GIS DATA ANALYSIS & SPATIAL MODELLING

Unit-6

Data Analysis and Spatial Modeling

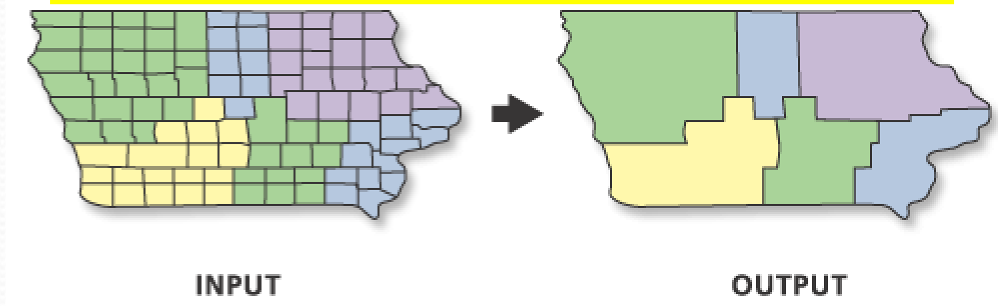
Dr. K.Palanivel
Assistant Professor
Centre for Remote Sensing
Bharathidasan University

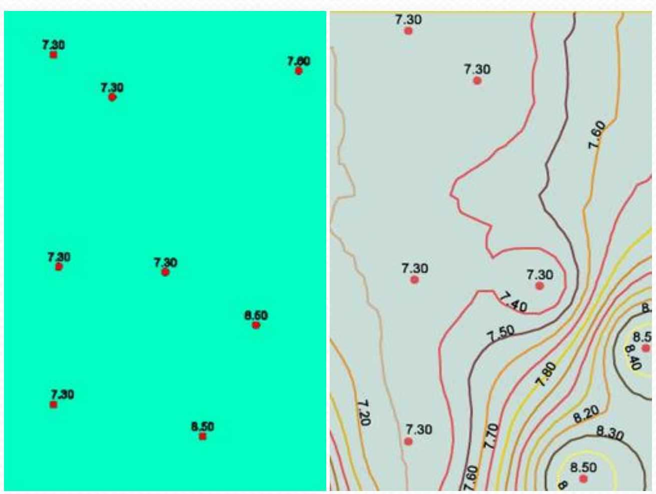
Unit-6

Data Analysis and Spatial Modeling: Simple data retrieval – Data retrieval through Boolean Logic – Map Overlaying and Cartographic Modeling (Two layers, Multiple layers, Binary, Index, Regression, and Process Models) – Overlay analysis, Capabilities (Point Operations, Regional Operations, Neighbourhood Operations) – Buffering – Cartographic Modeling using Natural Language Commands – Advantages and disadvantages of Carto modeling.

GIS DATA PREPARATION FOR ANALYSIS

Data Classification/Grouping , Regrouping,/Reclassification, Generalization





DISSOLVE

MERGE

UPDATE

CALCULATE RASTER ……

GIS ANALYSIS

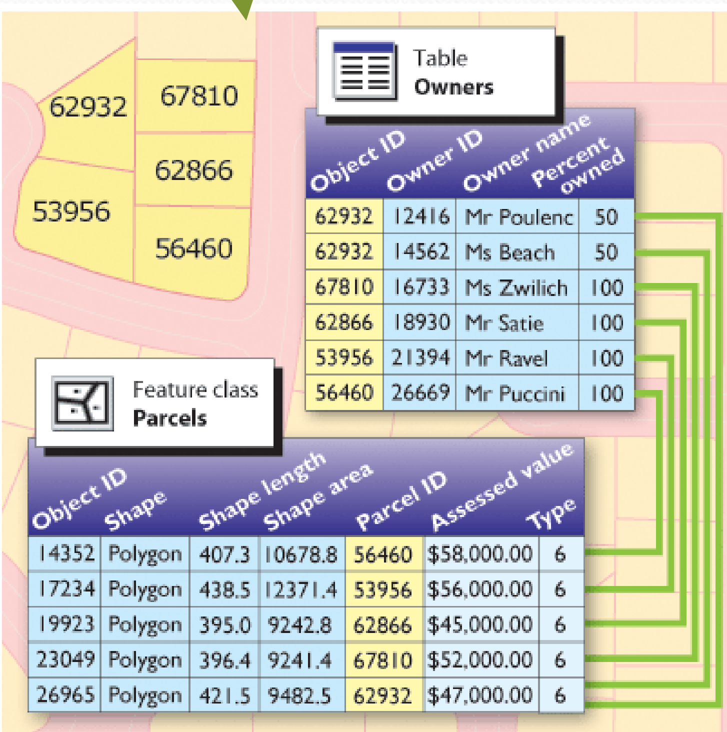
- GIS has the analysis capabilities ,that is not available with CAD
- Operates on Topology, spatial and/or non-spatial aspects combined
- The user can interactively work for getting the results
- Analysis capabilities range from
  - Retrieving subsets of information
  - up to univariate,
  - multivariate statistical and
  - spatial analysis using neighbourhood functions using interpolate methods
- Spatial modelling
- Simulation programs

Spatial Data

Unique Identifier (Unique Id)

Non Spatial Data

Provide Link using 'Join & Relate' methods



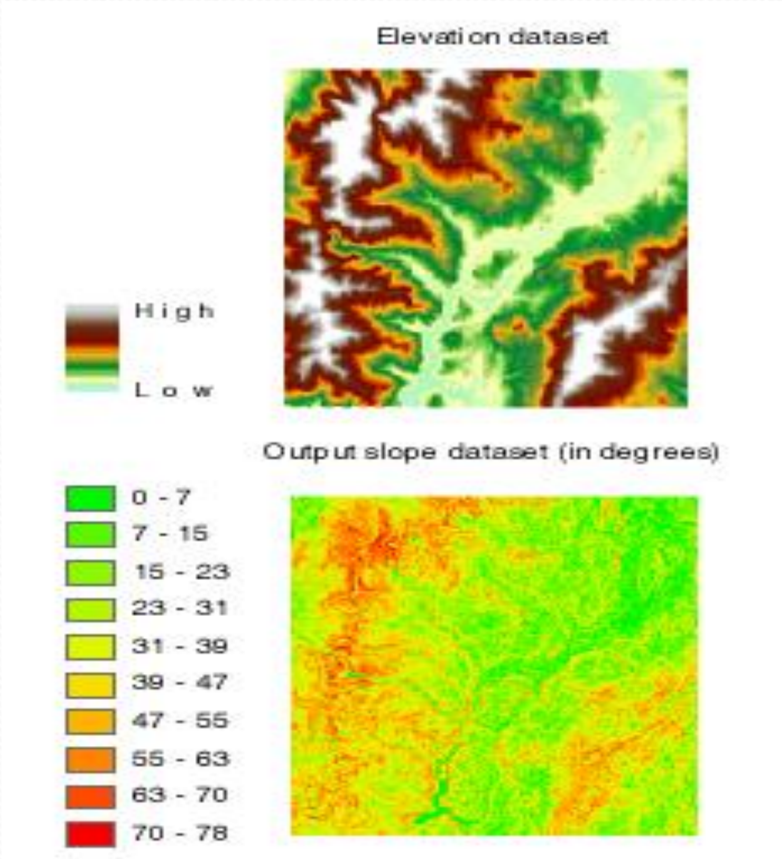
Object ID	Owner ID	Owner name	Percent owned
62932	12416	Mr Poulenc	50
62932	14562	Ms Beach	50
67810	16733	Ms Zwilich	100
62866	18930	Mr Satie	100
53956	21394	Mr Ravel	100
56460	26669	Mr Puccini	100

Object ID	Shape	Shape length	Shape area	Parcel ID	Assessed value	Type
14352	Polygon	407.3	10678.8	56460	$58,000.00	6
17234	Polygon	438.5	12371.4	53956	$56,000.00	6
19923	Polygon	395.0	9242.8	62866	$45,000.00	6
23049	Polygon	396.4	9241.4	67810	$52,000.00	6
26965	Polygon	421.5	9482.5	62932	$47,000.00	6

- It is necessary to establish a link between database and output that provides answer in different forms – maps, tables, figures, etc. using "**Unique_Identifier**"
- That link is any function used to convert data from one or more input maps into an output / **Action Plan** map
- Define the database – nature of geodatabase – **vector** / **raster**, and whether it contains **points**, **lines**, **polygons** & **associated attributes**, having **same projection**

GIS Data Analysis ...contd...

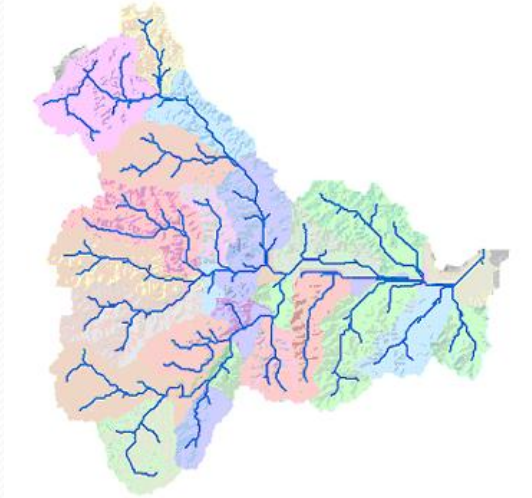
SLOPE ANALYSIS



Slope Categories – DEGREES , ASPECT, LENGTH

Intervisibility Analysis, etc.

DRAINAGE ANALYSIS



Delineation of Drainages,

Flow Direction, Flow Accumulation

Flow Length,

Demarcation of drainage basins,

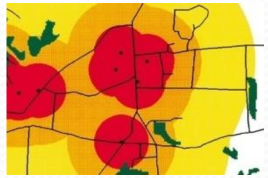
Runoff estimation,

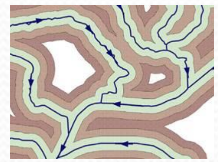
Quantum of soil erosion, etc.

Hydrological Modeling

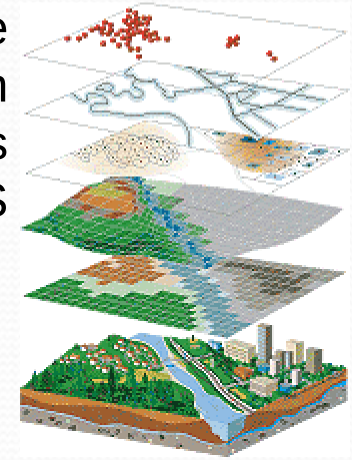
GIS Data Analysis & integration

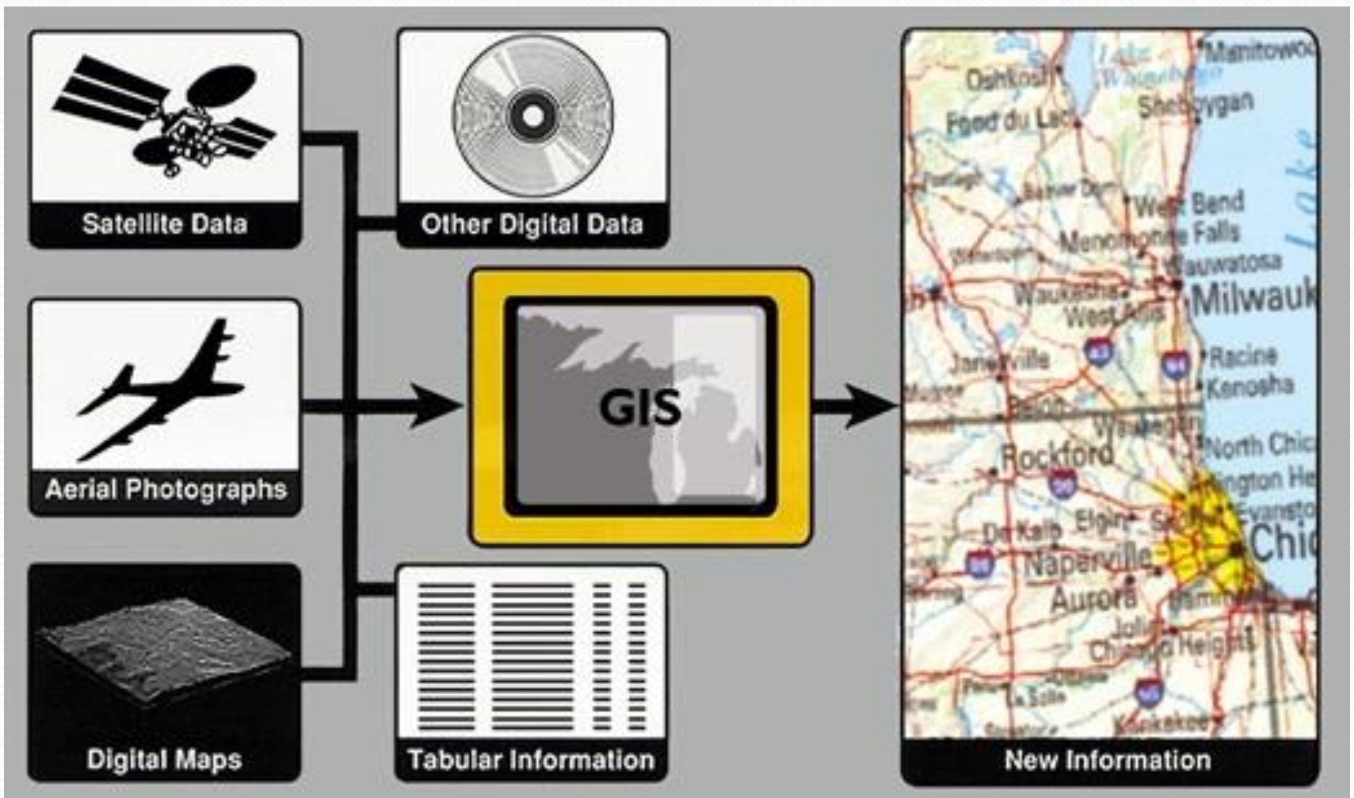
- Performing overlays
- Creating buffers
- Calculating statistics
- Merging datasets

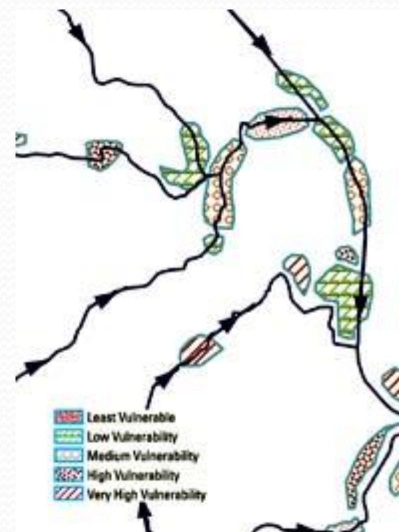




Data integration is the linking of information in different forms through a GIS



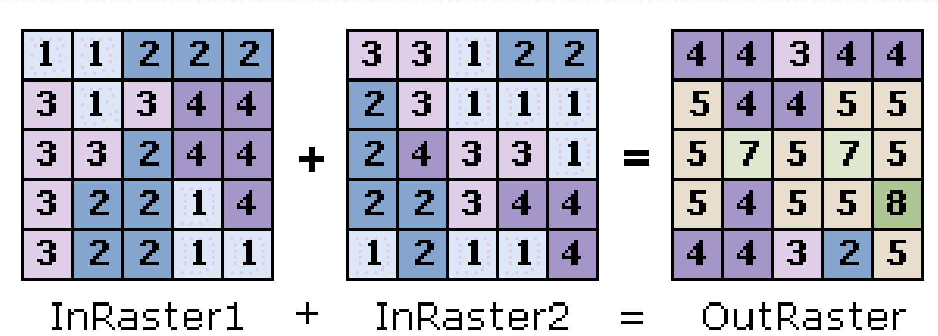


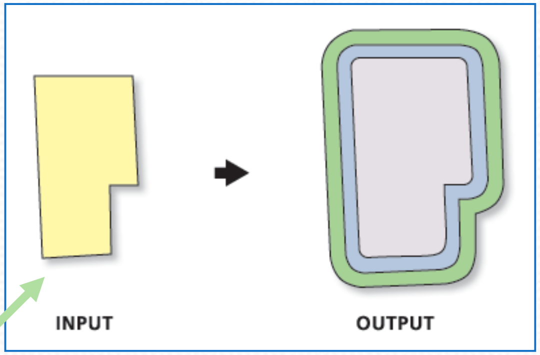


GIS Can do many operations / Analysis like add, subtract, multiply, divide, square, buffer, corridor, integrate multiple layers, etc.

Raster Based Addition

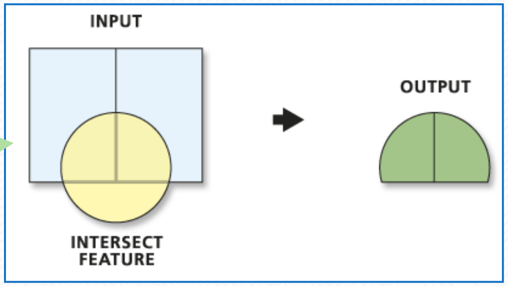
GO BACK



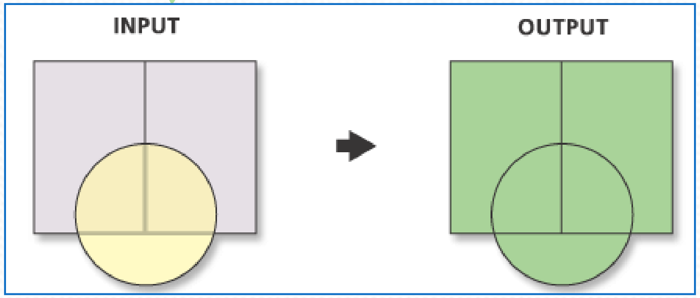


BUFFER

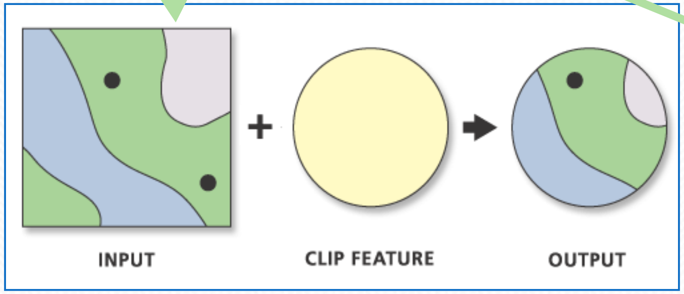
Vector Based Layer Integration



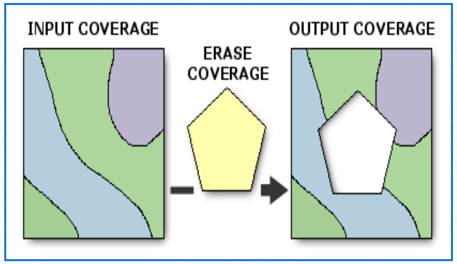
INTERSECT



UNION

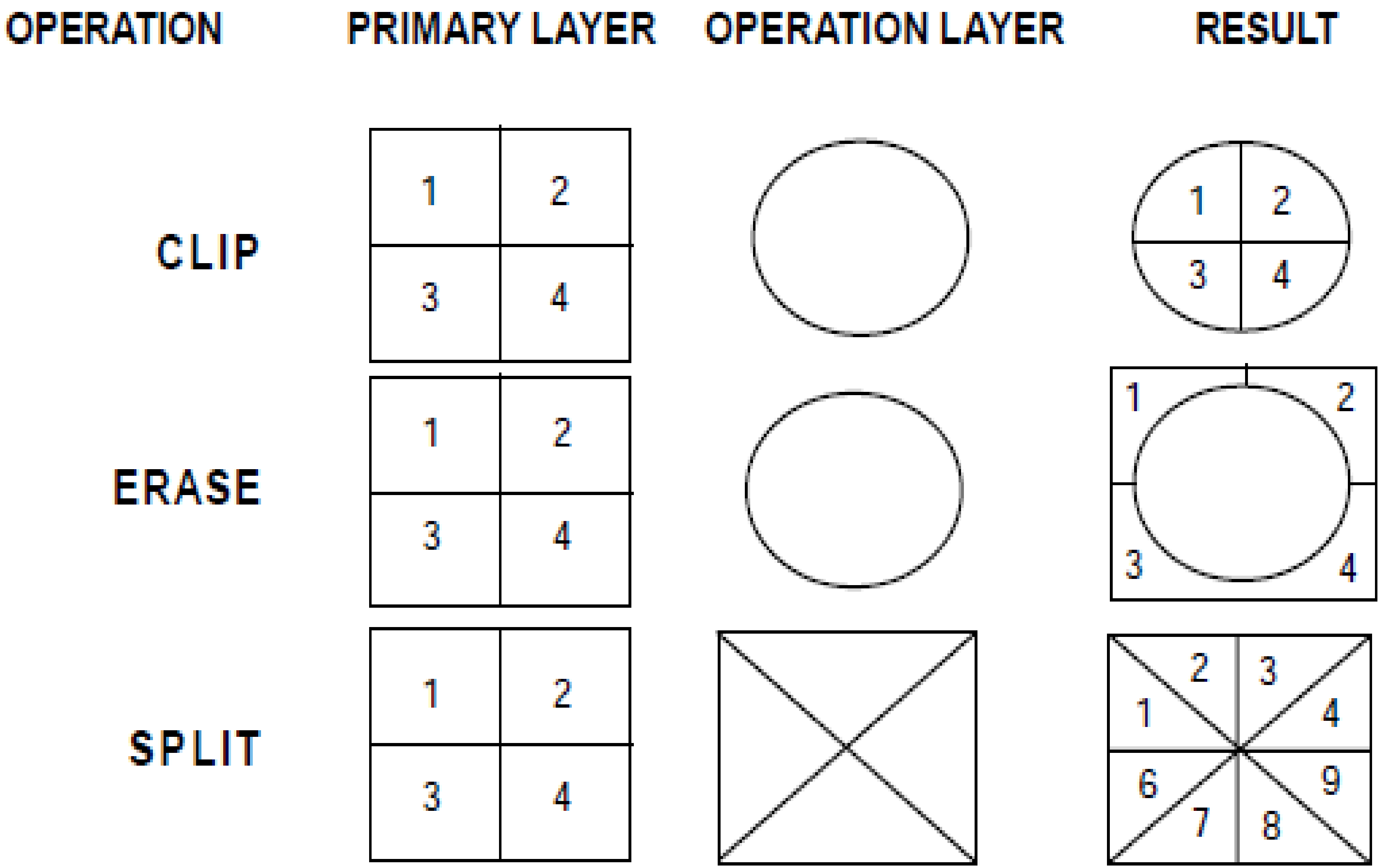


CLIP

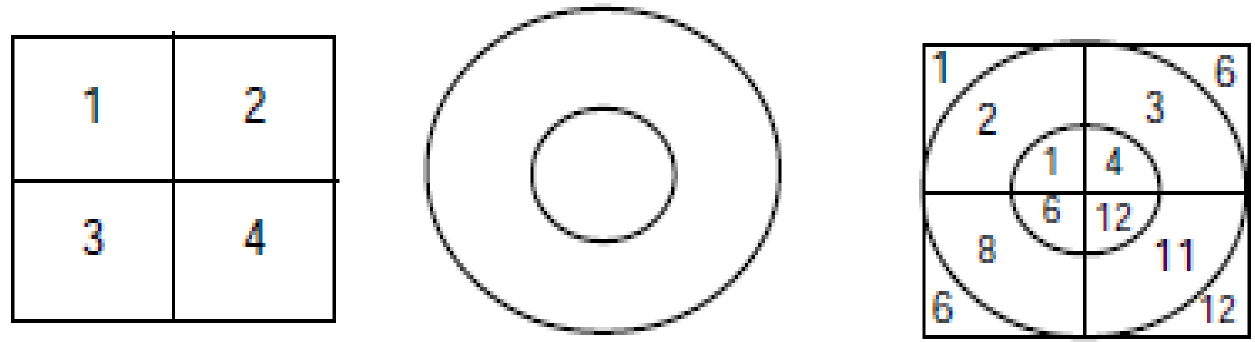


ERASE

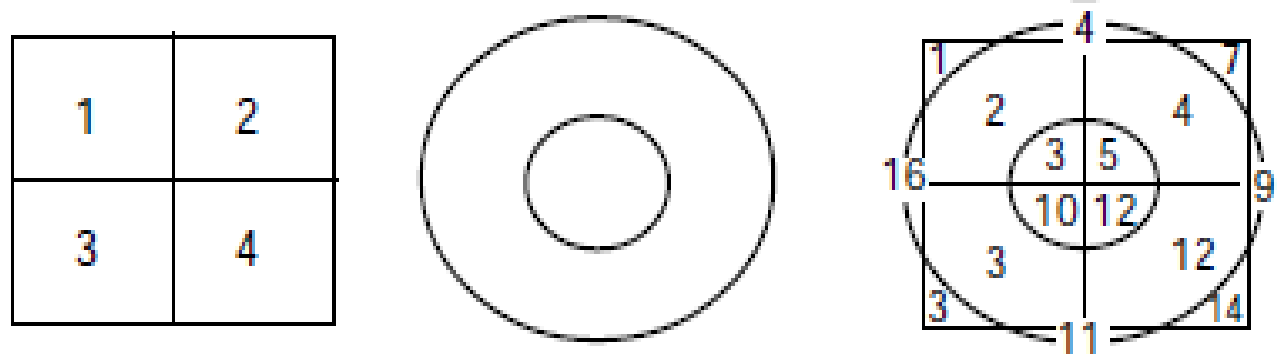
TYPES OF VECTOR OVERLAY OPERATIONS



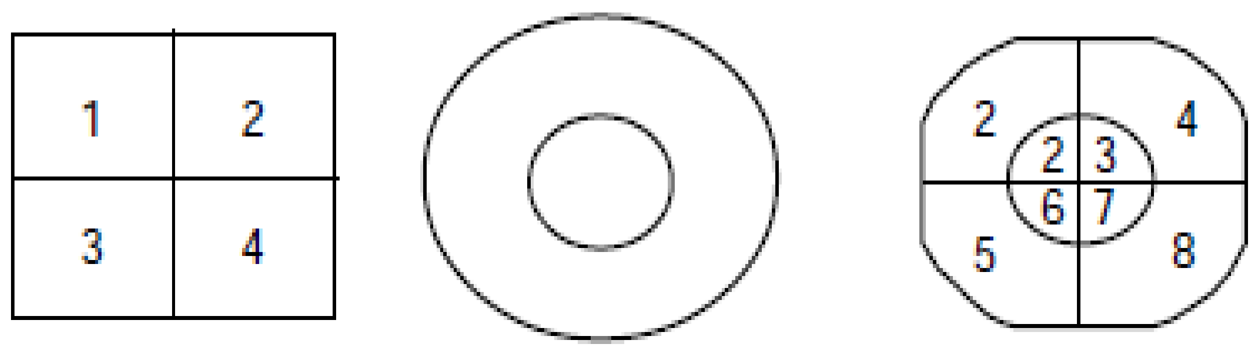
IDENTITY

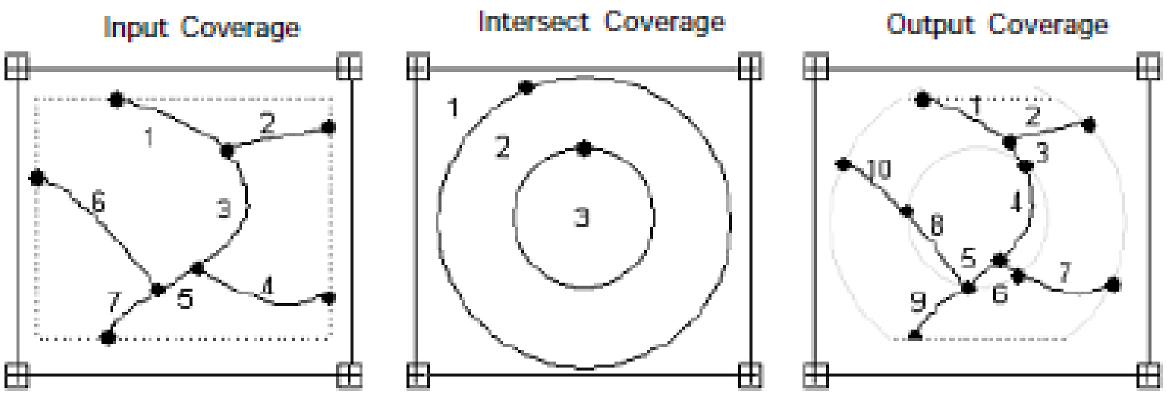


UNION



INTERSECT





INPUT COVERAGE	
#	ATTRIBUTE
1	A
2	B
3	A
4	C
5	A
6	D
7	A

INTERSECT COVERAGE	
#	ATTRIBUTE
1	
2	102
3	103

OUTPUT COVERAGE	INPUT COVERAGE		INTERSECT COVERAGE	
#	#	ATTRIBUTE	#	ATTRIBUTE
1	1	A	2	102
2	2	B	2	102
3	3	A	2	102
4	3	A	3	103
5	5	A	3	103
6	4	C	3	103
7	4	C	2	102
8	6	D	3	103
9	7	A	2	102
10	6	D	2	102

One of the overlay operation is AND (or INTERSECT) in vector layer operations, in which two coverages are combined.

Only those features in the area common to both are preserved.

Feature attributes from both coverages are joined in the output coverage.

Raster Based Overlay: Simple Addition

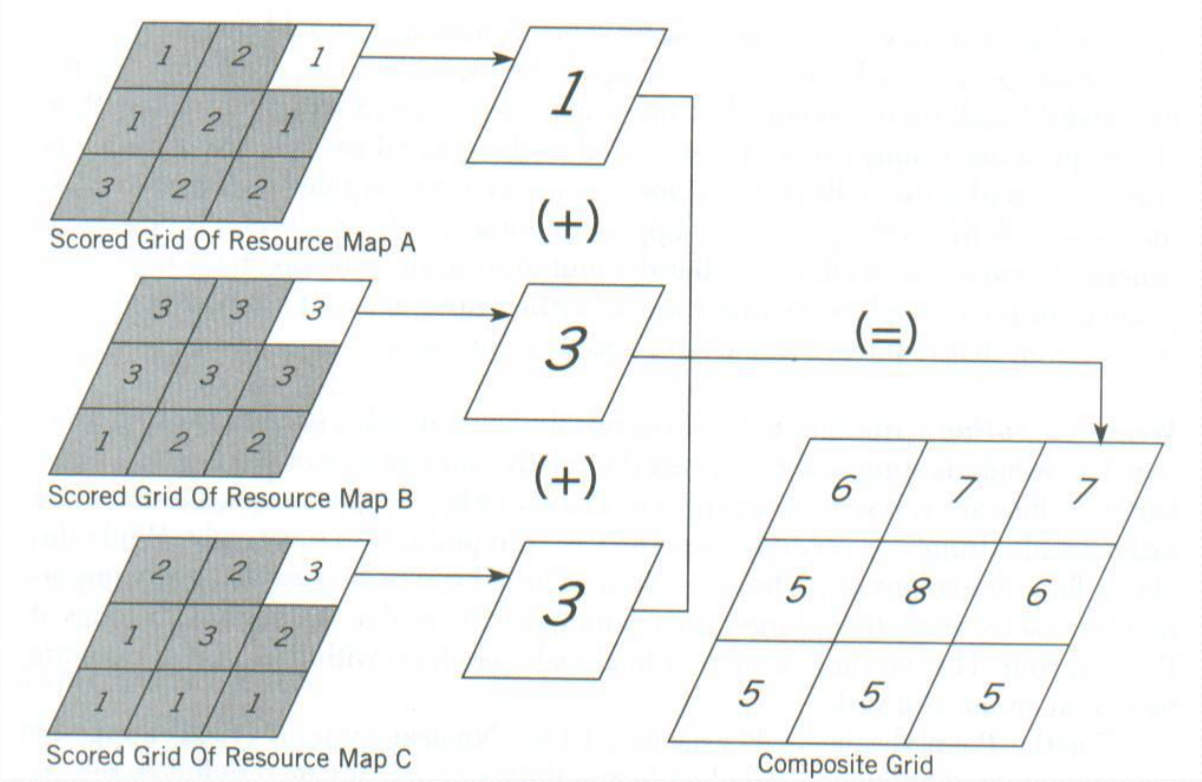


Image Source: Chrisman, Nicholas.(2002). 2nd Ed. *Exploring Geographic Information Systems.* p 144. fig. 5-12.

Raster Overlay: Boolean Combine

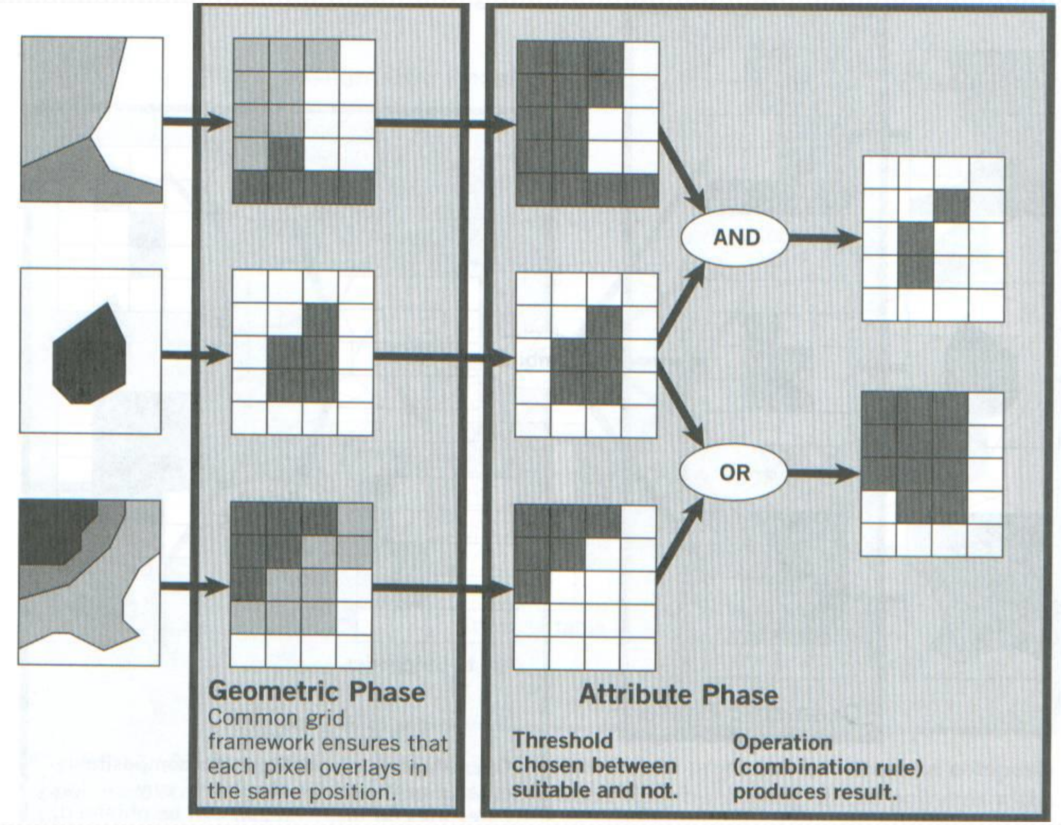
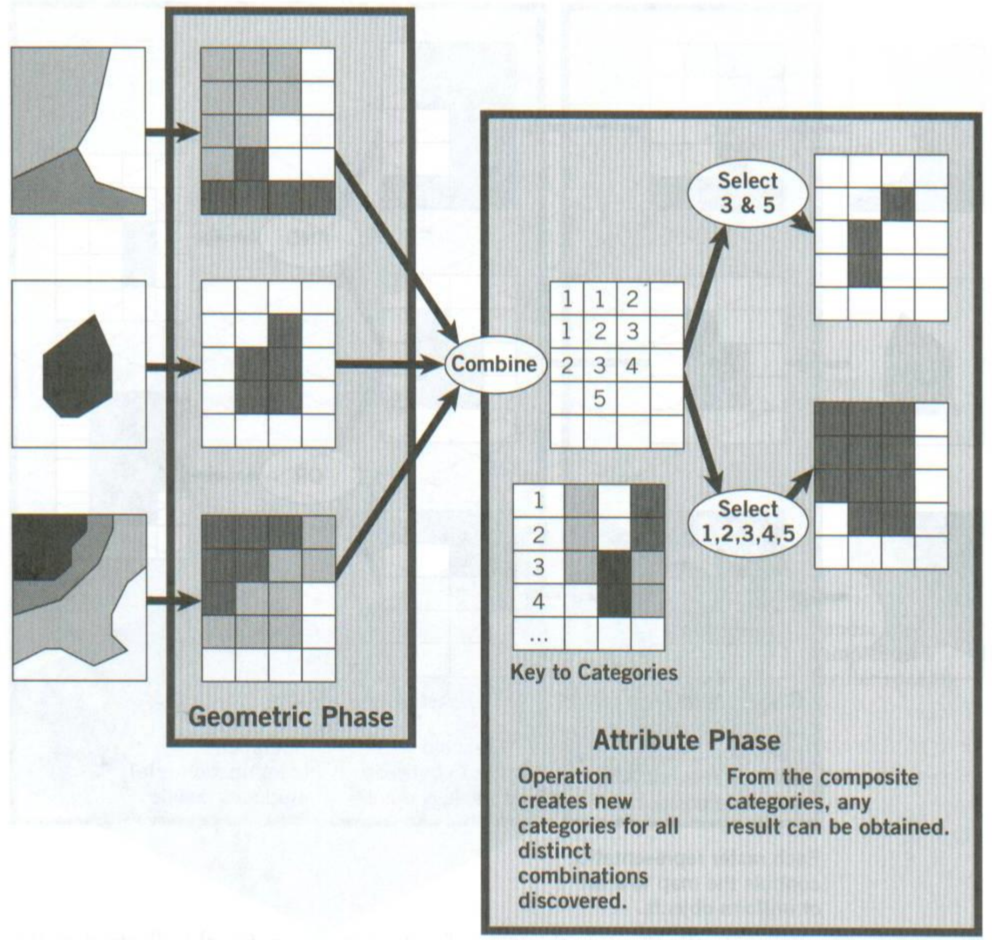


Image Source: Chrisman, Nicholas.(2002). 2nd Ed. *Exploring Geographic Information Systems*. p 125. fig. 5-3.

Raster Overlay: Composite Combine



Simple Data Retrieval

- **Boolean logic**
- **Reclassification and display**
- **Boolean operations on two or more maps**

Boolean logic

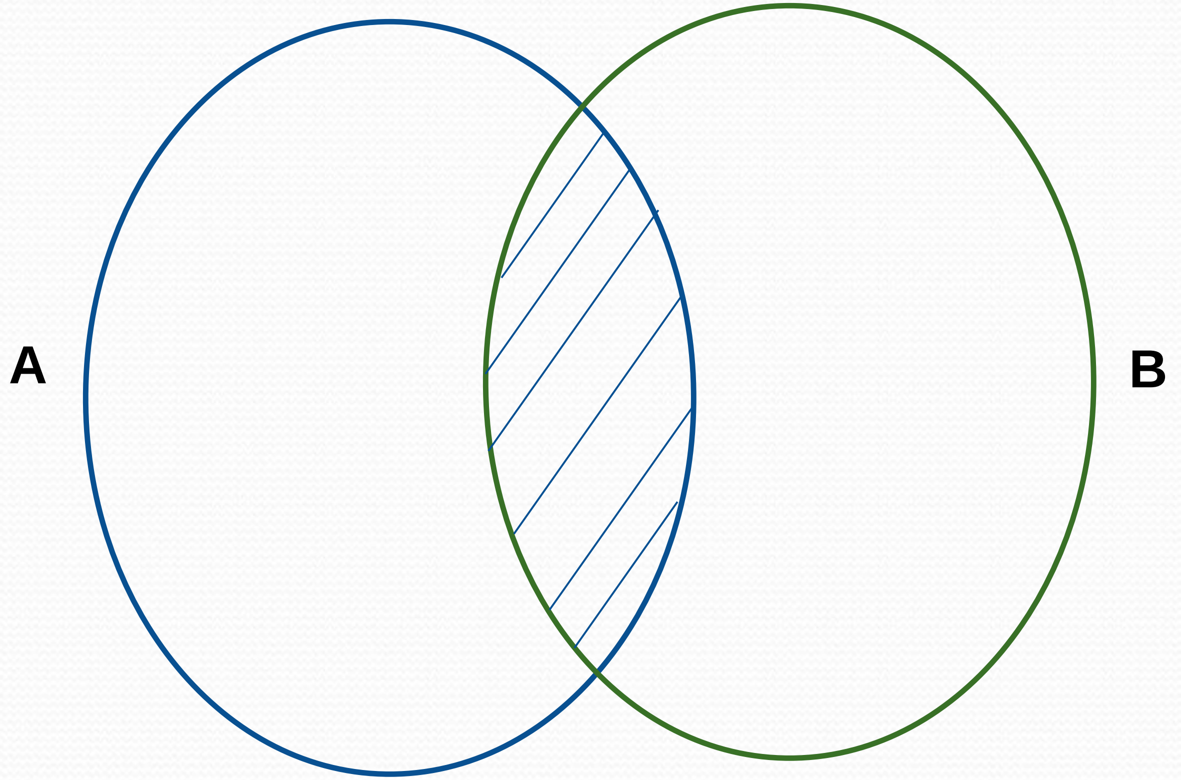
- Retrieve data by using the rules of Boolean logic to operate on the attributes and spatial properties
- Operators are : AND, OR, XOR, NOT
- The particular condition is ‘true’ or ‘false’
- Visual portray of Boolean logic is in the form of ‘Venn diagrams’
- A is the set of items having attribute ‘A’ and B is the set of items having attribute ‘b’, then the statements

A **AND** B, A **OR** B, A **XOR** B, A **NOT** B,

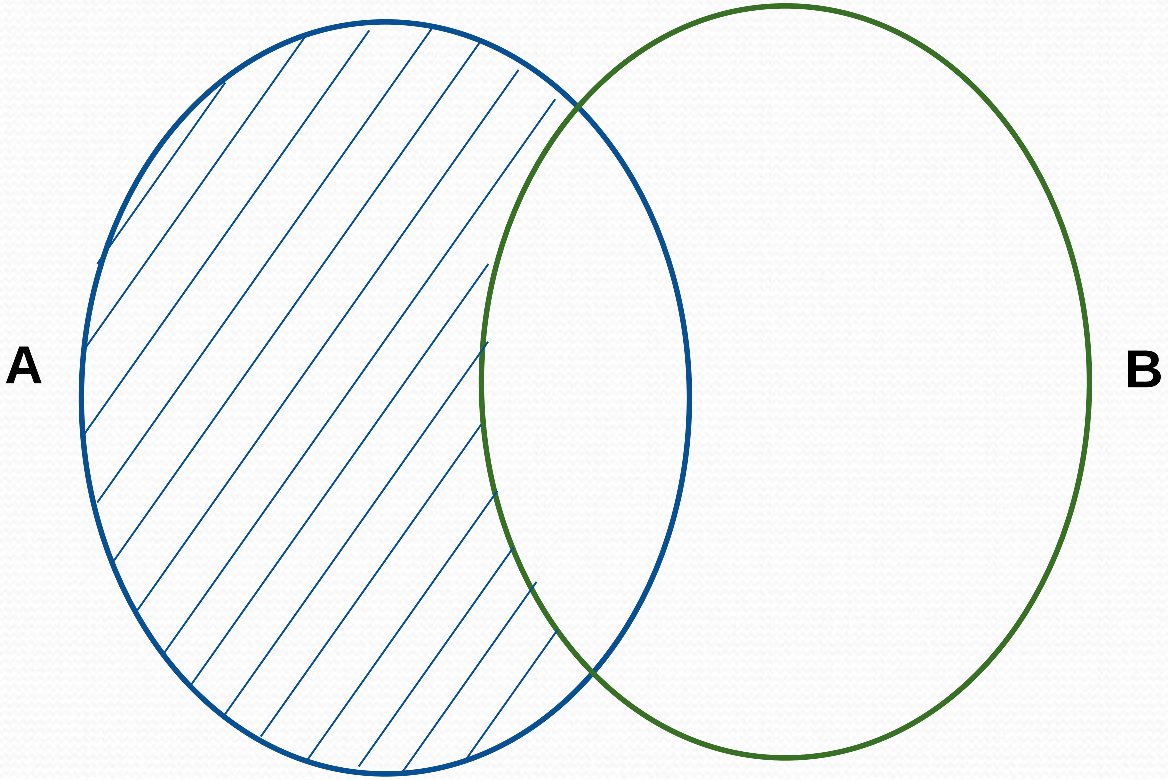
(A **AND** B) **OR** C, A **AND** (B **OR** C)

Will return items that have attribute combination covered by the **shaded portions** respectively

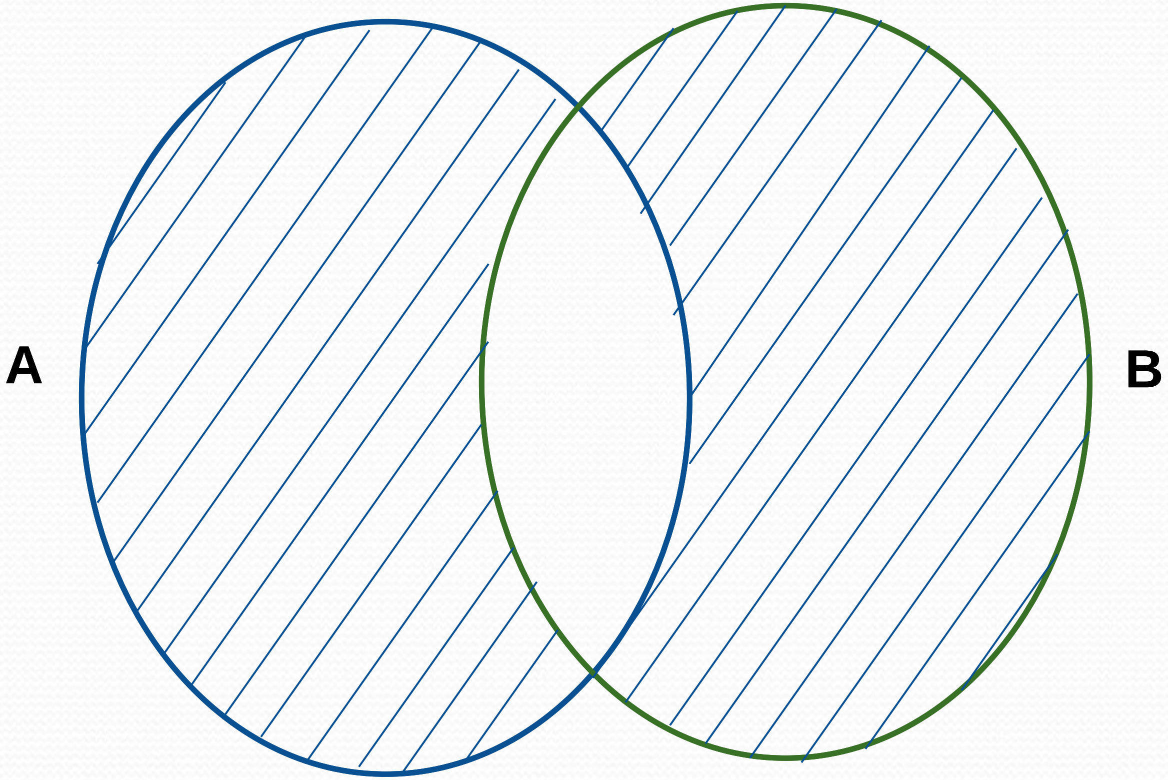
- On the basis of priority of AND with respect to OR
- Parentheses are usually used to represent the order of evaluation
- More complicated searches involve the shapes of areas, the properties of the boundaries of areas or the neighbouring areas.
  - All areas of woodland bordering urban areas
  - All open areas that have a particular shape and area bounded by hedgerow (a row of bushes)



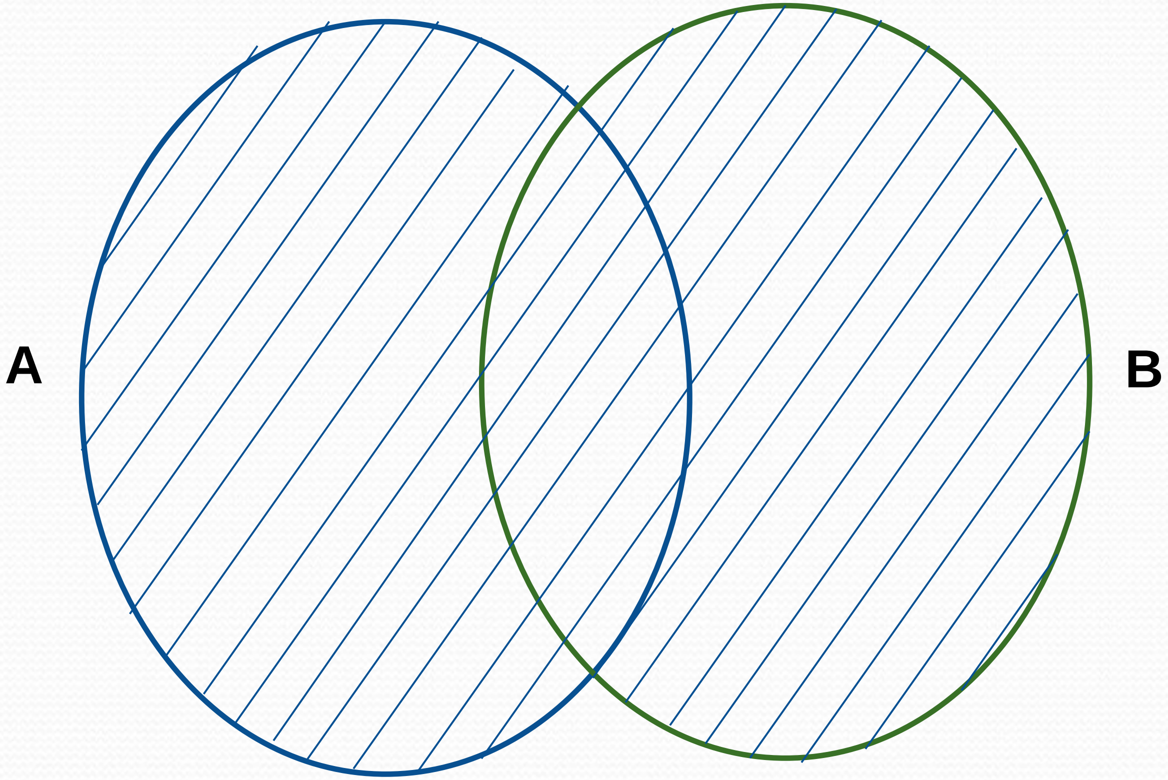
A AND B



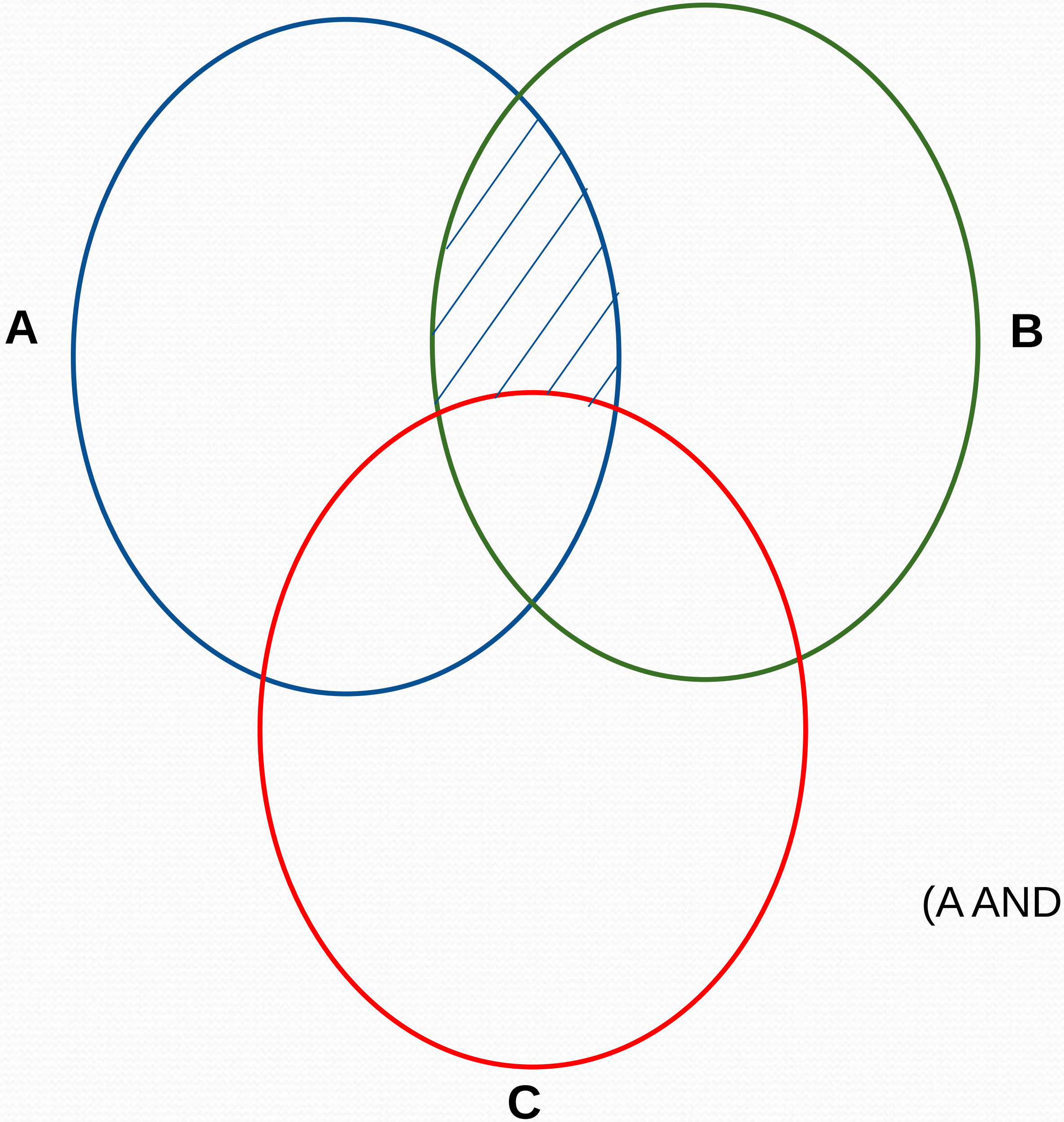
A NOT B



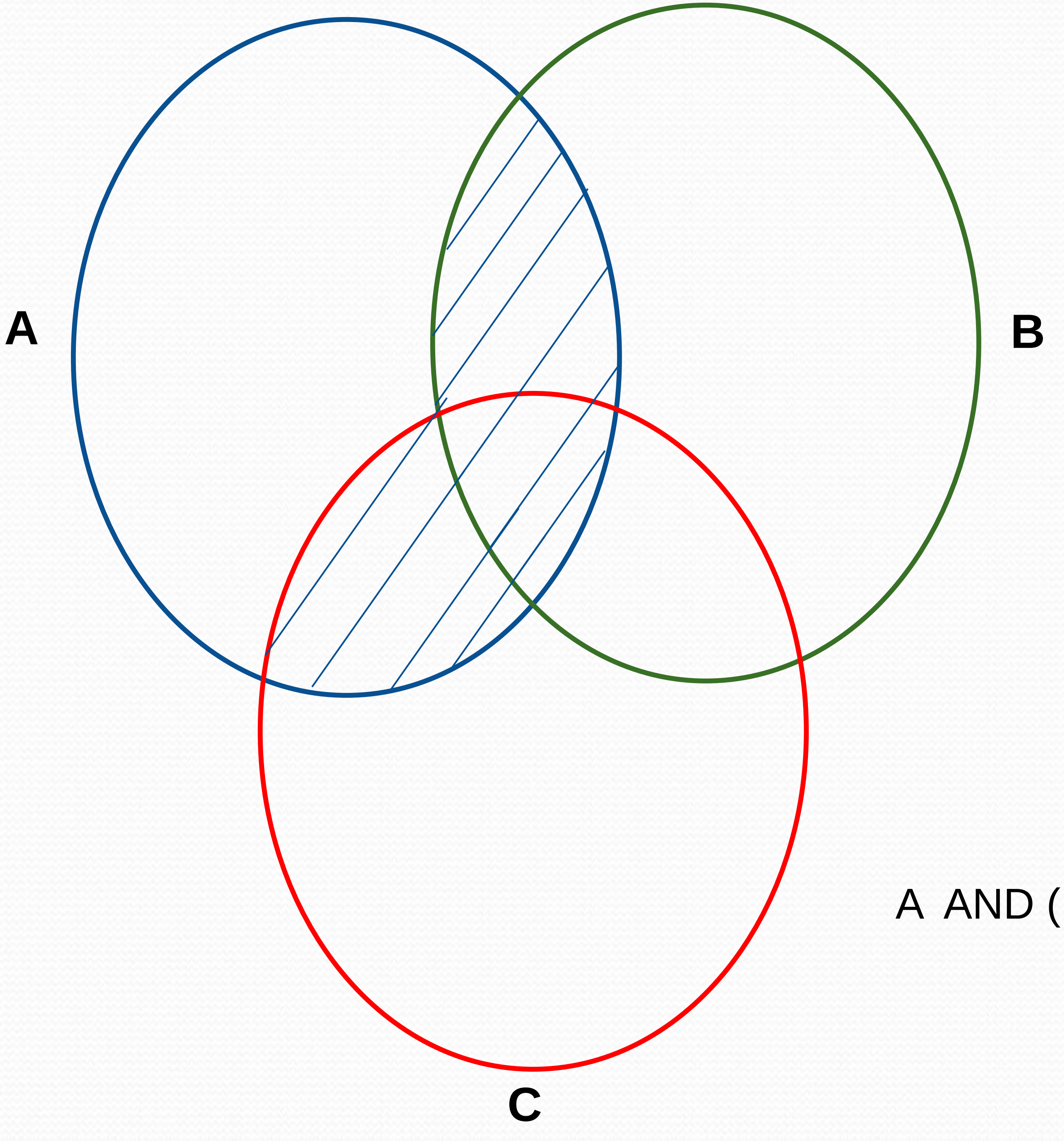
A XOR B



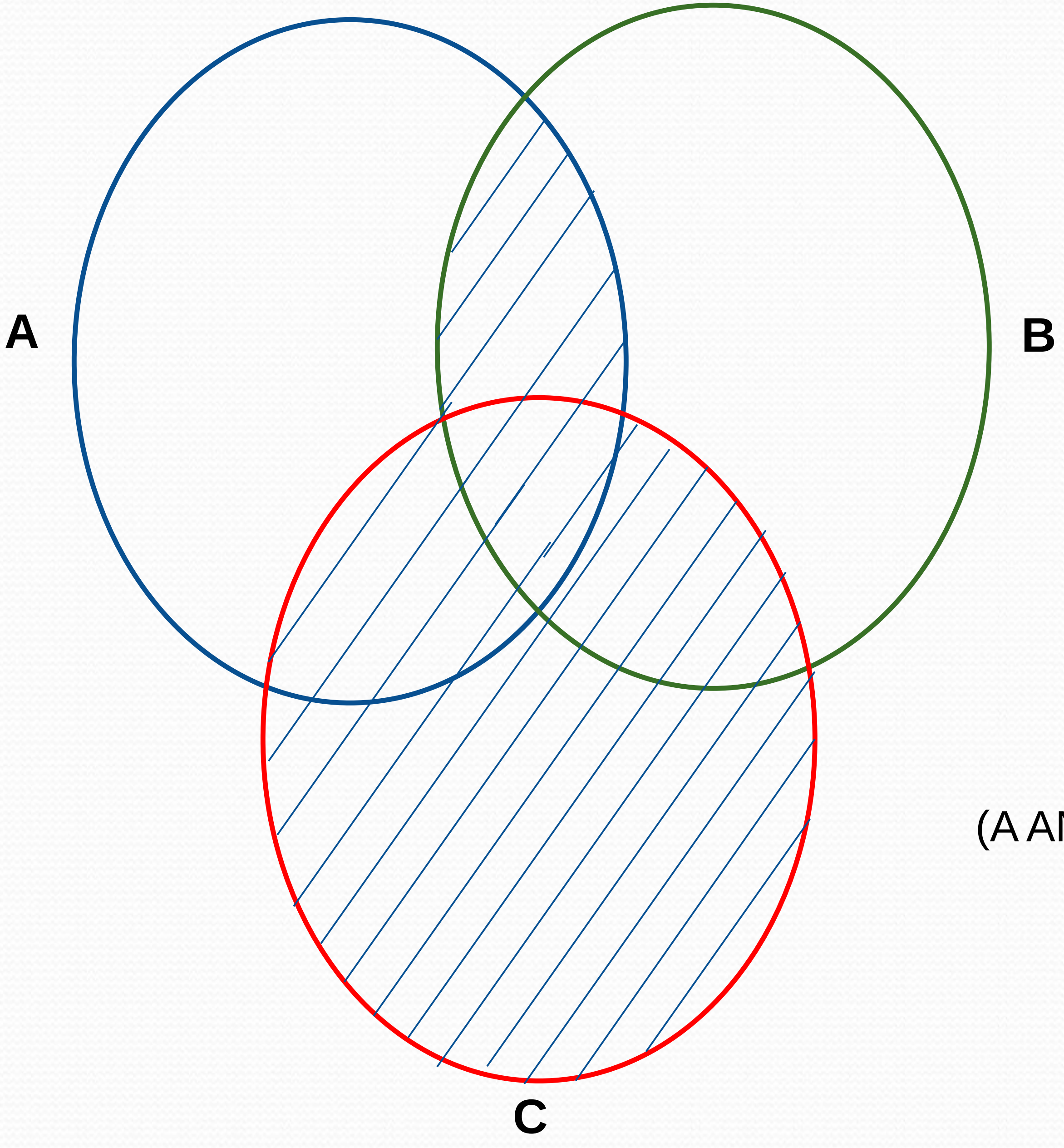
A OR B



(A AND B) NOT C



A AND (B OR C)



(A AND B) OR C

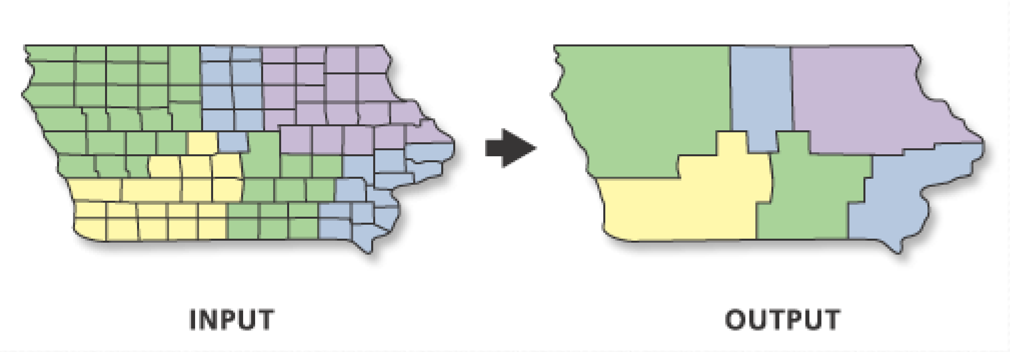
Reclassification and display

- After the retrieval of required data, reclassification is necessary
- So that redisplay of the retrieved data can be done which will exactly reflect the user's perception
- The similar polygons in adjacent places / boundary shared polygons having same codes / values / attributes,
- Then Dissolving has to be done
- Then, recompute the polygon structure again

Due to the multiple fractionated classes of polygon groups with 1000's of features, it is important to provide clarity for representing the output.

Similar to processing of data prior to analysis, post processing of out put is also necessary.

Data Classification/Grouping , Regrouping,/Reclassification



Finally using option like,

DISSOLVE

MERGE

UPDATE

CALCULATE RASTER ……

Simplified user friendly outputs can be presented.

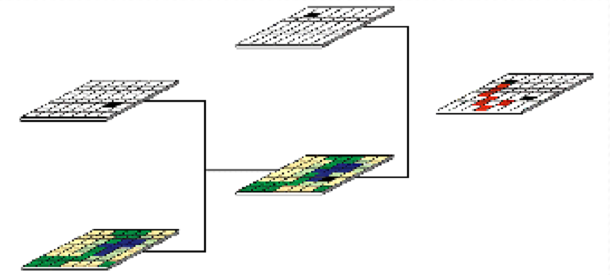
GIS Modeling

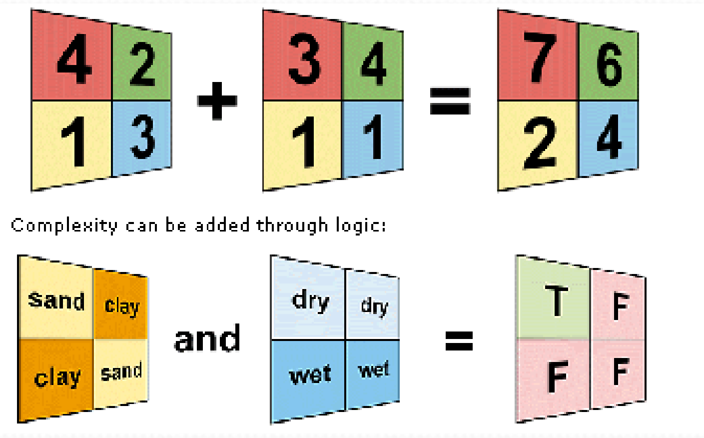
- A model is a representation of reality.

- Due to the inherent complexity of the world and the interactions in it, models are created as a simplified, manageable view of reality.

- Models help to understand, describe, or predict how things work in the real world.

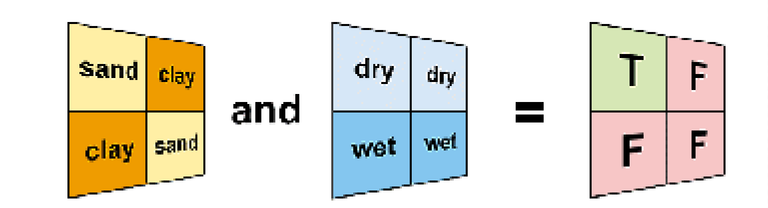
- **Binary model**
- **Process model**
- **Index model**
- **Regression model**
- **…**
- **…**





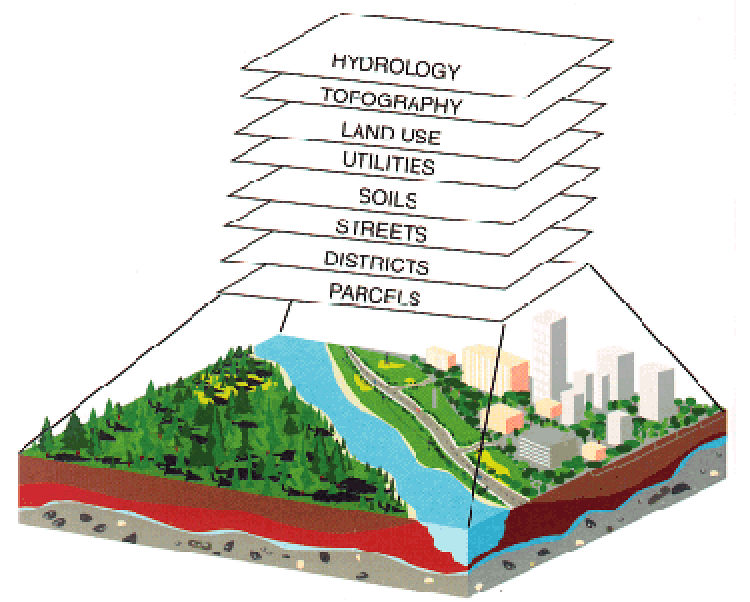
GIS MODELLING ...CONTD...

Binary Model



Yes / No, True / False, 1 / 0, etc.

Process Model



Lookup Table (LUT)

Combined Attribute table containing the values / classes available from the both thematic layers involved in GIS integration.

1st **Thematic Map** / Layer having 3 classes – e.g. Thickness of Top soil (TTS) High, Moderate & Less

+

2nd **Thematic Map** / Layer having 3 classes – e.g. Water level (WL) - Deep, Moderate & Shallow

	TTS-H	TTS-M	TTS-L
WL-D	DH	DM	DL
WL-M	MH	MM	ML
WL-S	SH	SM	SL

In **Binary model**, Thematic layers can be classified with only two classes, Yes - No. In this case, once the GIS integration is performed then, the number of combined classes in the resultant layers can be easily determined by a simple formula,

$$\boxed{N_{rc} = 2^n - 1}$$

Where,

n = no. of layers having Y/N classes involved in GIS integration analysis.

$\mathbf{N_{rc}}$ = resultant number of classes/combinations

LUT for GIS Integration of Two layers in Binary model

	TTS-1	TTS-0
WL-1	1 1	1 0
WL-0	0 1	0 0
4 combination s = **R1**		

R1 / Sl	SLP-1	SLP-0
1 1	1 1 1	1 1 0
1 0	1 0 1	1 0 0
0 1	0 1 1	0 1 0
0 0	0 0 1	0 0 0
8 combination s = **R2**		

R2 / Gm	GEM-1	GEM-0
1 1 1	1 1 1 1	1 1 1 0
1 0 1	1 0 1 1	1 0 1 0
0 1 1	0 1 1 1	0 1 1 0
0 0 1	0 0 1 1	0 0 1 0
1 1 0	1 1 0 1	1 1 0 0
1 0 0	1 0 0 1	1 0 0 0
0 1 0	0 1 0 1	0 1 0 0
0 0 0	0 0 0 1	0 0 0 0
16 combination s = **R3**		

GIS Integration of **Three layers** in binary model

First step of GIS integration in **Binary model** will result 3 + 1 (1-unwanted) classes or combinations.

In the **Second step** of GIS integration, 7 +1(unwanted) classes are resulted.

$$N_{rc} = 2^3 - 1$$

Similarly, for the GIS Integration involving **Four layers** in **Binary model** will result into ($2^4 - 1$) = **15** + 1 (1-unwanted) classes or combinations.

DATA SORTING And PERFORMING STATISTICAL OPERATIONS

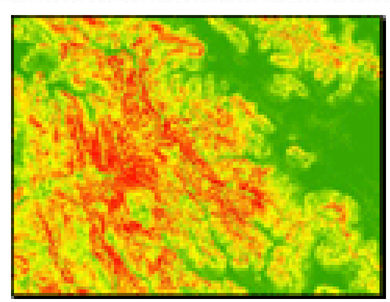
- Data Sorting
- Mean, Mode, Std. Deviation, Correlation, Factor analysis, Multi variate analysis, Multiple Regression analysis, etc.

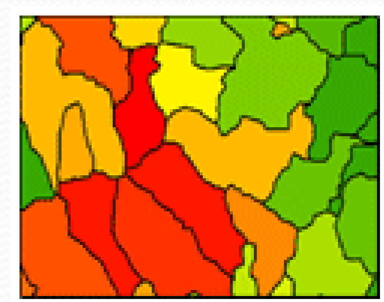
In order to perform statistical analysis in GIS, it is necessary to have samples of uniform length (equal number of samples) for all the thematic maps.

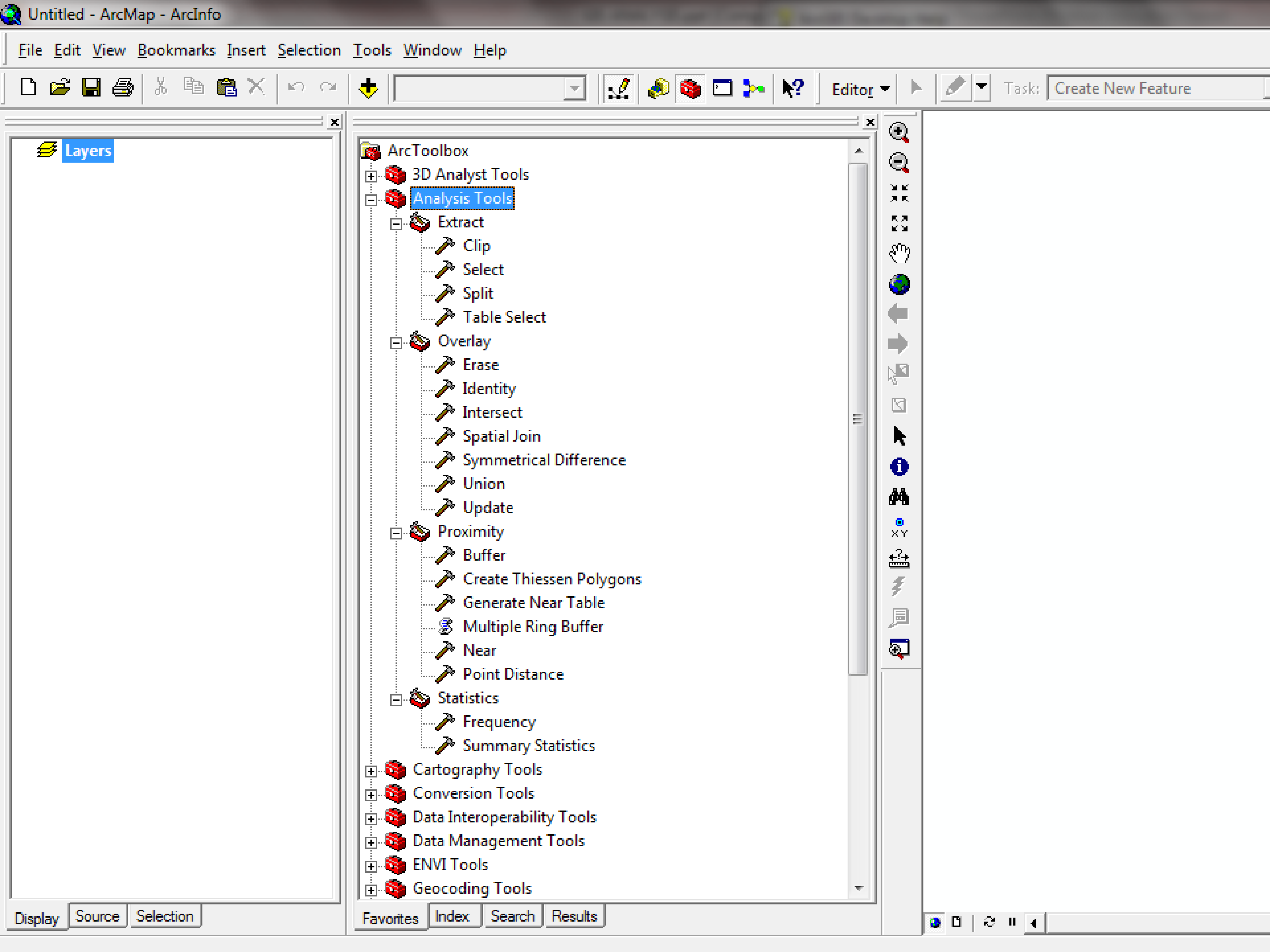
Hence, it would be better to convert the vector data to Raster format for statistical analysis.

DATA CONVERSION

Raster ⇄ Vector






Untitled - ArcMap - ArcInfo
File Edit View Bookmarks Insert Selection Tools Window Help
Editor
Task: Create New Feature
Layers
ArcToolbox
3D Analyst Tools
Analysis Tools
Extract
Clip
Select
Split
Table Select
Overlay
Erase
Identity
Intersect
Spatial Join
Symmetrical Difference
Union
Update
Proximity
Buffer
Create Thiessen Polygons
Generate Near Table
Multiple Ring Buffer
Near
Point Distance
Statistics
Frequency
Summary Statistics
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
ENVI Tools
Geocoding Tools
Display
Source
Selection
Favorites
Index
Search
Results

Unit-7

Network Modeling: Networking and Dynamic Segmentation – Applications, Minimum Distance Model, Maximum Covering Model (P-median model), Urban Transportation Planning Model.

Network Analysis / Model

- **Network** – Consists of connected linear features
    - E.g. Road, Railway, Bicycle paths, Streams & Shorelines.
- **Dynamic Segmentation**
    - A data model that is built upon lines of a network and allows the use of real-world co-ordinates with linear measures such as mileposts
    - To link linearly referenced data such as accidents & pavement conditions to geographically referenced road network
- For this, the networks need to be segmented and labelled by inserting nodes in proper locations
- For both **NW/DS** - Appropriate attributes are to be collected as per the real-world applications and entered in computer for 'Linking'.
    - For e.g., to use road network for Path Finding / Resource Distribution, etc., - we need the following attributes:

- Travel time
- Impedance at turn
- One-way streets
- Linear measures – to be added with arcs and nodes of the network – can be directly linked to road segments.

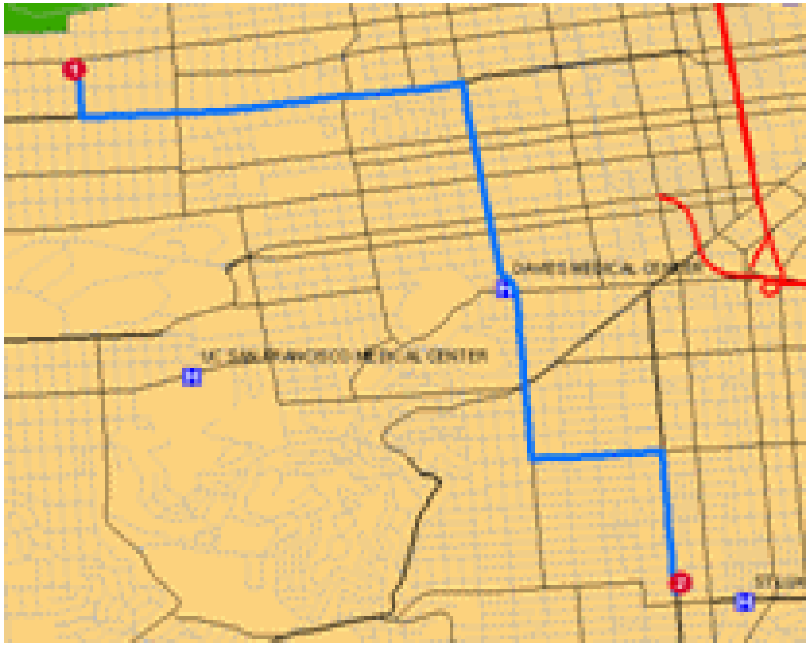
- 1. Network – is a line coverage, which is topology based and has the appropriate attributes for the flow of objects such as traffic.
  - Linear Feature + Attribute data, such as
    - Impedance values assigned to
      - network links – link impedance
      - Turns – turn impedance
    - One-way streets and
    - Overpasses and under passes.

Link impedance

- A link refers to a segment separated by two nodes in a road network

NETWORK ANALYSIS

Route Identification



NETWORKS AND APPLICATIONS

Networks are simple. They are comprised of two fundamental components, *edges and junctions*

Edges are streets, transmission lines, pipe, and stream reaches.

Junctions are street intersections, fuses, switches, service taps, and the confluence of stream reaches.

Edges connect at junctions

Railroad schedules its trains to efficiently link with intermodal container trucks.

- **A parcel delivery service optimizes its package delivery on a street system.**

- An electric utility locates where power outages originate based on telephone calls received from affected customers.
- An environmental agency analyzes water samples collected from streams to trace contaminant flow.
- A regional transportation agency uses traffic data to plan future highway construction.
- A school district finds optimum bus routes to pick up children and deliver them to school.
- A driver uses a mapping system with a GPS receiver mounted in the car to find the best way to get to a destination.

THE GEOMETRIC NETWORK

A geometric network is a collection of features that comprise a connected system of *edges and junctions.*

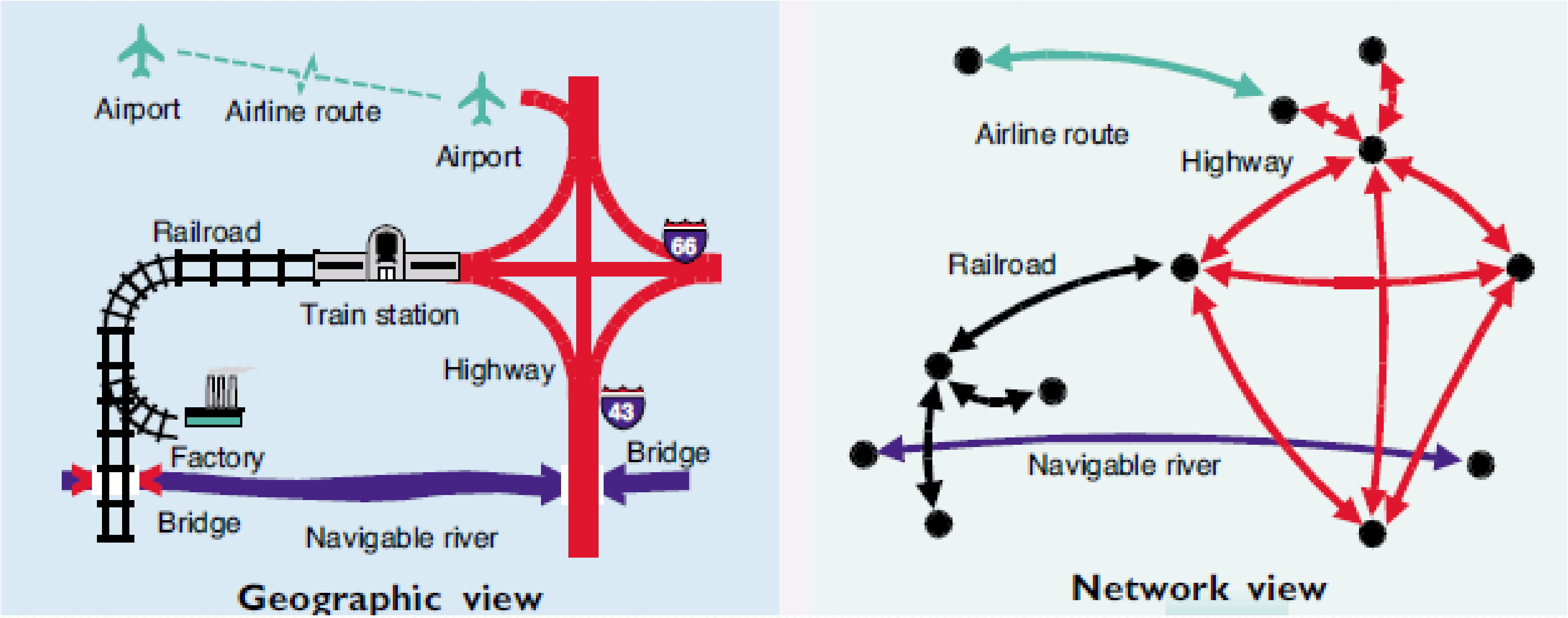
***An edge has two junctions and a junction* can be connected to any number of edges.**

network features: simple junction feature, complex junction feature, simple edge feature, or complex edge feature

THE LOGICAL NETWORK

Like a geometric network, a logical network is a collection of connected edges and junctions.

The key difference is that a logical network does not have coordinate values.

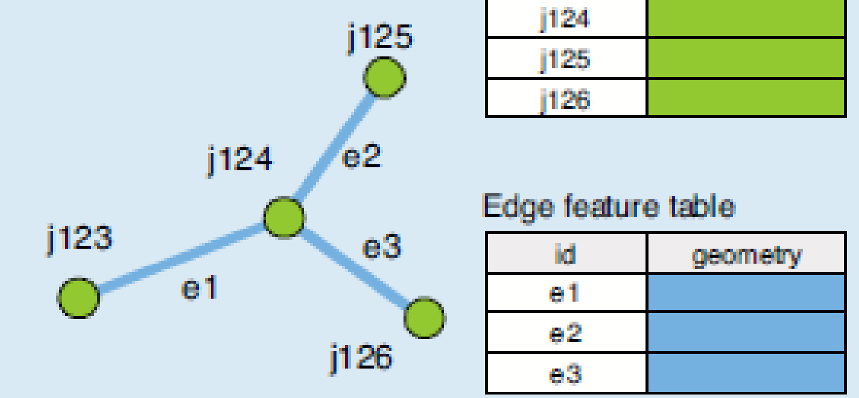

Airport
Airline route
Airport
Railroad
Train station
66
Highway
43
Factory
Bridge
Bridge
Navigable river
Geographic view
Airline route
Highway
Railroad
Navigable river
Network view

HOW FEATURES CONNECT

The center-piece of a logical network is the connectivity table, which describes how network elements are connected.

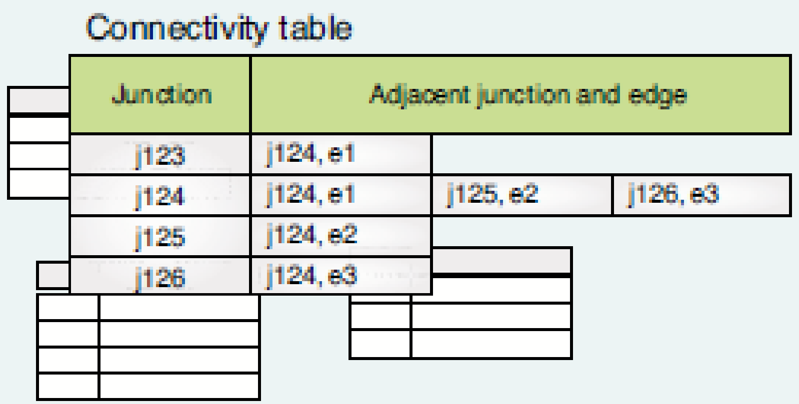
For every junction in the network, the connectivity table lists the adjacent junctions and edges—junctions at the other end of the connected edge.

Geometric network



A geometric network contains the features that participate in a network. Feature classes contain either edge features or junction features.

Logical network



A logical network contains the connectivity of the network. The connectivity table lists all the adjacent junctions to a given junction, along with the edge that connects them.

CONNECTIVITY RULES

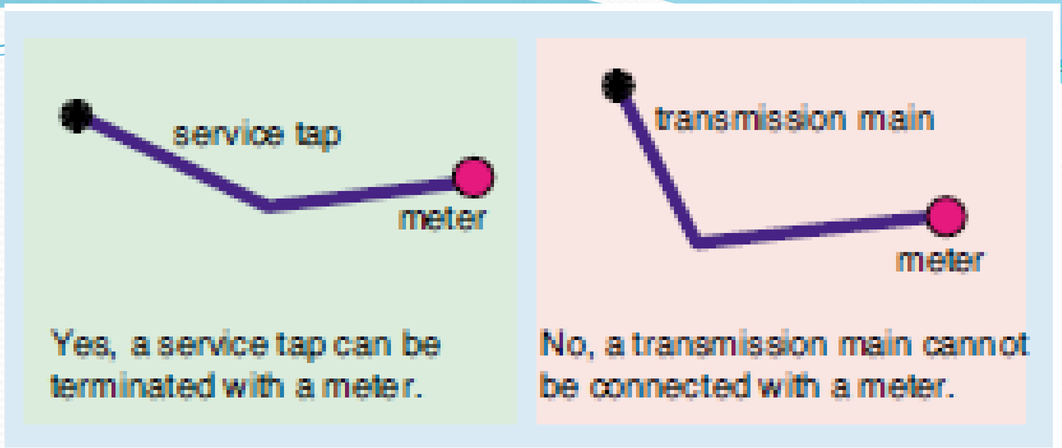
In most networks, not all edges can connect to all other junctions. Also, not all edges can connect to all other edges through a specified junction.

For example, a hydrant lateral in a water network can connect to a hydrant, but not to a service lateral.

Similarly, a 10-inch transmission main can only connect to an 8-inch transmission main through a reducer.

Edge–junction rule

This rule constrains which types of junctions can connect to a type of edge.

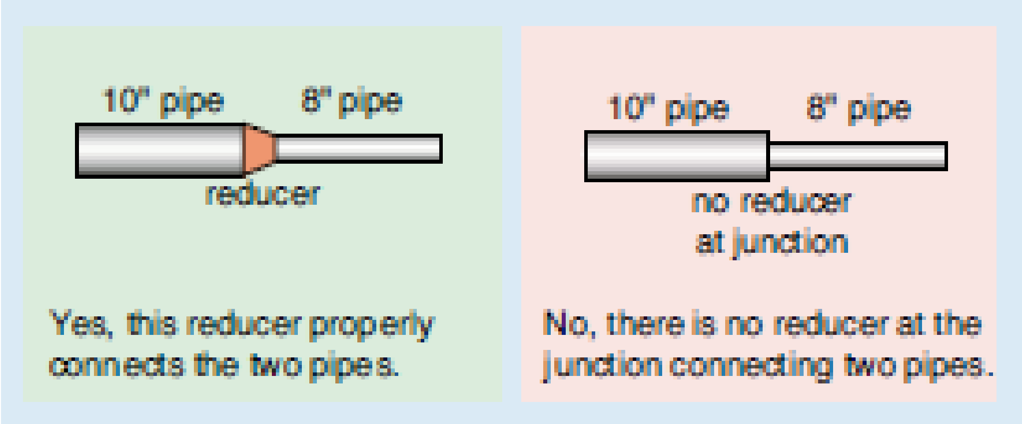


Meters can only connect to low-voltage lines.

Edge–edge rule

This rule establishes which combinations of edge types can connect through a given junction.

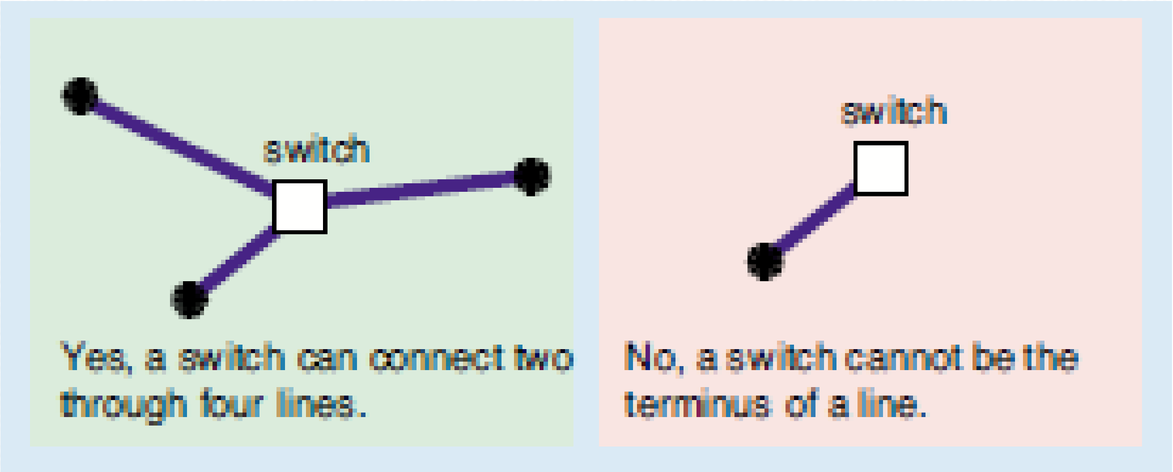
Two pipes of different diameters can be connected only through a properly sized reducer



Edge–junction cardinality

This rule lets you restrict the count (cardinality) of edges that connect at a junction.

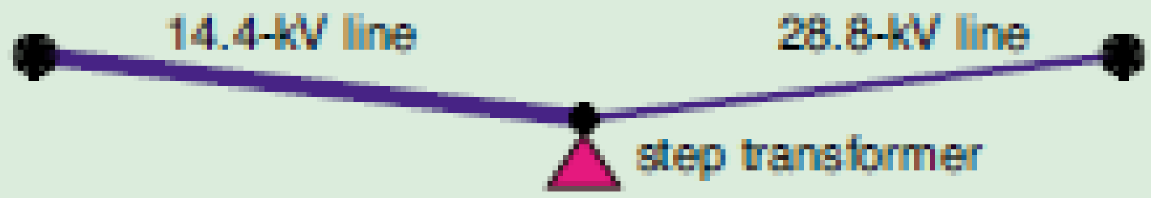
A certain type of switch might be designed to accept between two and four lines. You can precisely define the acceptable range of lines that can be connected at a junction



Default junction type

When you connect one type of edge to another, you can specify a default junction type to be inserted.

When a 14.4-kV line is added to an end-junction of a 28.8-kV line, a step-down transformer with the correct electrical ratings is assigned to the junction



When you connect a 14.4-kV electric distribution line to a 28.8-kV line, a properly sized transformer can be automatically inserted.

NETWORK FLOW

Networks are of two type: utility network and transportation network.

In a *transportation network*, depends on the need will be changed "will of their own."

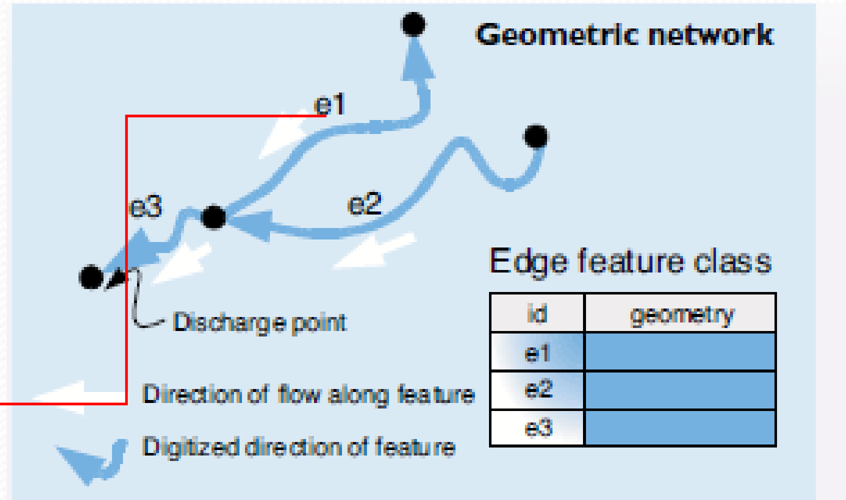
In a *utility network*, the commodity that flows through the network—water, electricity, oil—has no will of its own.

The network imposes flow direction by its configuration of sources, sinks, and switches

Logical network

Edge element table

Feature class	Feature ID	Sub-ID	Element ID	Flow direction
1	e1	1	10	Against
1	e2	1	11	With
1	e3	1	12	With



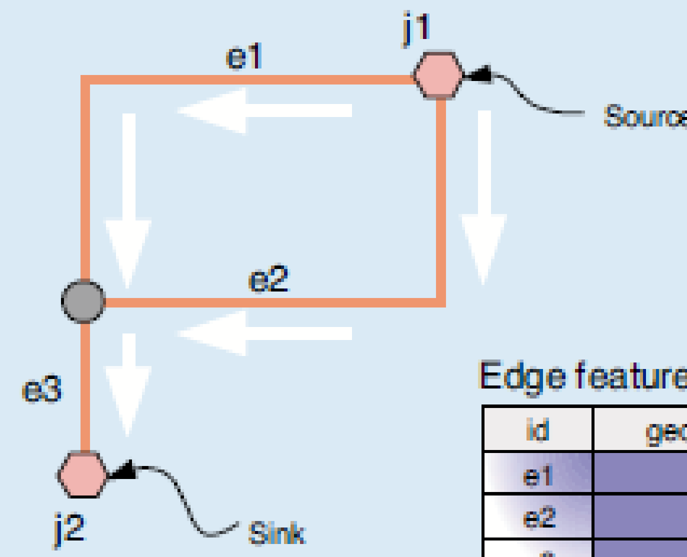
SOURCES AND SINKS

In a utility network, sources and sinks are used in determining flow direction.

Source is a junction from which a commodity flows, such as a well-head pump.

Sink is a junction where all commodity flow terminates, such as a wastewater treatment plant.

Geometric network



Edge feature class

id	geometry
e1	
e2	
e3	

Flow junctions feature class

id	Ancillary role	geometry
j1	Source	
j2	Sink	

Junction features can have an ancillary role of source, sink, or neither. The role is stored in an attribute of the feature class, which is accessed by the establish flow direction method.

DISABLED FEATURES

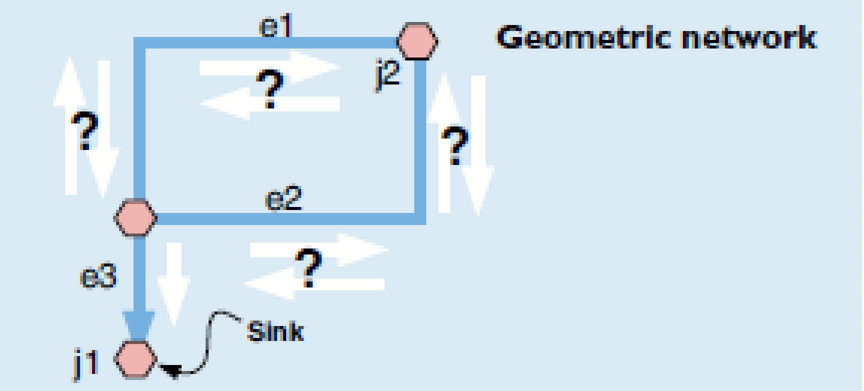
All features participating in a network have an enabled/disabled state.

Nothing flows into or out of the feature.

Disabled features are useful for representing open electrical switches or closed valves.

INDETERMINATE FLOW

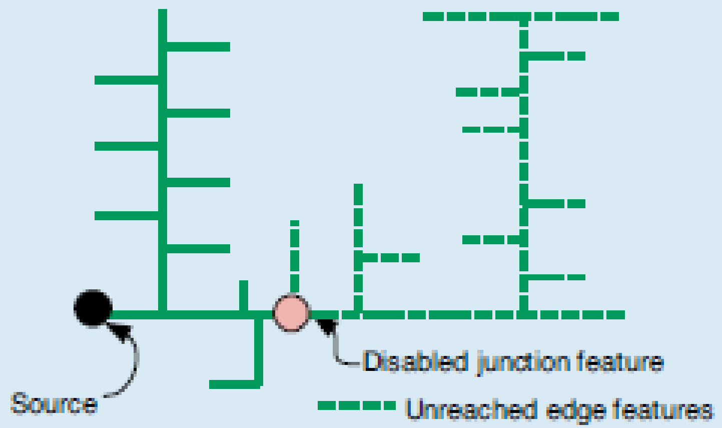
Indeterminate flow occurs when the established flow direction method cannot determine in which direction, commodities flow in a network



It may not be possible to determine the direction of flow given a configuration of sources, sinks, and enabled features. In this example, it is impossible to determine the flow through edges e1 and e2 because they form a cycle. Changing junction j2 to a source would break the cycle.

UNINITIALIZED FLOW

When a flow is isolated because the edges are disconnected from the rest of the network (that has flow), flow is said to be uninitialized.



When establishing flow direction, edge features may be unreached because they are disconnected from the rest of the network. In this example, the unreached edges are disconnected because one of the junction features–a valve–is disabled.