

What is GIS?

"Geographic Information System"

□ A new technological tool



□ A Geospatial information handling strategy

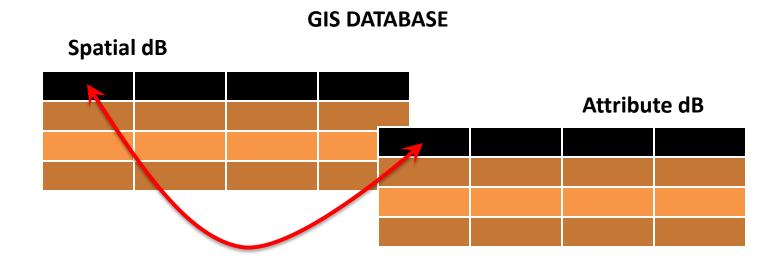
The objective is

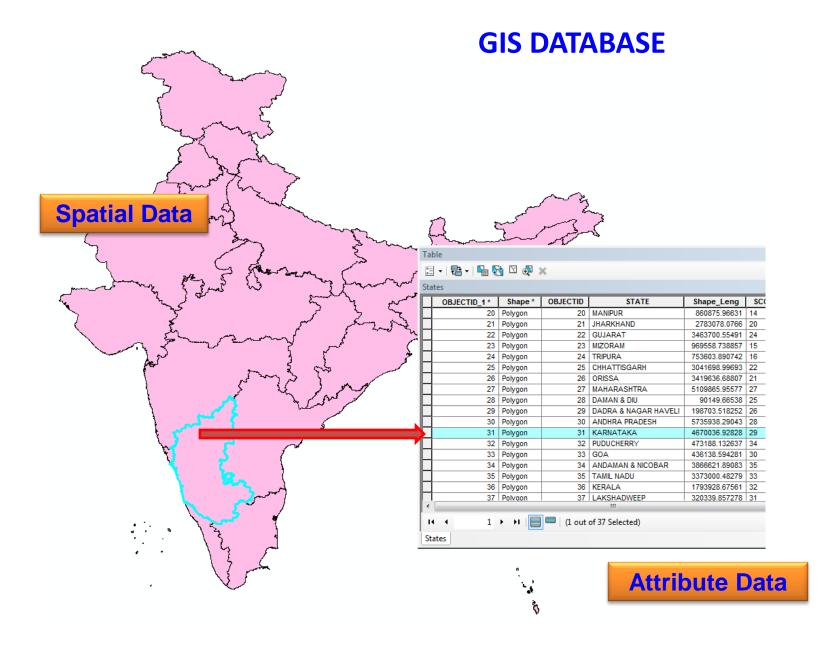
"to improve overall decision making".

What is GIS?

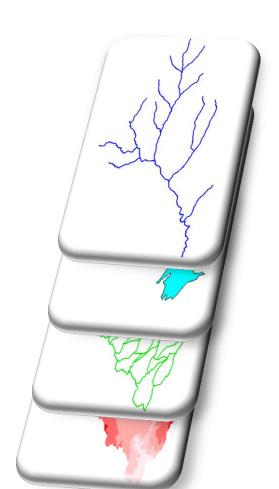
A GIS does not hold maps or pictures.

It holds a database from which the data can be displayed in a variety of views.

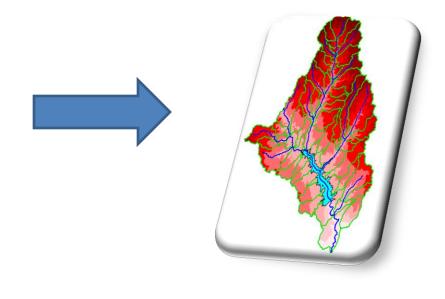




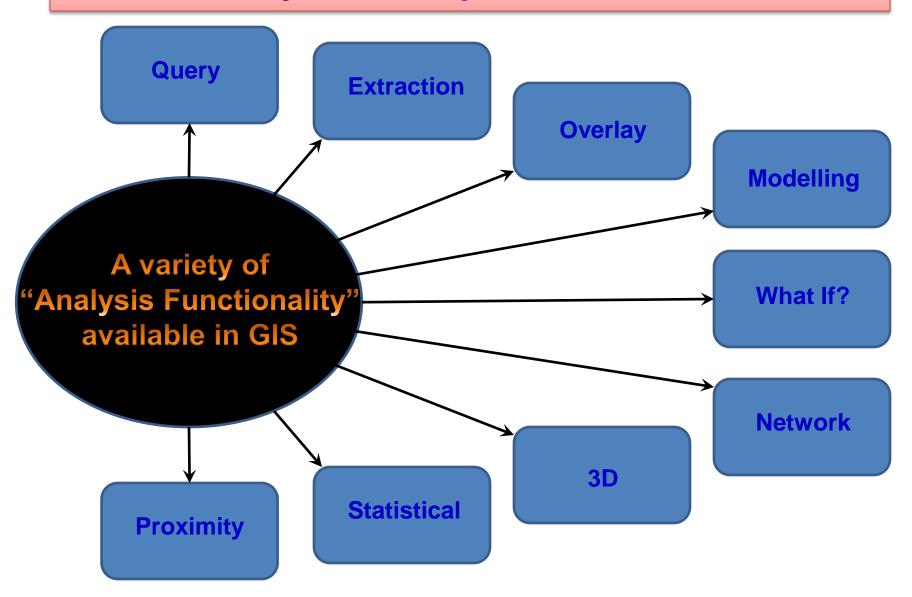
AN INTEGRATED VIEW



Layers are integrated using explicit location on the earth's surface, thus geographic location is the organizing principal.



Why GIS is a powerful tool?



GIS became a powerful technological tool because of its analysis techniques. A variety of tools are available for analysis in GIS. However, the techniques are different for different spatial models.

Raster analysis tools are different from Vector analysis tools.

The appropriate analysis tool(s) should be used depending on the spatial data model.

"Geoprocessing" is a GIS operation used to manipulate spatial data.

The whole set of **Vector Analysis Tools** are categorised into following groups.

- Query Techniques
- Extraction Techniques
- Proximity Techniques
- Overlay Techniques
- Statistical Techniques

Query Techniques

Attribute Query

Popular in DBMS, **Structured Query Language (SQL)** is a powerful language used to define one or more criteria that can consist of attributes, operators, and calculations for **selecting features and attribute records** from GIS database.

General syntax of a SQL statement is

SELECT <attribute name(s)> **FROM WHERE** <condition statement>

Attribute Query

SQL Operators: Used to frame out conditional statement in SQL.

Arithmetic operators

```
Addition ( + )
Subtraction ( - )
Multiplication ( * )
Division ( / )
Modulus ( % )
```

Logical operators

NOT "STATE NAME" = 'Goa'

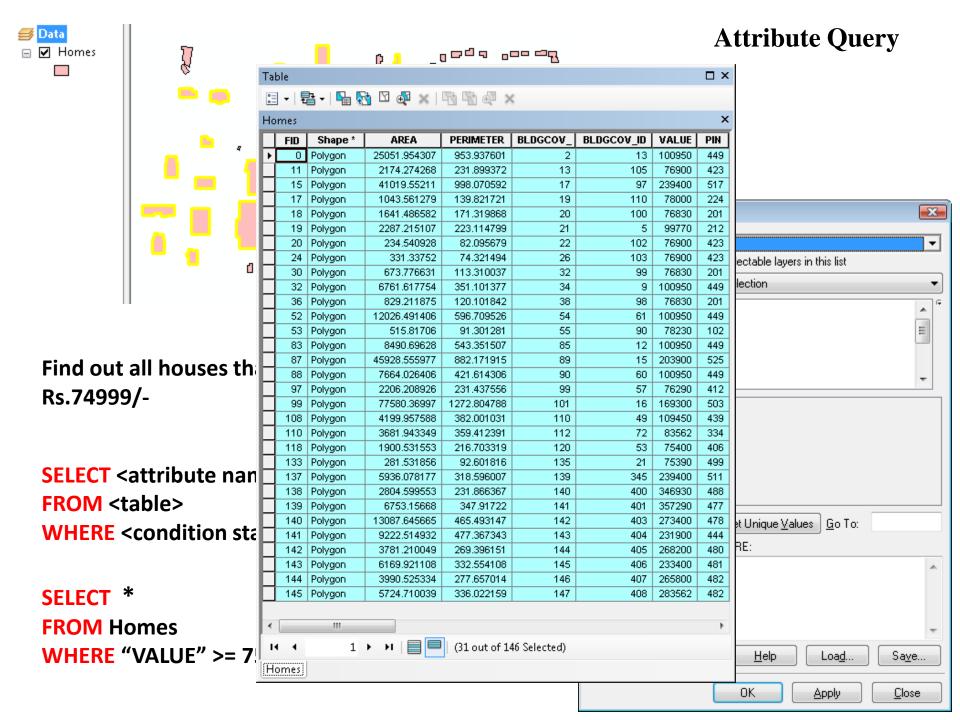
```
AND
"AREA" > 1500 AND "Type" = 'Residential'

OR
"AREA" > 1500 OR "Type" = 'Residential'

XOR
"AREA" > 1500 XOR "Type" = 'Residential'

NOT
```

```
Relational / Comparison operators
Equal ( = )
Not equal ( != or <> )
Greater than ( > )
Greater than or equal ( >= )
Less than (<)
Less than or equal ( <= )
[NOT] EXISTS
EXISTS (SELECT * FROM roads WHERE "Type" = 'NH')
[NOT] IN
"STATE NAME" IN ('Bihar', 'Goa', 'Odisha', 'Asom')
"STATE NAME" IN (SELECT "STATE NAME" FROM states WHERE
"POP" > 5000000)
IS [NOT] NULL
```



Query Techniques

Spatial Query

A **spatial query** is a special type of database query supported by spatial databases. This type of query lets you

select features from one layer based on their location relative to features in the other layer.

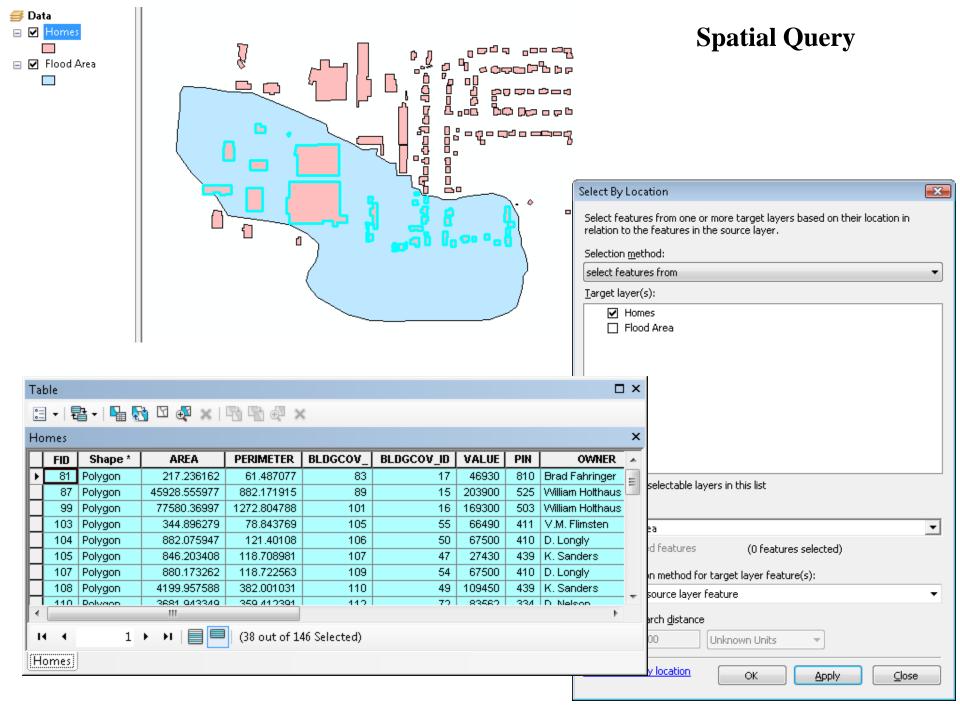
They allow for the use of geometry data types such as points, lines and polygons and that these queries consider the spatial relationship between these geometries.

For instance, if you want to know how many homes were affected by a recent flood and you mapped the flood area, you could select all the homes that are within the flooded area.

Target layer: The layer in which features will be selected.

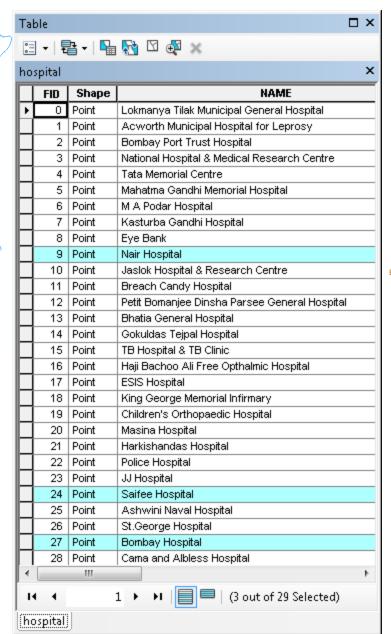
(Homes layer)

Source / Selector layer: The layer that is used to determine the selection based on its topological relationship to the target layer (Flood Area layer).

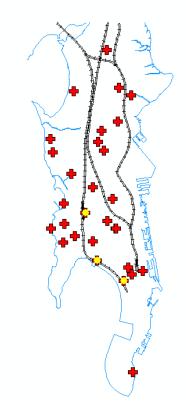


Layers ☐ ✓ railwayline

SPATIAL QUERY







SPATIAL SELECTION METHODS

Spatial Selection Method	Meaning	
Intersect	Select any feature in the target layer that geometrically shares a common part with the source feature(s)	
Are within a distance of	First creates a buffer(s) with a size equal to the distance specified around the source feature(s), then selects all the features intersecting the buffer	
Contain	The geometry of the source feature must fall inside the geometry of the target feature including its boundaries.	
Completely Contain	Source feature must fall inside the geometry of the target feature, excluding the target's boundaries (the boundaries cannot touch).	
Are within	The geometry of the target feature must fall inside the geometry of the source feature including its boundaries.	
Are completely within	The target feature must fall within the geometry of the source feature excluding the source's boundaries (the boundaries cannot touch). This is the reverse operator from Completely contain.	
Have their centroid in	A target feature will be selected by this operator if the centroid of its geometry falls into the geometry of the source feature or on its boundaries.	
Share a line segment with	Target layer features share a line segment with the Source layer feature.	
Touch the boundary of	Target layer features touch the boundary of the Source layer feature.	
Are identical to	Target layer features are identical to the Source layer feature. Two features are considered identical if their geometries are strictly equal.	
Are crossed by the outline of	Target layer features are crossed by the outline of the Source layer feature.	

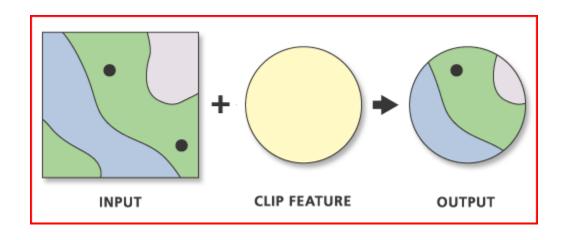
Feature Extraction Techniques

Extraction tools are used to create a subset of spatial database

- **CLIP**
- **ERASE**
- **❖** SPLIT
- **❖** SELECT
- *** TABLE SELECT**

- CLIP
- ERASE
- ❖ SPLIT
- **SELECT**
- ***** TABLE SELECT

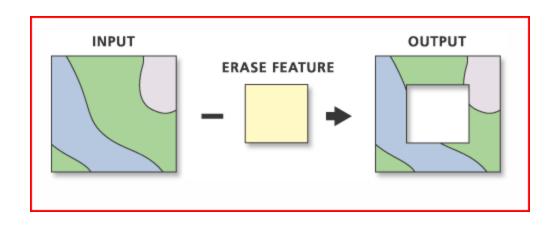
Clip is used to cut out a portion of a layer using one or more of the features in another feature class as a "cookie cutter". This is useful for creating a new geographic subset out of larger feature class.



INPUT FEATURES	CLIP FEATURES	OUTPUT FEATURES
Polygons	Polygons	Polygons
Lines	Lines / Polygons	(Coincident) Lines
Points	Points / Lines / Polygons	(Coincident) Points

- ❖ CLIP
- ERASE
- ❖ SPLIT
- SELECT
- TABLE SELECT

Only those portions of the Input Features falling outside the Erase Features outside boundaries are copied to the Output Feature Class.



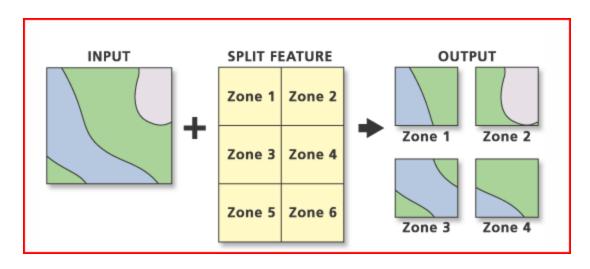
A **polygon erase** feature can be used **to erase polygons, lines, or points** from the input features

A **line erase** feature can be used **to erase** <u>lines or points</u> from the input features

A **point erase** feature can be used **to erase points** from the input features.

- CLIP
- ERASE
- SPLIT
- SELECT
- TABLE SELECT

The spatial extraction of features by clipping portions of the input feature class into multiple feature classes.

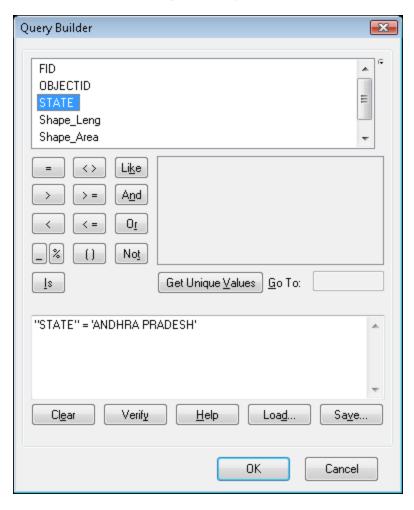


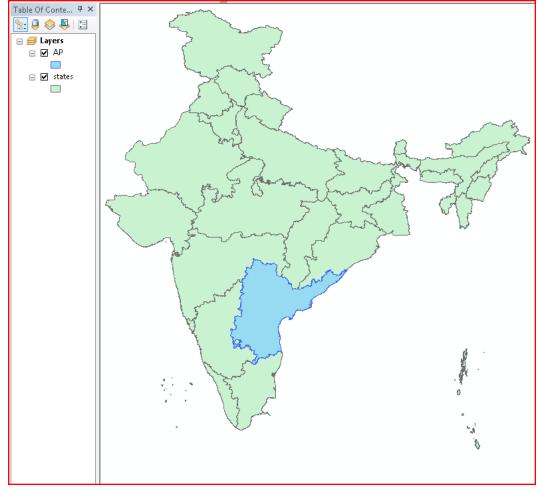
The INPUT features are split into four OUTPUT feature classes based on four of six overlaying SPLIT FEATUREs. These six split features correspond to six unique split field (character data type) values.

The character field used to split the input features. This field's values identify the split features used to create each output feature class. The split field's unique values provide the output feature classes' names.

- CLIP
- ERASE
- ❖ SPLIT
- SELECT
- TABLE SELECT

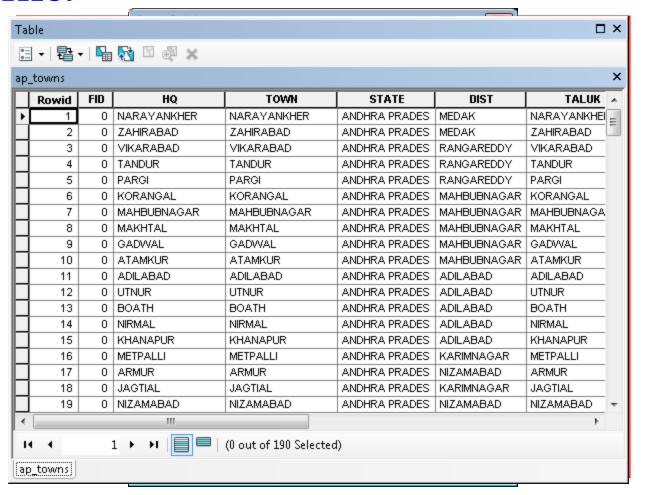
Extracts features from an input feature class or input feature layer, using a **SQL expression** through **Query Builder** and stores them in an output feature class.





- ❖ CLIP
- ERASE
- ❖ SPLIT
- ❖ SELECT
- TABLE SELECT

Extracts **selected attributes** from an input table based on an attribute query and stores them in the **output table**. Remember no spatial features will be extracted to output.



Proximity Analysis Techniques

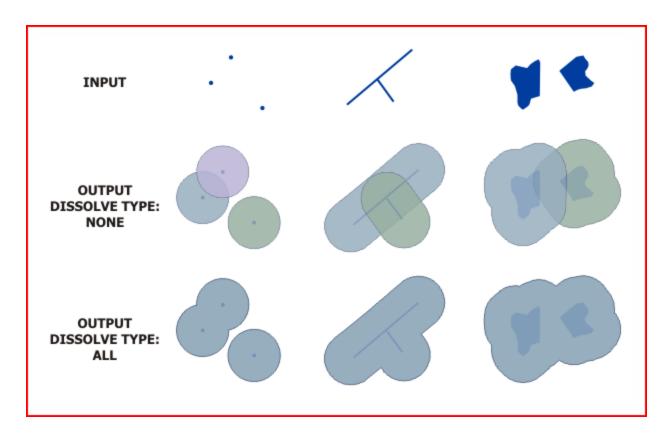
The Proximity techniques are used to determine the proximity (closeness or distance) of features within one or more feature classes or between two feature classes.

- **BUFFER**
- **❖ MULTIPLE BUFFERS**

Creates buffer polygons around input features to a specified distance. The dissolve can be performed to remove overlapping buffers.

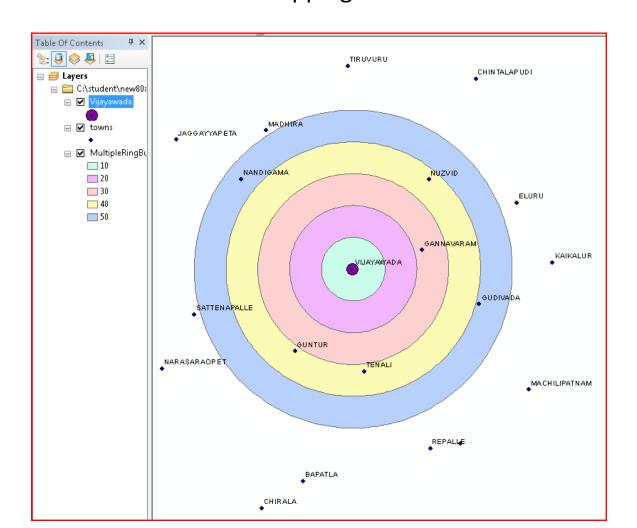
Negative buffer distance means inside the polygon.

Variable buffer distance can be specified as the name of one of the **attribute** in the table.



- BUFFER
- MULTIPLE BUFFERS

Creates **multiple buffers** at specified distances around the input features. These buffers can optionally be merged and dissolved using the buffer distance values to create non-overlapping buffers.



Overlay Analysis Techniques

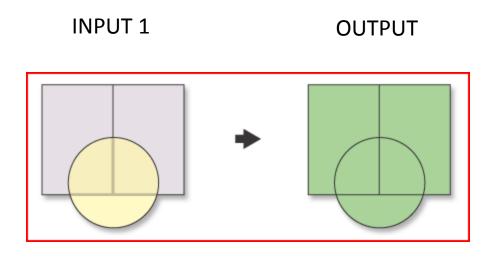
The Overlay toolset contains tools to overlay multiple feature classes to combine, erase, modify, or update spatial features in a new feature class. New information is created when overlaying one set of features with another.

- **UNION**
- **❖ INTERSECT**
- ***** IDENTITY
- **SYMMETRICAL DIFFERENCE**
- **UPDATE**

- UNION
- INTERSECT
- IDENTITY
- SYMMETRICAL DIFFERENCE
- **UPDATE**

Computes a geometric intersection of the Input Features. All features will be written to the Output Feature Class with the attributes from the Input Features, which it overlaps.

All input 1 and input 2 (union layer) must have **polygon geometry**.



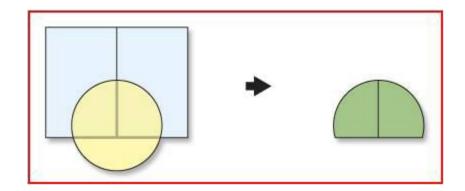
INPUT 2 (Union Layer)

- UNION
- INTERSECT
- ***** IDENTITY
- SYMMETRICAL DIFFERENCE
- **UPDATE**

The Intersect tool calculates the geometric intersection of any number of feature classes and feature layers. The features or portion of features that are common to (intersect) all inputs will be written to the Output Feature Class.

INPUT 1	OUTPUT

Input 1	Input 2	Output
Point	Polygon	Point
Arc	Polygon	Arc
Polygon	Polygon	Polygon



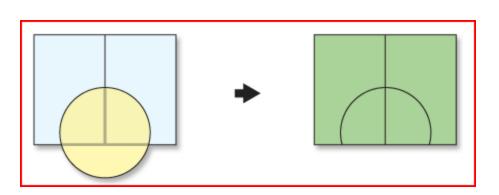
INPUT 2 (Intersect Layer)

- UNION
- INTERSECT
- **❖** IDENTITY
- SYMMETRICAL DIFFERENCE
- **UPDATE**

Computes a geometric intersection of the Input Features and Identity Features. The Input Features or portions thereof that overlap Identity Features will get the attributes of those Identity Features.

The Input Features must be point, multipoint, line, or polygon. The Identity Features must be polygons.

Input 1	Input 2	Output
Point	Polygon	Point
Arc	Polygon	Arc
Polygon	Polygon	Polygon



OUTPUT

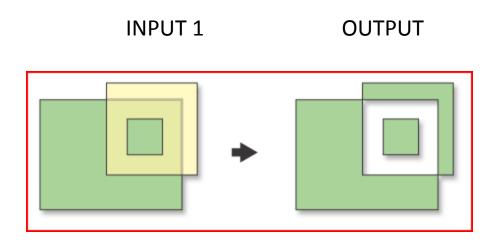
INPUT 2 (Identity Layer)

INPUT 1

- UNION
- INTERSECT
- IDENTITY
- SYMMETRICAL DIFFERENCE
- ❖ UPDATE

Features or portions of features in the input and update **features which do not overlap** will be written to the Output Feature Class.

Both inputs must have **polygon** geometry.

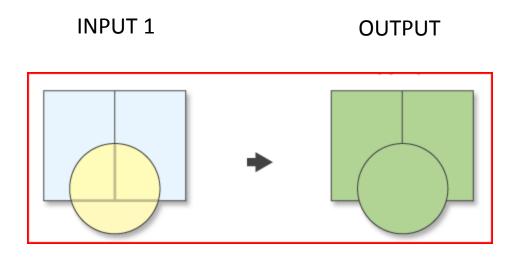


INPUT 2

- UNION
- INTERSECT
- **❖** IDENTITY
- SYMMETRICAL DIFFERENCE
- **❖** UPDATE

The attributes and geometry of the Input Features are updated by the Update Features or layer that they overlap.

Both inputs must have **polygon** geometry.



INPUT 2 (Update Layer)

Statistical Analysis Techniques

Statistical analysis in GIS is of two types:

- Non-spatial statistical analysis
- Spatial statistical analysis

Non-spatial statistics are used to analyze attribute (numerical) data associated with features.

E.g., Understanding the distribution of values for a particular attribute by summarizing the data.

Charts and graphs, such as a histogram or Q-Q plots are also used for analyzing non-spatial data. In all cases, only the values are analyzed.

The locations of the features with which the values are associated and any spatial relationships between the features are not considered.

Spatial statistics, on the other hand, focus on the spatial relationships between features.

Spatial statistics are used for

- Analyzing Patterns
- Mapping Clusters
- Measuring Geographic Distributions
- Modelling Spatial Relationships

Statistical functions analyze the underlying data and give you a measure that can be used to confirm the existence and strength of the pattern.

Geostatistics

Geostatistics are a type of **spatial statistics**. Kriging, for example, is a very powerful geostatistical technique that goes beyond interpolation, looking not only at nearby features to predict values where you don't have sample data, but actually **utilizing spatial relationships** to give you stronger, more accurate predictions.

Geostatistics are ideal for analyzing and predicting the values associated with nearly any kind of spatially continuous phenomena.

The **semivariogram** is a statistical measure of the rate of change with distance, for attributes that vary in space.

The semivariogram depicts the spatial autocorrelation of the measured sample points. Once each pair of locations is plotted, a model is fit through them.

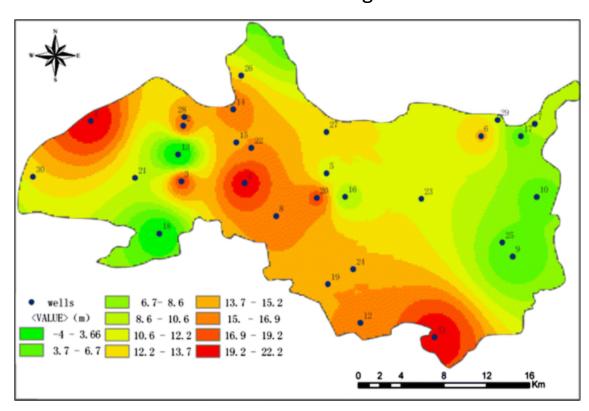
The Steps in a Geo-statistical Study

- Exploratory Spatial Data Analysis
 (Spatial Data Variability, Spatial Data Dependence, Global Trends & Outliers)
- 2) Structural data analysis (Calculation and Modeling of Variograms)
- 3) Making predictions (kriging or simulations)
- 4) Analysis of Result

"Based on the geo-statistical theory data of **30** groundwater level observation wells were used to estimate the decline of groundwater level in study area (Beijing piedmont). Seven different interpolation methods (inverse distance weighted interpolation, global polynomial interpolation, local polynomial interpolation, tension spline interpolation, ordinary Kriging interpolation, simple Kriging interpolation and universal Kriging interpolation) were used for interpolating groundwater level between 2001 and 2013. Cross-validation, absolute error and coefficient of determination (R²) was applied to evaluate the accuracy of different methods.

Groundwater level drawdown during 2001 and 2013.

"The result shows that simple Kriging method gave the best fit".

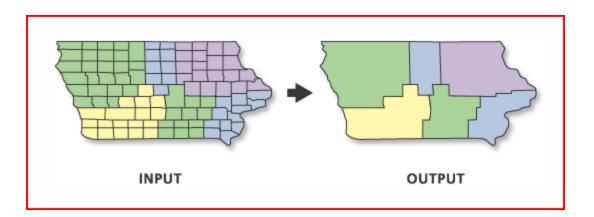


Tools for Generalization

Tools in the Generalization toolset can be used to aggregate or eliminate features.

DISSOLVE

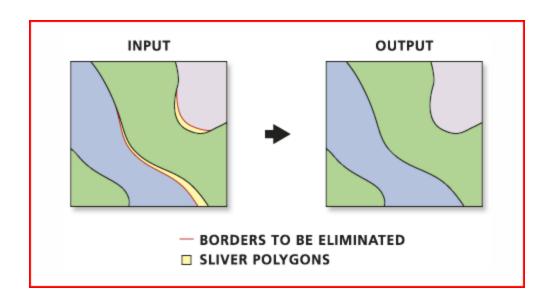
Aggregates features based on specified attributes.



Tools for Generalization

ELIMINATE

Eliminates polygons by merging them with neighboring polygons that have the largest area or the longest shared border. Eliminate is often **used to remove small sliver polygons** that are the result of overlay operations.

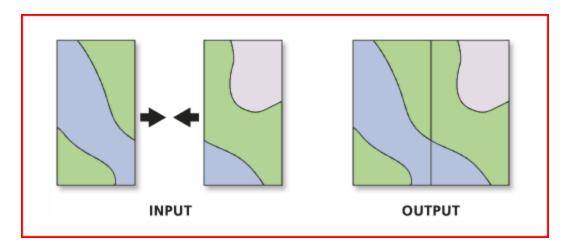


General Tools

MERGE

Combines multiple input datasets of the same data type into a **single, new output** dataset. This tool can combine point, line, or polygon feature classes or tables. (Input data sources need not be adjacent; overlap is allowed.)

All input datasets must be of the same type.

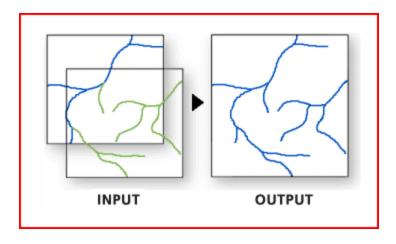


Edgematching is used to align features along the edges of adjacent layers.

General Tools

APPEND

Appends multiple input datasets **into an existing target dataset**. Input datasets can be point, line, or polygon feature classes, tables, rasters, annotation feature classes, or dimensions feature classes.





Raster vs. Vector Analysis Techniques

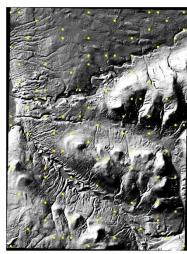
- Analysis using vector model is different from that of raster
 - some operations are more accurate
 - some operations are slower
 - some operations are faster
- Vector GIS normally uses topological properties
- Raster GIS uses map algebra, pixel counts, etc.

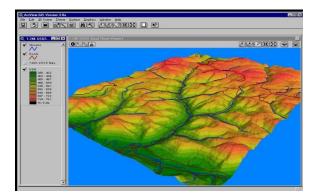
Why to use Raster GIS?

Raster GIS is often used because:

- Raster is better suited for spatially continuous data like elevations
- Raster is better for creating visualizations and modeling environmental phenomena
- Other continuous data may include: pH, air pressure, temperature, salinity, etc.
- Raster data is a simplified realization of the world, and allows for fast and efficient processing







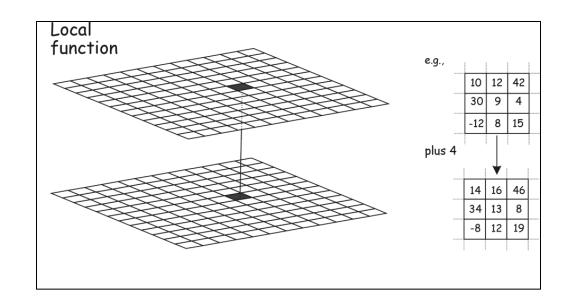
A raster GIS performs geoprocessing tasks on a grid based realization of the world

Operations on Raster Layers

- Local Operations (Cell by cell)
- Focal Operations (Neighborhood)
- Zonal Operations (Region-specific)
- Global Operations (Entire layer)

Local Operations

- New layer is a function of two or more input layers
- Output value for each cell is a function of the values of the corresponding cells in the input layers
- Neighboring or distant cells have no effect



Focal Operations

Focal functions apply one calculation to all input grid cells within a "focus"

input grid					
24	50	7	41	32	
30	27	8	22	39	
14	16	21	16	41	
38	6	44	8	7	
36	7	32	30	32	
38	20	28	28	21	
32	2	13	35	49	

output grid					
32.8	24.3	25.8	24.8	33.5	
26.8	21.9	23.1	25.2	31.8	
21.8	22.7	18.7	22.9	22.2	
19.5	23.8	20.0	25.7	22.3	
24.2	27.7	22.6	25.6	21.0	
22.5	23.1	21.7	29.8	32.5	
23.0	22.2	21.0	29.0	33.3	

focal mean

$$(27 + 8 + 22 + 16 + 21 + 16 + 6 + 44 + 8) / 9 = 18.7$$

Zonal Operations

Zonal functions apply one calculation to all input grid cells within each zone

input grid

53	57	33	10	14
78	31	12	22	55
32	9	9	85	26
6	54	33	85	94
75	25	76	49	27
48	16	67	23	89
36	46	82	97	74
45	86	44	42	35

zone grid

1	1	1	2	2
1	1	1	2	2
1	1	1	2	2
5	5	1	2	2
5	5	1	2	2
5	5	10	91	91
5	5	10	91	91
5	5	10	10	10

423	423	423	467	467
423	423	423	467	467
423	423	423	467	467
437	437	423	467	467
437	437	423	467	467
437	437	270	283	283
437	437	270	283	283
437	437	270	270	270

zonal sum

$$(53 + 57 + 33 + 78 + 31 + 12 + 32 + 9 + 9 + 33 + 76) = 423$$

Global Operations

Global functions apply a single calculation to all input grid cells

input grid					
12	49	20	46	26	
6	37	20	46	28	
49	45	38	15	6	
23	5	35	17	17	
22	48	3	45	23	
42	14	23	32	42	
31	23	41	2	40	

oucput grid				
-0.537	-0.954	0.913	0.902	0.763
-0.279	-0.644	0.913	0.902	0.271
-0.954	0.851	0.296	0.65	-0.279
-0.846	-0.959	-0.428	-0.961	-0.961
-0.009	-0.768	0.141	0.851	-0.846
-0.917	0.991	-0.846	0.551	-0.917
-0.404	-0.846	-0.159	0.909	0.745

global sine

$$sin(12) = -0.537$$

Raster Overlay Operations

(MAP ALGEBRA)

Map algebra is a cell by cell combination of raster layers using mathematical operations

- Unary one layer
- Binary two layers

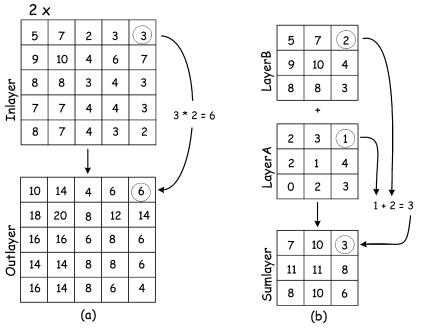
All basic mathematical operations can be used

- Addition
- Subtraction
- Multiplication
- Division
- Max
- Min
- **–** ...

virtually any mathematical operation

Strong analytical functions

- Outgrid = grid1 + grid2
- Outgrid = grid1 * 2
- Outgrid = sin(grid1)
- Outgrid = costallocation(sourcegrid, costgrid, accumgrid, backgrid)
- Outgrid = con(>5 (ingrid1),0,ingrid1)
- Outgrid = select(grid1, 'VALUE = 10')
- slp dem = slp grid * dem



© Paul Bolstad, GIS Fundamentals

Raster Overlay Operations

(MAP ALGEBRA)

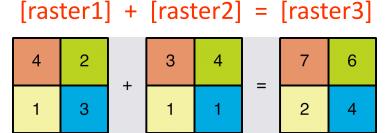
- Output cell value is the result of an arithmetic operation on the input layers
- e.g., if a and b are input layers and c is an output layer,

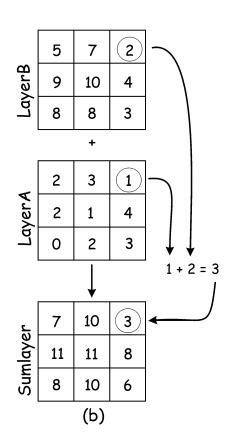
$$c = a + b$$

 $c = a * b$

.

Also, any kind of function such as Average, Sum, Min, Max, Standard Deviation, etc.

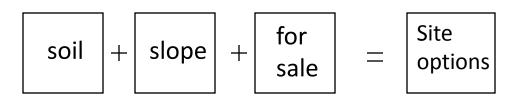




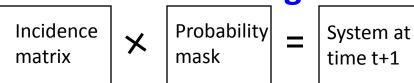
Raster Modelling

- Relating multiple rasters
- Processes may be:
 - Local: one cell only
 - Neighborhood: cells relating to each other in a defined manner
 - Zonal: cells in a given section
 - Global: all cells

Suitability modelling



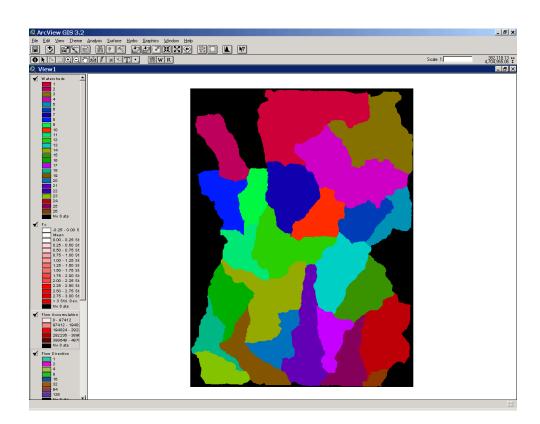
Diffusion Modelling



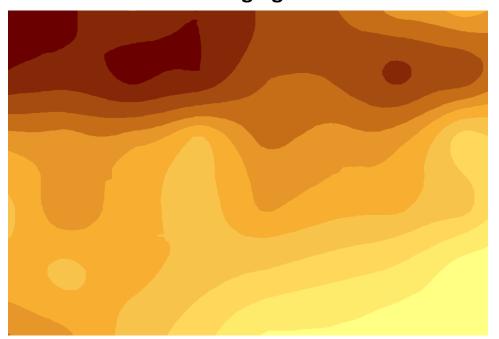
Connectivity Modelling

Watershed Delineation

- > Fill Sinks in DEM
- Compute Flow Direction
- Compute Flow Accumulation
- Generate Watershed

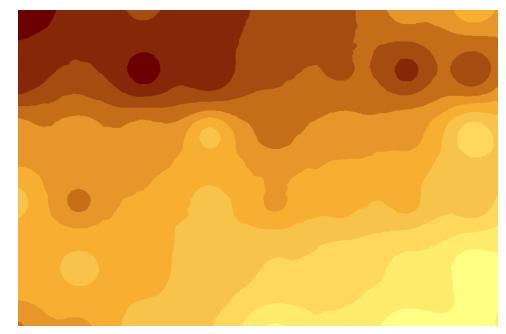


Kriging



Surface Generation (interpolation)

IDW

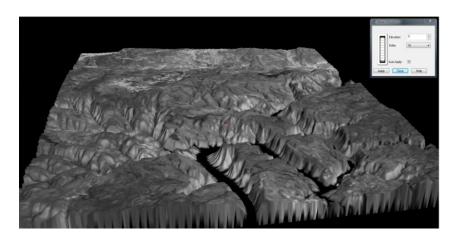


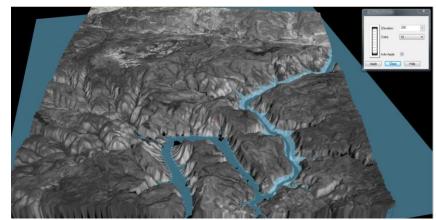
Flood Effect Simulation

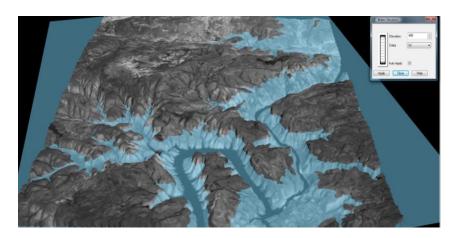
Water elevation: Normal

Water elevation: 200 m

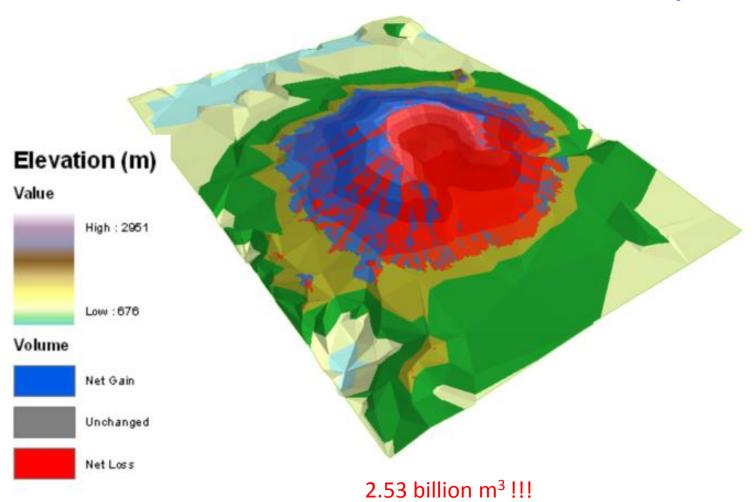
Water elevation: 400 m







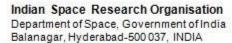
Volumetric Analysis





T.S. Viswanadham

Manager, TOFM / TEOG / MSA
National Remote Sensing Centre



Tel (Off.): +91 40 2388 4567 Fax: +91 40 2387 8816

Email: viswanadham_ts@nrsc.gov.in