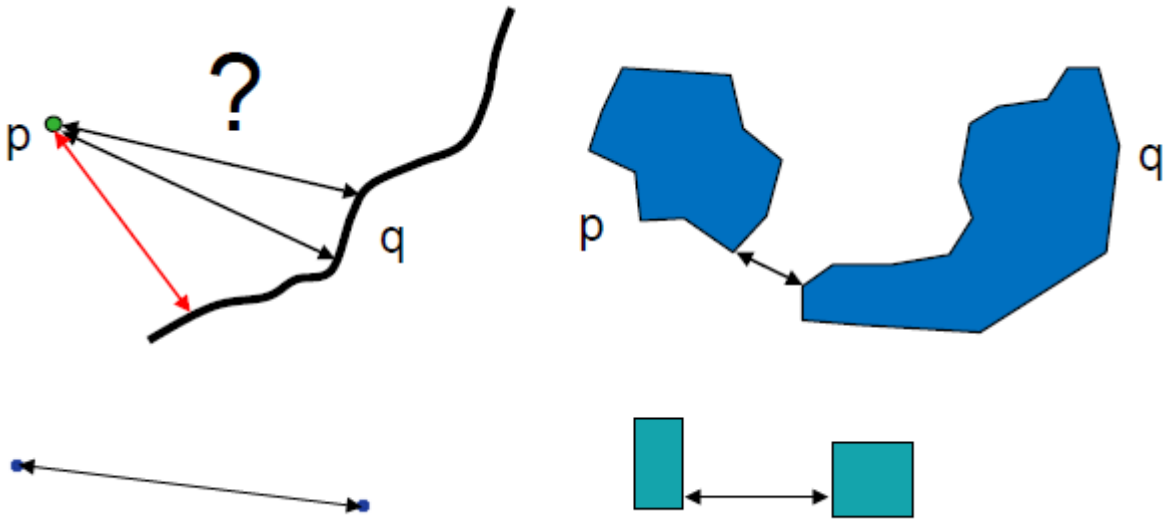
- Classification of functions
  - Measurements, retrieval and classification
  - Overlay functions
  - Neighborhood functions
  - Connectivity functions

Measurement

- These functions allow geometry based calculation like distances, length, area, perimeter, centroid (vector, raster)
- E.g.
  - Find average distance between points
  - Find population density
- Geometric Measurements only

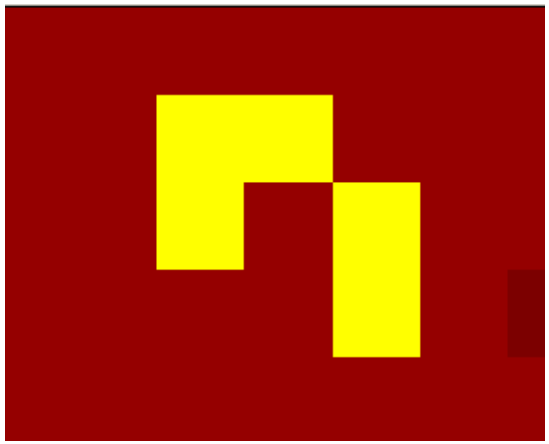
Continued…

- Distance between two points → Pythagorean distance function
- If one or both features are not a point →*minimal distance*



Measurement - Raster

- Raster measurements include: location, distance and area size
- Location of an individual cell → derived from anchor point and resolution
- Area size → number of cells * cell size
- Distance → standard distance function applied to the locations of their mid-points



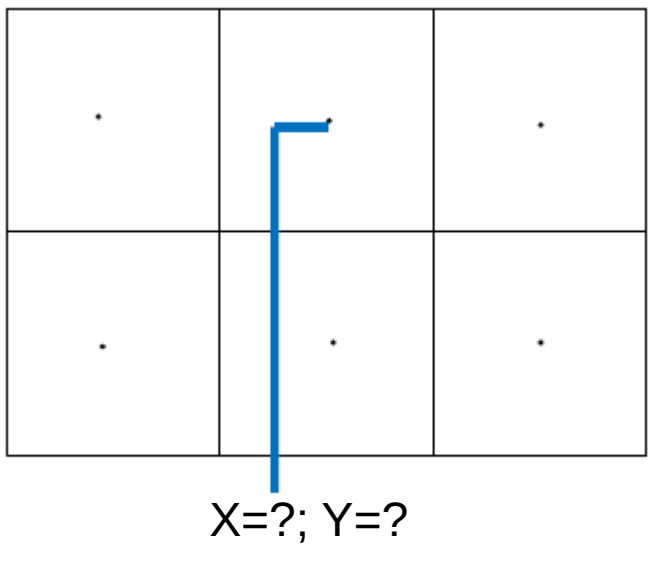
Cell size: 30 m X 30 m

900 * 5 = 4500 m^2

Continued…

- Location of individual cell is derived from anchor point and resolution
- Cell's location is defined by its midpoint

X=1554
Y=203538



Note: Resolution=20x20 m

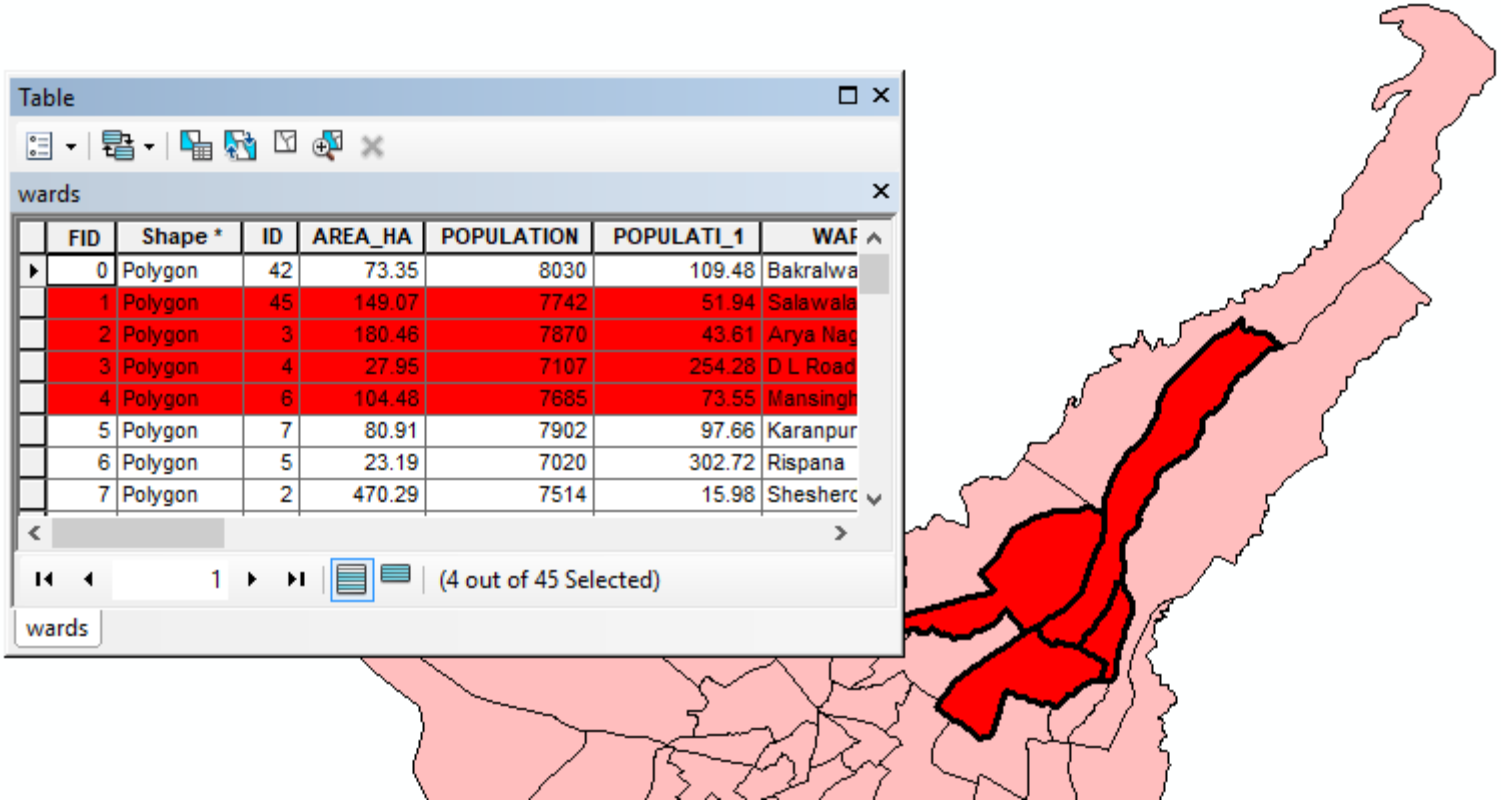
Retrieval

- Facilitates searching of the data.
- Can be based on:
  - Attribute values
  - Location
- E.g. find all the cities having population more than 0.1 million
- E.g. find length of border of India

Spatial selection queries

- Interactive
- Spatial Selection by Attribute conditions
  - Relational operators
  - Logical operators
  - Combining attribute conditions
- Spatial selection using topological relationships
  - Selecting features that are inside selection objects
  - Selecting features that intersect
  - Selecting features adjacent to selection objects
  - Selecting features based on their distance

Interactive

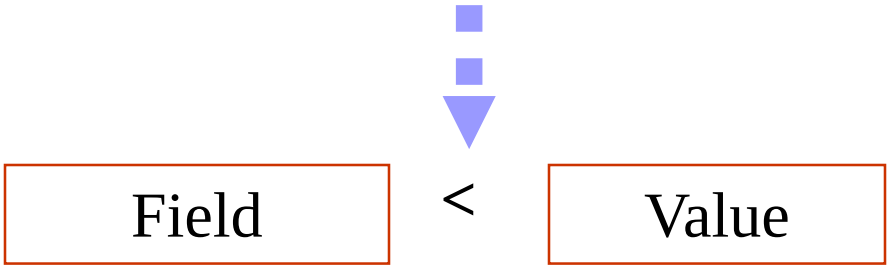


Spatial Selection by Attribute conditions

- Define a **selection condition** on the features attributes in a query language, such as SQL.
- Display the result both on the map and in the attribute table.

Syntax – Relational operators

< (less than)
= (equals)
<= (less or equal than)
> (greater than)
>= (greater than or equal)
<> (does not equal)



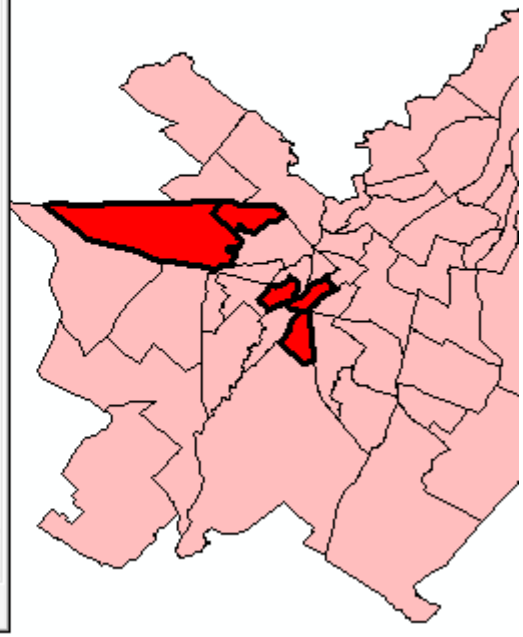
Syntax - Relational operator

Population < 7000

	FID	Shape *	ID	AREA_HA	POPULATIO	POPULATI_1	WAI
	12	Polygon	36	25.48	7654	300.39	Dhama
	13	Polygon	35	23.92	7241	302.72	Lunya I
	14	Polygon	31	20.97	6907	329.38	Jhanda
	15	Polygon	33	22.83	8091	354.4	Khudbu
	16	Polygon	32	19.77	6883	348.15	Inderes
	17	Polygon	24	77.19	7777	100.75	Idgah
	18	Polygon	25	27.22	7375	270.94	Shivaji
	19	Polygon	28	112.09	7393	65.96	Patel N
	20	Polygon	20	78	7302	93.62	Ballupu
	21	Polygon	22	181.7	7299	40.17	Rajend
	22	Polygon	23	41.46	6935	167.27	Devsur
	23	Polygon	26	30.5	7073	231.9	Gandhi
	24	Polygon	27	40.48	7208	178.06	Laxmar
	25	Polygon	29	40.34	7384	183.04	Lakkhi

1 (5 out of 45 Selected)

wards



Syntax – logical operators

AND	(returns true if both expressions a and b are true)
OR	(returns true if one or both of the expressions a and b is true)
NOT	(returns true if expression is false)

Area < 400,000	(atomic condition)
(Area < 400,000) AND (landuse =Forest)	(composite cond.)
(Area < 400,000) OR (landuse = Forest)	(composite cond.)
NOT (landuse = Forest)	(negate condition)

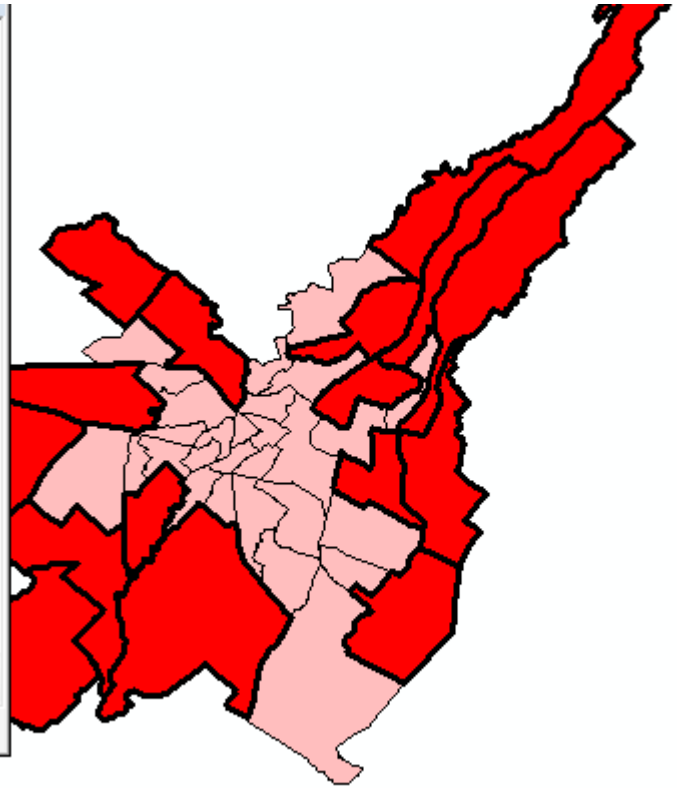
Combining attribute conditions

(Area > 100) AND (Population < 8000)

FID	Shape *	ID	AREA	POPULATION	POPULATI_1	WA
31	Polygon	12	51.98	7932	152.6	Dhara
32	Polygon	39	86.98	8249	94.84	Race
33	Polygon	10	145.01	8030	55.38	Dalar
34	Polygon	14	651.99	7630	11.7	Bhan
35	Polygon	11	270.08	7030	26.03	Rajiv
36	Polygon	8	261.04	7239	27.73	Adho
37	Polygon	13	481.5	8252	17.14	Ajabp
38	Polygon	15	383.32	7900	20.61	Majra
39	Polygon	19	236.97	6864	28.97	Maha
40	Polygon	18	263.88	7881	29.87	Basa
41	Polygon	17	202.41	8068	39.86	Kanw
42	Polygon	16	347.37	7347	21.15	Niran
43	Polygon	1	475.17	7341	15.45	Rajpu
44	Polygon	21	194.31	7764	39.96	Kaula

0 (16 out of 45 Selected)

wards



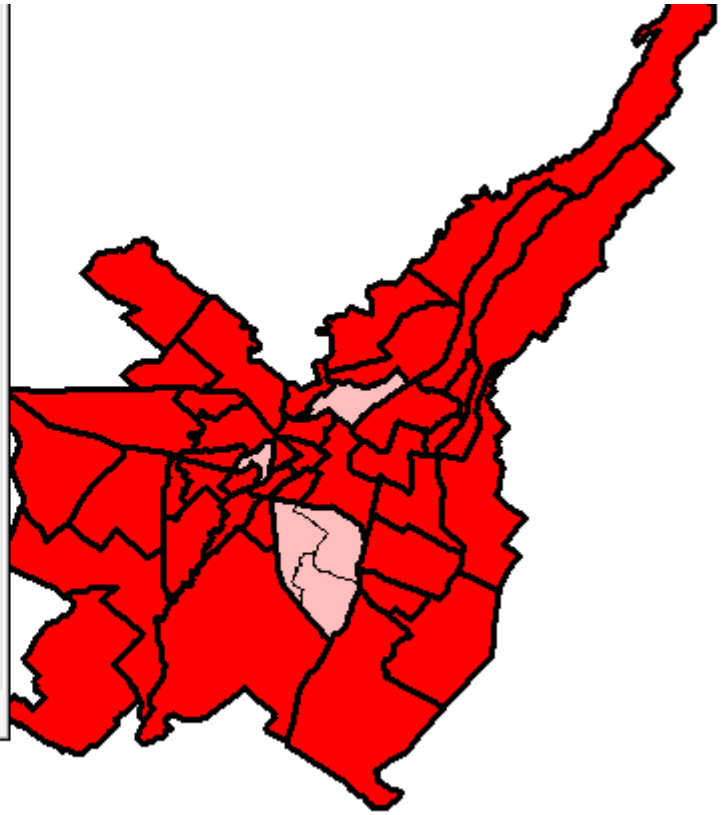
Combining attribute conditions

(Area > 100) OR (Population < 8000)

	FID	Shape *	ID	AREA	POPULATION	POPULATI_1	WA
	31	Polygon	12	51.98	7932	152.6	Dhara
	32	Polygon	39	86.98	8249	94.84	Race
	33	Polygon	10	145.01	8030	55.38	Dalar
	34	Polygon	14	651.99	7630	11.7	Bhan
	35	Polygon	11	270.08	7030	26.03	Rajiv
	36	Polygon	8	261.04	7239	27.73	Adho
	37	Polygon	13	481.5	8252	17.14	Ajabp
	38	Polygon	15	383.32	7900	20.61	Majra
▸	39	Polygon	19	236.97	6864	28.97	Maha
	40	Polygon	18	263.88	7881	29.87	Basa
	41	Polygon	17	202.41	8068	39.86	Kanw
	42	Polygon	16	347.37	7347	21.15	Niran
	43	Polygon	1	475.17	7341	15.45	Rajpu
	44	Polygon	21	194.31	7764	39.96	Kaula

40 (40 out of 45 Selected)

wards



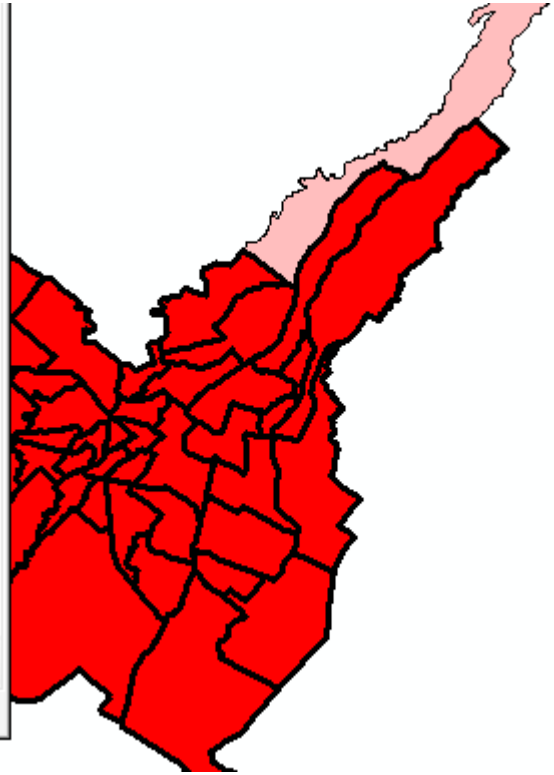
Negate Condition

NOT ‘Ward Name’ = ‘Rajpur’

WARD NAME	POPULATI_2	DE
Dharampur	Moderate Population Density	141-210 Pers
Racecourse (N)	Low Population Density	71-140 Persc
Dalanwala (S)	Very Low Population Density	Up to 70 Pe
Bhandaribagh	Very Low Population Density	Up to 70 Pe
Rajiv Nagar	Very Low Population Density	Up to 70 Pe
Adhoiwala	Very Low Population Density	Up to 70 Pe
Ajabpur	Very Low Population Density	Up to 70 Pe
Majra	Very Low Population Density	Up to 70 Pe
Maharanibagh	Very Low Population Density	Up to 70 Pe
Basant Vihar	Very Low Population Density	Up to 70 Pe
Kanwali	Very Low Population Density	Up to 70 Pe
Niranjanpur	Very Low Population Density	Up to 70 Pe
Rajpur	Very Low Population Density	Up to 70 Pe
Kaulagarh	Very Low Population Density	Up to 70 Pe

0 (44 out of 45 Selected)

wards



Searching for NULL values

‘Ward Name’ is NULL

‘Ward Name’ is NOT NULL

Like Operator

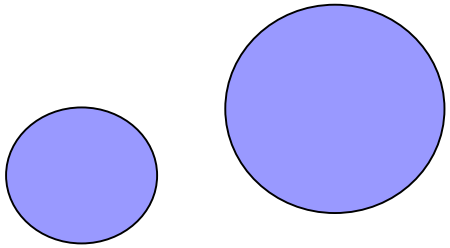
- 'Ward Name' LIKE 'Ra%'
- 'Ward Name' LIKE '%ra'
- 'Ward Name' LIKE '%ra%'

Spatial Relationships

- Disjoint
- Meet
- Equal
- Inside
- Covered by
- Contains
- Covers
- Overlap

Disjoint

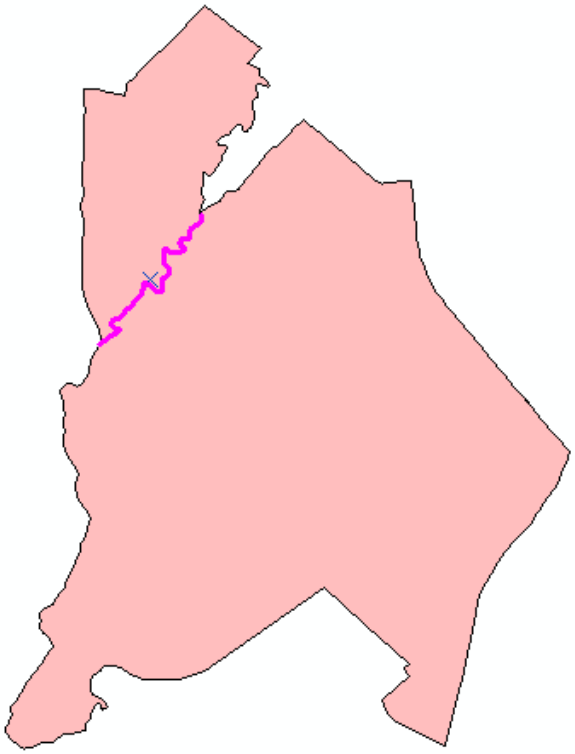
Geometries neither touch nor overlaps



Meet

- Geometries touch at the boundary but not from the interior

- Poly-poly
- Line-line
- Poly-line
- Poly-point
- Line-point



Questions

- Which of the following is not spatial analysis?
  - Network analysis
  - TIN Creation
  - Spatial Query
  - Attribute Query
- Which of the following is/are component/s of spatial data analysis?
  - Toposheet
  - Selection criteria
  - Visualisation
  - All of the above

Questions

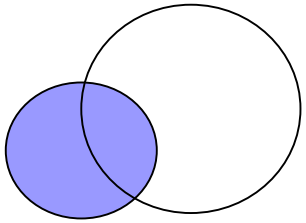
- Distance between two cells of a raster is?
  - Minimum distance between those cells
  - Maximum distance between the cells
  - Distance between their mid points
  - None of the above
- Which of the following wards will be selected by 'Ward Name' LIKE 'Ra%'
  - Rajpur
  - rajpur
  - kaRapur
  - Pura

Questions

- Which of the following is correct about disjoint relationship?
  - Geometries neither touch nor overlaps
  - Geometries touch at the boundary from outside
  - Geometries are completely coincident
  - Intersecting boundaries
- Which of the following is correct about meet relationship?
  - Geometries neither touch nor overlaps
  - Geometries touch at the boundary from outside
  - Geometries are completely coincident
  - Intersecting boundaries

Overlap

- Intersecting boundaries



- Poly-poly
- Poly-line
- Line-line

Equal

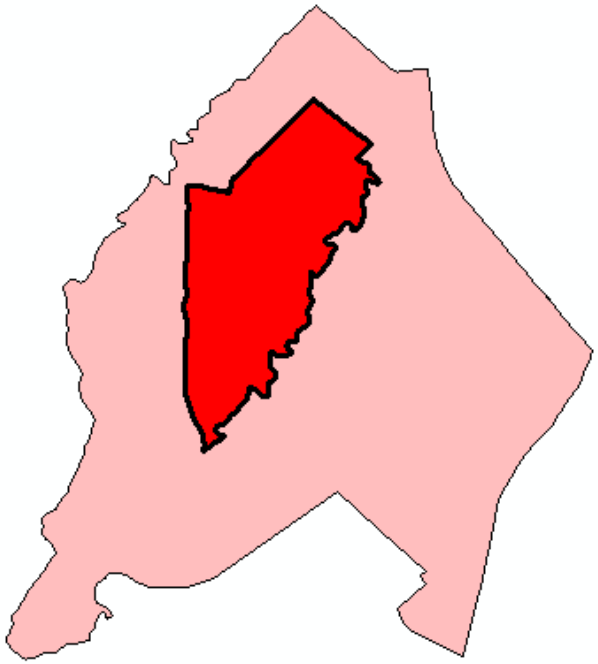
- Geometries are completely coincident.

- Poly-poly
- Line-line
- Point-point

Inside

- Object B is inside Object A (Object B will be selected) e.g. select all the schools inside a city.

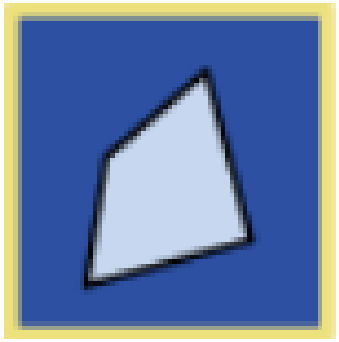
- Poly-poly
- Poly-line
- Poly-point



Contains

- Object A completely contains object B (object A will be selected)
- Select the city which contains IIRS

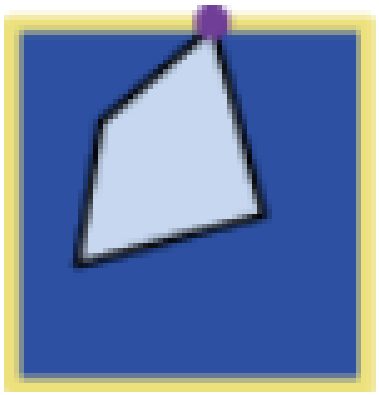
- Poly-poly
- Poly-line
- Poly-point
- Line-point



Covers

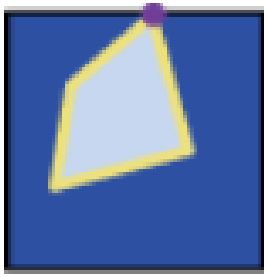
- Object A covers B (object A will be selected)
- E.g. select the city which covers IIRS
- The interior of an object is completely inside the other object and the boundaries intersect.

- Poly-poly
- Poly-line

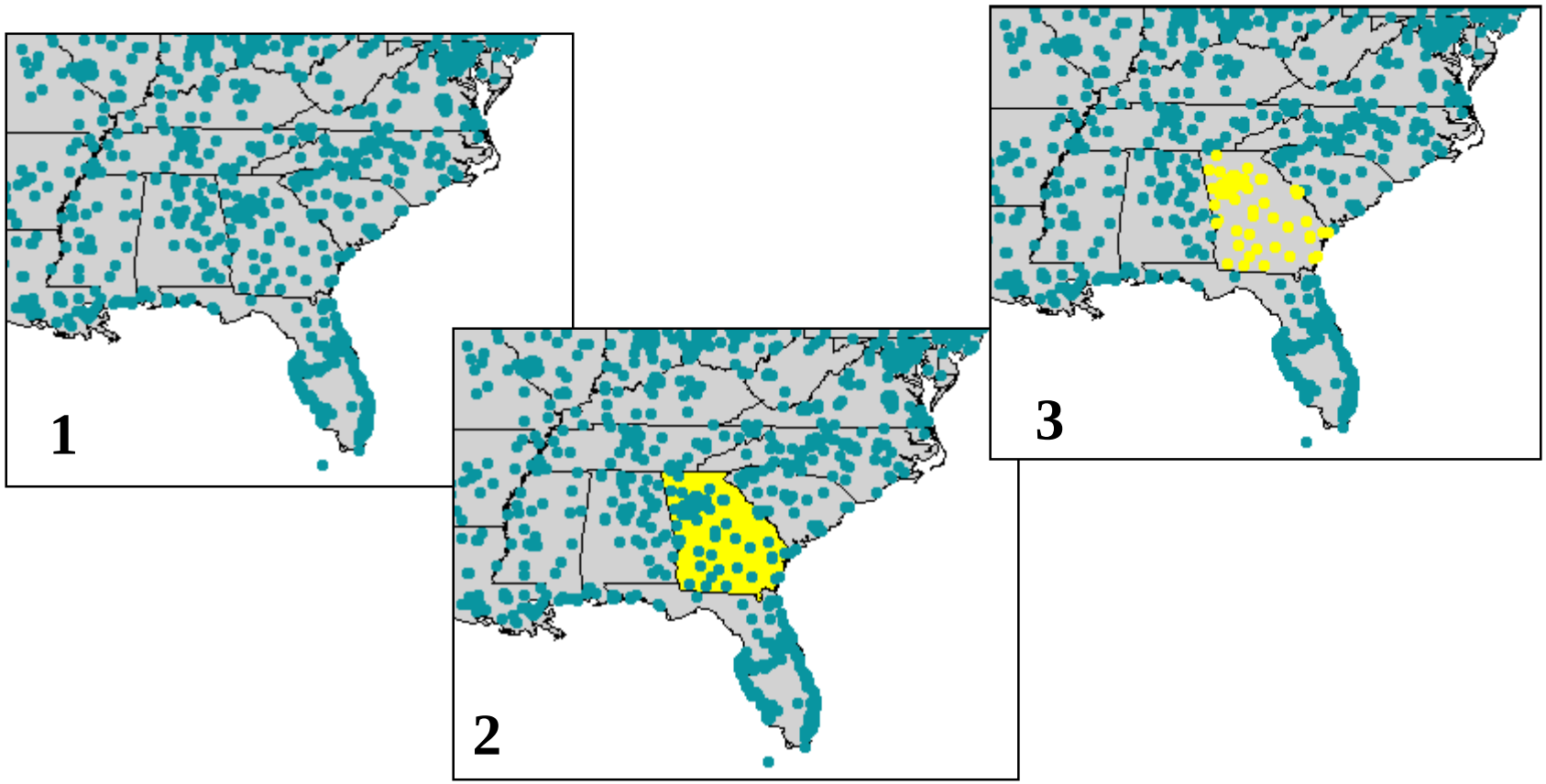


Covered by

- Object B is Covered by object A (object B will be selected)
- Opposite of covers
- Poly-poly
- Poly-line

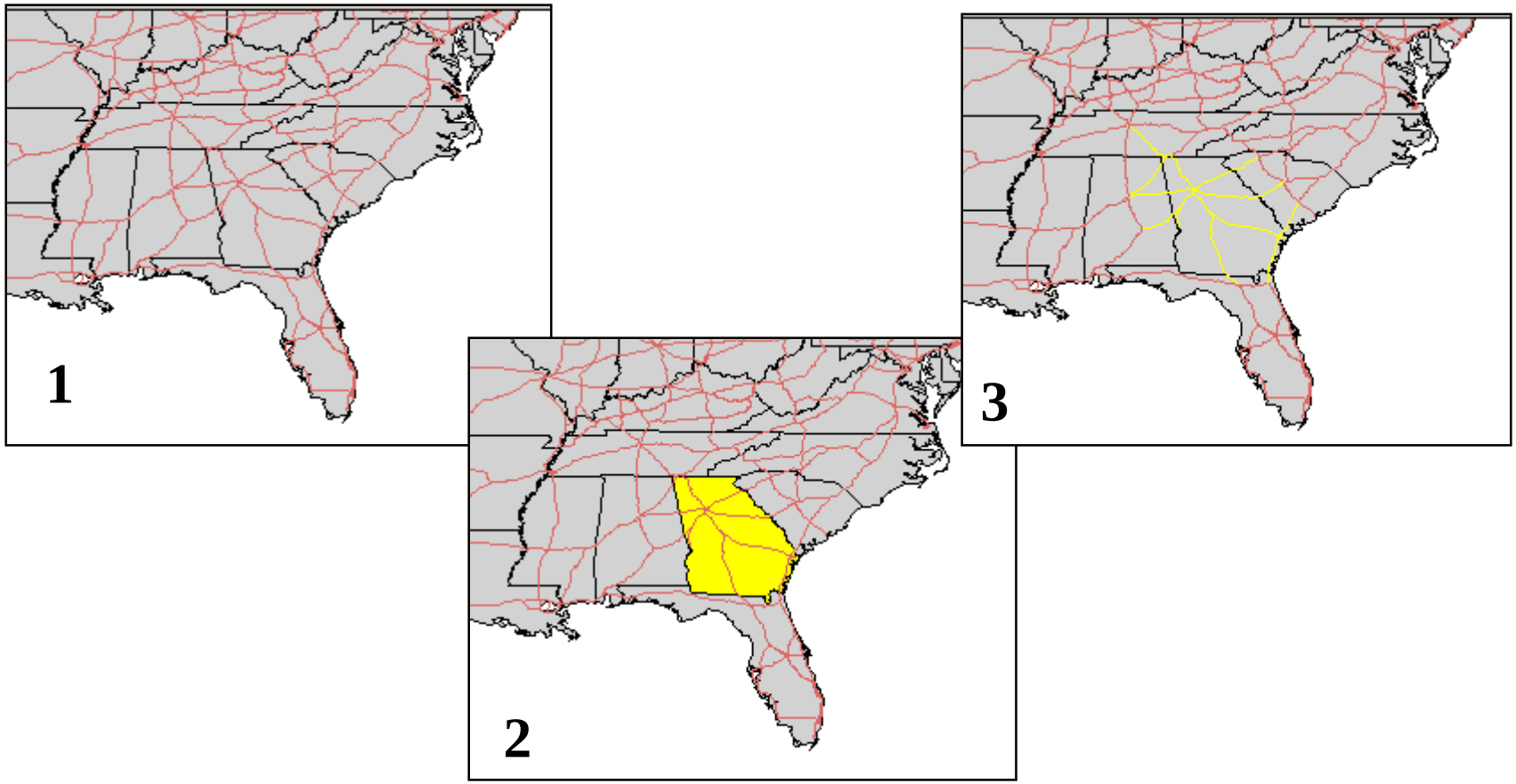


Selecting features that are inside selected objects



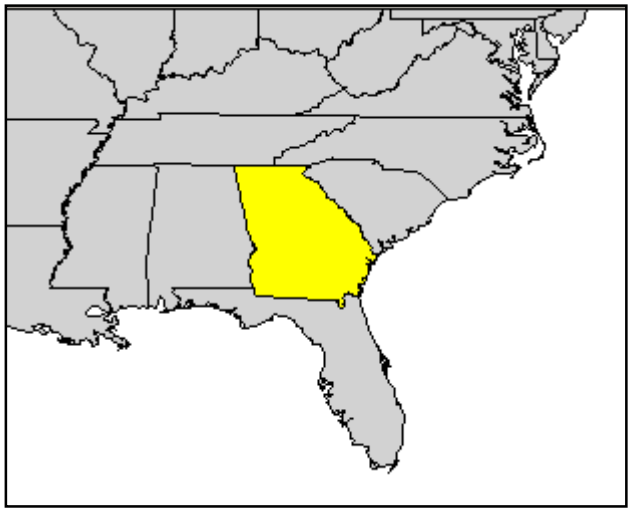
Select all cities that are located in the state Georgia. (INSIDE RELATIONSHIP)

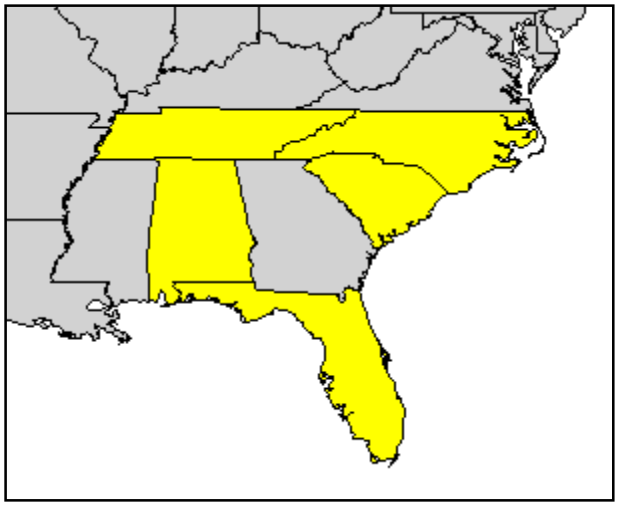
Selecting features that intersect



Select the highways that run (partly) through the state Georgia. First select the state Georgia, then select all the highways that intersect the selected state. (OVERLAP)

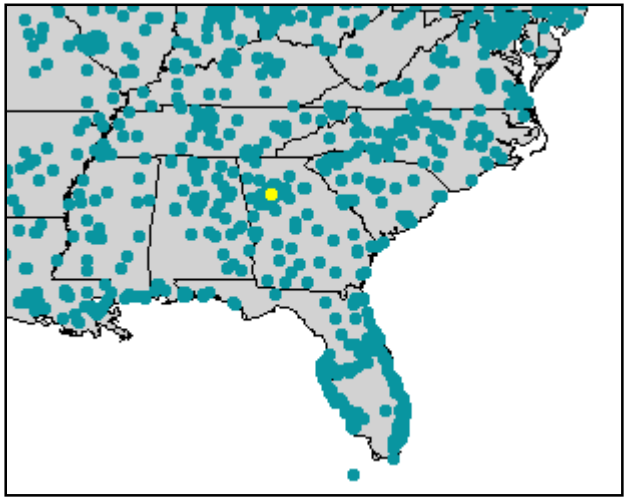
Select features adjacent to selected objects

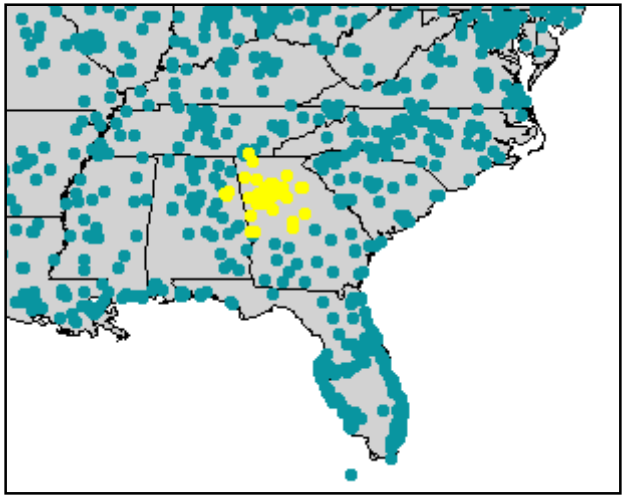




Select all the states that are neighbors of the state Georgia. (MEET RELATIONSHIP)

Selecting features based on their distance





Select all cities within a distance of 100 miles of Atlanta

Combining various selection techniques

Attribute Selection → Selection using Topological relationships → Attribute Selection

Select hospitals in the Building layer

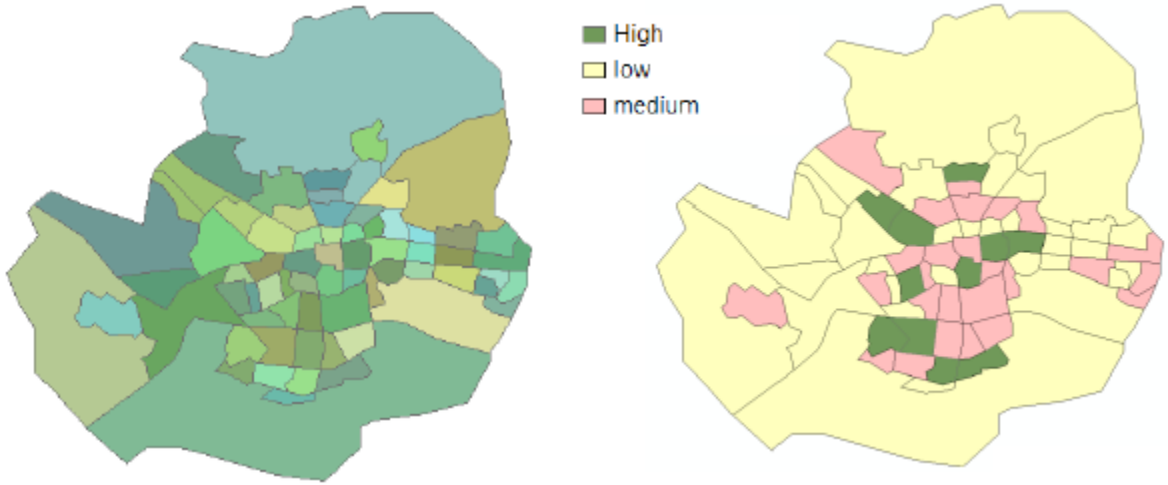
Select all roads that are within a distance of 200 meters from the selection in the Building layer

From the selected roads, Select the major roads

Question: Select all major roads that are located within a distance of 200 meter from a hospital.

(Re) Classification

- Allows (re)assignment of features to a class based on attribute value (vector data)
- (Re)Assignment of group of pixels to a class based on pixel value (raster data)



Classification is a technique of purposefully removing detail from an input data set in order to reveal patterns.

(Re) Classification

- Remove detail from an input dataset to reveal important spatial patterns.
- Reduce the number of classes and eliminate details.
- If the input dataset itself is the result of a classification we call it a reclassification.
- Reclassify data in different systems or for different purposes.
- Assign codes based on specific attributes.

(Re)Classification - procedure

Example:

- soil types reclassified into soil suitability for agricultural purpose.
- House hold income classification:
  - low
  - below average
  - average
  - above average
  - high.

Classification

- Classification - Reclassification
- Vector Classification with post processing
- User controlled classification
  - Classification table
- Automatic classification
  - Equal interval technique
  - Equal frequency technique

User controlled

- In user-controlled classification we indicate the classification attribute and the classification method.
- This is normally done via a classification table.

Two Examples of classification tables:

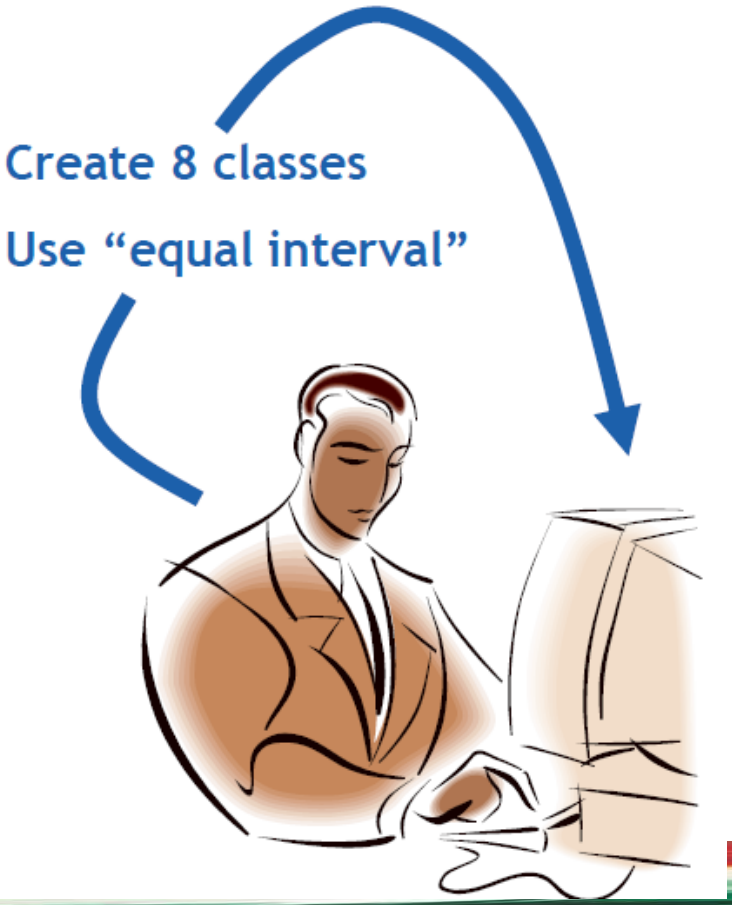
Old value	New value
391 - 2474	1
2475 - 6030	2
6031 - 8164	3

Code	Old value	New value
10	Planned Residential	Residential
20	Industrial	Commercial
30	Commercial	Commercial

The top table, the original values are ranges, in the lower table the old values already were a classification.

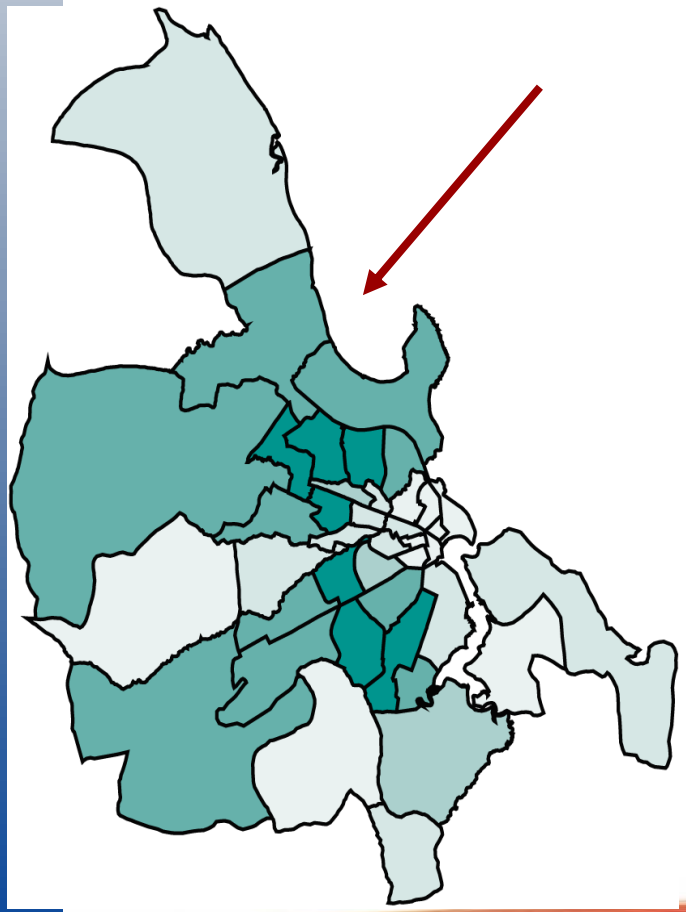
Automatic

- User specifies the number of output classes.
- Computer decides the class break points.
  - Equal frequency
  - Equal interval etc.

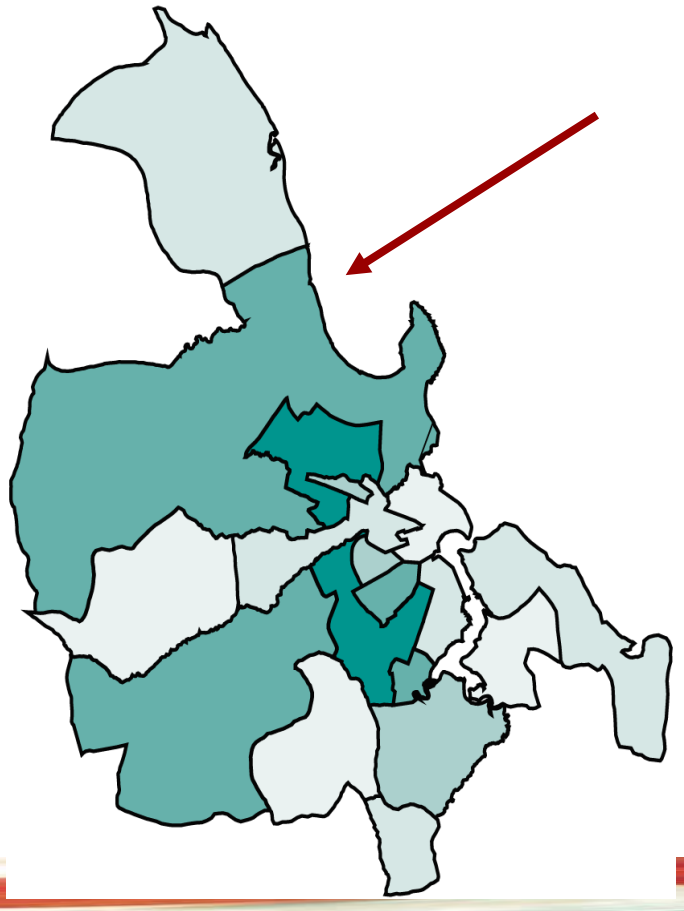


(Re)Classification - merge

Five classes of house hold income with original polygons intact



Five classes of house hold income with original polygons in the same categories were merged (boundary dissolved)



Differences between vector and raster data

	Vector	Raster
Geometric or topological change	**No.** Because only polygon (line, point) attributes are changed	**No**. Because only pixel attributes are changed.
Post processing: Spatial merging, aggregation or dissolve	**Yes.** For example, neighbour polygons with the same category are merged into one bigger feature.	**No.**

Thank You

Contact Details of the Faculty:

Email - prabhakar@iirs.gov.in
Tel - +91 1352524129