Intro to Spatial Science & Technology

Lec 9: Spatial Data Models

Dr. K. S. Rajan
Lab for Spatial Informatics, IIIT Hyderabad
Feb 6th 2024

1

- Lets revisit the queries we talked about last class.
- Definition of a GI System

- First step,
  - What is a Spatial Entity / Object ?
  - What is a Spatial Phenomena ?

2

What are Spatial Objects?

- Building
- Lake
- …
- Car
- Mobile

- Location
- Occupy some space
- Interact/Affect the neighbours
- …

What are Spatial Phenomena?

- Disease Spread
- Snow Avalanche
- Flooding
- …..

- What is important in the phenomena?
  - objects/elements
  - their interactions and its types
  - Environment in which it functions
  - …

Riddle

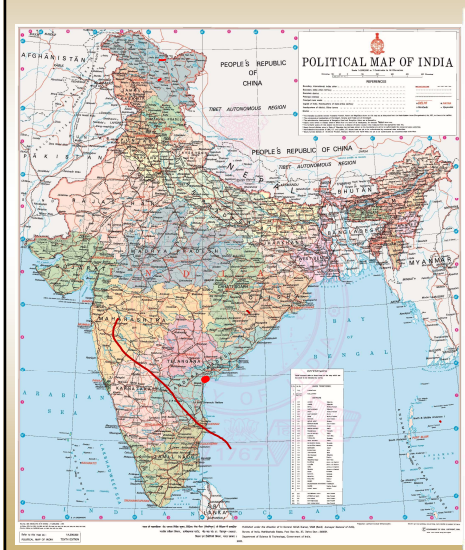
What has
mountains but no trees,
rivers but no water,
cities but no buildings ?

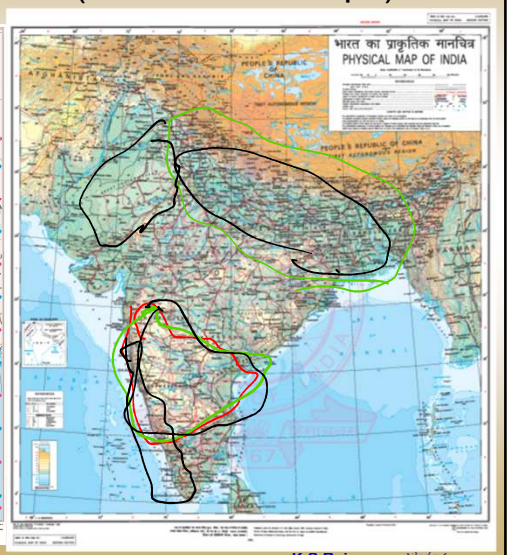
Map shows
- Features on the earth surface
- represented through symbology
- indicates distance, area, shape, direction

5

Map Sheets

- shows features of India (Source: SoI Maps)



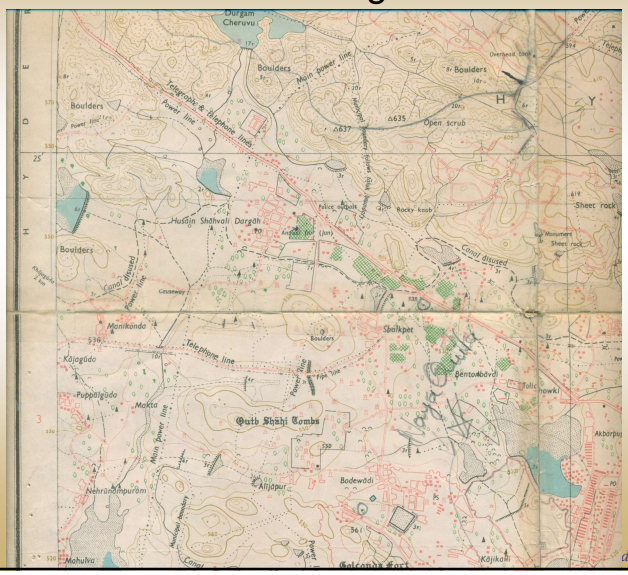


6

CS4.408 Spatial Informatics

Sol Toposheet view

- of Golconda Fort and its neighbourhood



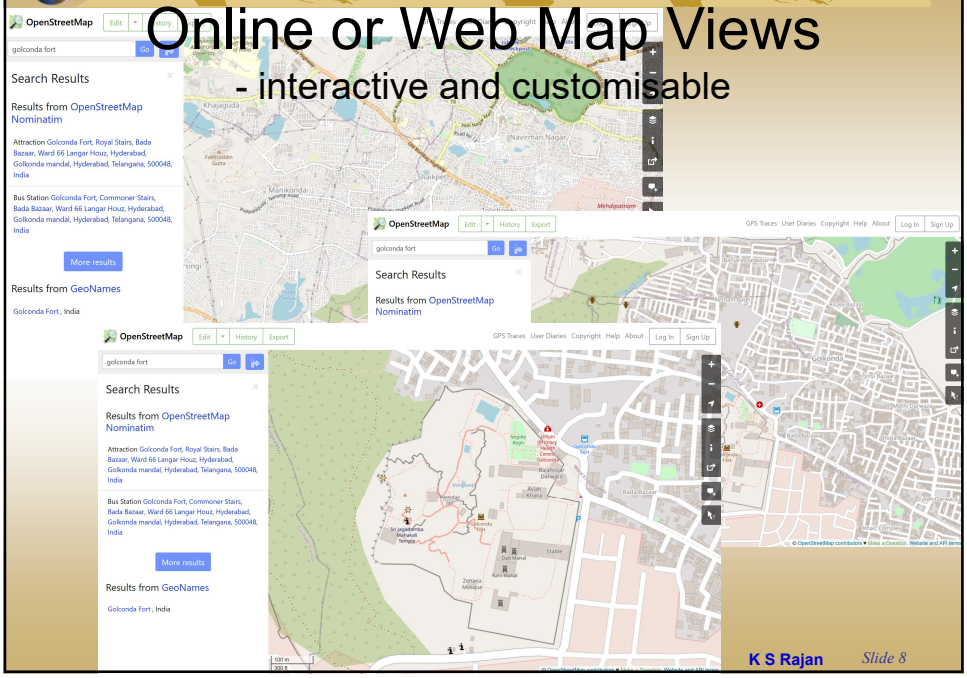
de 7

7

CS4.408 Spatial Informatics

Online or Web Map Views

- interactive and customisable



K S Rajan Slide 8

8

GeoSpatial Technology

BROAD DEFINITION: GIS is a system of hardware, software and procedures designed to support the capture, management, manipulation, analysis, modeling and display of spatially-referenced data for solving complex planning and management problems

- although many other computer programs can use spatial data (e.g. AutoCAD and statistics packages), GISs include the additional ability to perform spatial operations

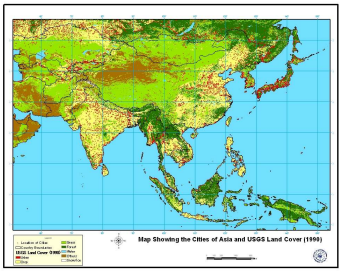
CURRENT STATUS: A field of Science and Systems that brings the Spatial and temporal context of the data into understanding, visualizing, analyzing, modelling and simulating phenomena that exhibit such characteristics, thus providing a more dynamic view and understanding of these phenomena.

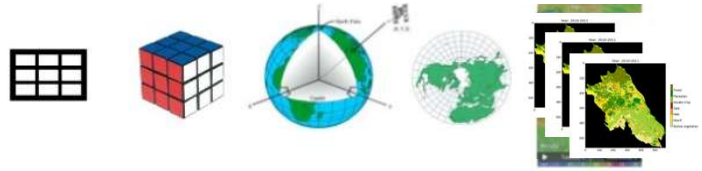
What is an Information System ?

- an Information System is a set of processes, executed on raw data, to produce information which will be useful in decision-making
  - a chain of steps leads from observation and collection of data through analysis
  - an information system must have a full range of functions to achieve its purpose, including observation, measurement, description, explanation, forecasting, decision-making

Geo-Spatial Data

((geo)spatial) data





Tabular	Spatial	Geospatial	Geospatial Projected	Spatiotemporal
	xyz	Xyz lat/long/H	xy(H) EN(H)	xy(H)t

Image Modified based on Talk by Thierry Gregorius Delivered to the BCS Data Management forum, an overview of GIS/Geospatial

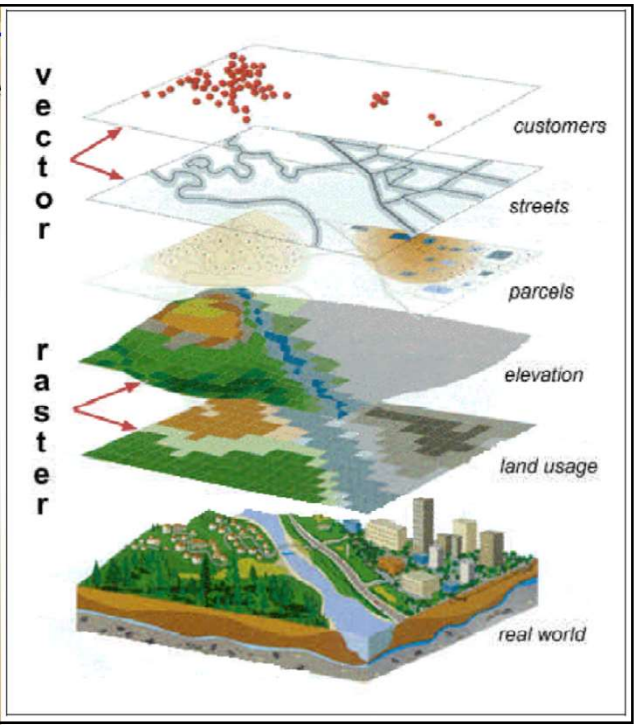
15

Need to capture and represent the Features

- Data Layers

Commonly called as **"Thematic Layers"**

As each layer represents a theme



16

Why should the features be captured in separate layers?

Multiple info

1. Same location – different Attributes

Analysis / Visualization

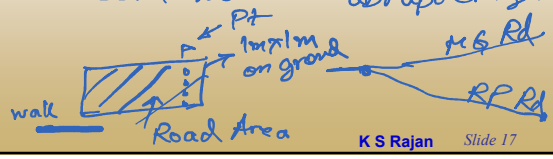
2. Separation – ease of understanding (view, un-clutter)

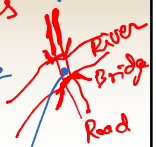
3. Feature Attrib properties (eg: how long/wide is the river)

Different

4. Spatial Objects → Different extents / bnds
   - Geometry ← Discrete / Continuous. → boundary abrupt change

5. Geometry of the feature

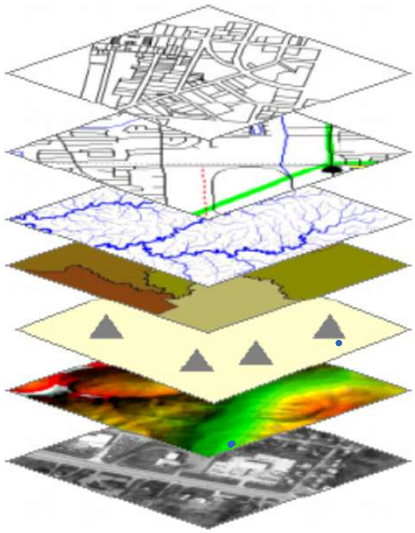




17

Layers in a GIS

One more example is here



Land Ownership

Transportation

Surface Waters

Demographics

Geodetic Control

X, Y, Z

Elevation

Aerial Imagery

18

GeoSpatial Data Model
- Data Structure Design & Purposes

- Efficiently access and retrieve data from database;
- Minimize the required storage space; and
- Support the desired GIS operations

Model for handling Spatial Objects

- **How do we describe Spatial Object features?**
  - **Spatial data**
    - (which describes location (where))
  - **Attribute data** which specifies characteristics at that location
    (what, how much, and when)

- **How to handle multiple Spatial Objects having the same geography but different characteristics or attributes?**
  - **Layers** based on similar characteristics
    (e.g hydrography, elevation, water lines, sewer lines)
  - How does this help?

Can All **Spatial Entities** be handled the same way?

- Assumes identical occurrences can be classified
- Each entity type must be unique (no ambiguity)
  - e.g., detached house classified under house; not industrial building
- Some entities may need to be categorized
- e.g., roadways as a class: with categories for national highways, urban roads, private roads
- Entity type also known as qualitative data
  - or in statistics the 'nominal scale'

So we implicitly do Entity type classification

Some Example of Spatial Objects

- ***areas***:
  - *unbounded:* landuse, market areas, soils, rock type
  - *bounded:* city/county/state boundaries, ownership parcels, zoning
- ***networks***:
  - roads, transmission lines, streams
- ***points***:
  - *fixed:* wells, street lamps, addresses
  - *moving:* cars, fish, deer
- ***continuous***: elevation, rainfall, ocean salinity

Object: Spatial & Non-Spatial Component

Object
- Identity
- Type code

Spatial Component
- Point
- Line
- Area (polygon / grid cells)
- 2.5 D / 3D

Attribute Component
- Attribute values
- Relations
- Quality
- ....

Questions?

*Can **Continuous** and **Discrete** Spatial Objects be handled the same way?*

How to handle changing spatial features across time??

Source: Bernhardsen, Tor. (1999). 2nd Ed. *Geographic Information Systems: An Introduction*, p 43.

GIS DATA STRUCTURES

There are two fundamental approaches to the representation of the spatial component of geographic information:

Vector Model

(Points, Lines, Polygons)

Raster Model

(Surfaces)

The real world can only be depicted in a GIS through the use of models that define **phenomena in a manner that computer systems can interpret**, as well perform meaningful analysis.

Vector Model

- The first model of indicating geographical space, called vector, allows us to give specific spatial locations explicitly.
- The vector data structure is representative of dimensionally as it would appear on a map (DeMers, 1997).
- The vector data model provides for the precise positioning of features in space.
- Based on analytical geometry, a vector model builds a complex representation from primitive objects for the dimensions: **points, lines and areas**.

Vector Data Models/Structures

- One model for representing geographic space
- Spatial locations are explicit
- Relationships between entities/objects are implicit
- Points associated with single set of coordinates (X, Y)
- Lines are a connected sequence of coordinate pairs
- Areas are a sequence of interconnected lines whose 1st & last coordinate points are the same

In Summary- **Vector described as a quantity with a starting coordinate and an associated displacement and direction.**

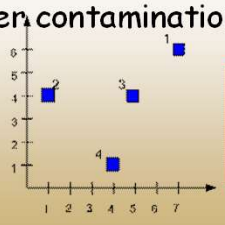
Concept of Vector

- The directional linear connection between two points;
- The root of vector data model lies in cartography;
- The basic elements of vector data model are **points;**
- X,Y coordinates of points, lines and polygons can be represented with unlimited precision;
- Complex spatial objects are created by connecting primitive objects;
- Entity data and attribute data kept in separate files, perhaps a DBMS, which links them

Primitive Objects - Points, Lines, & Polygons

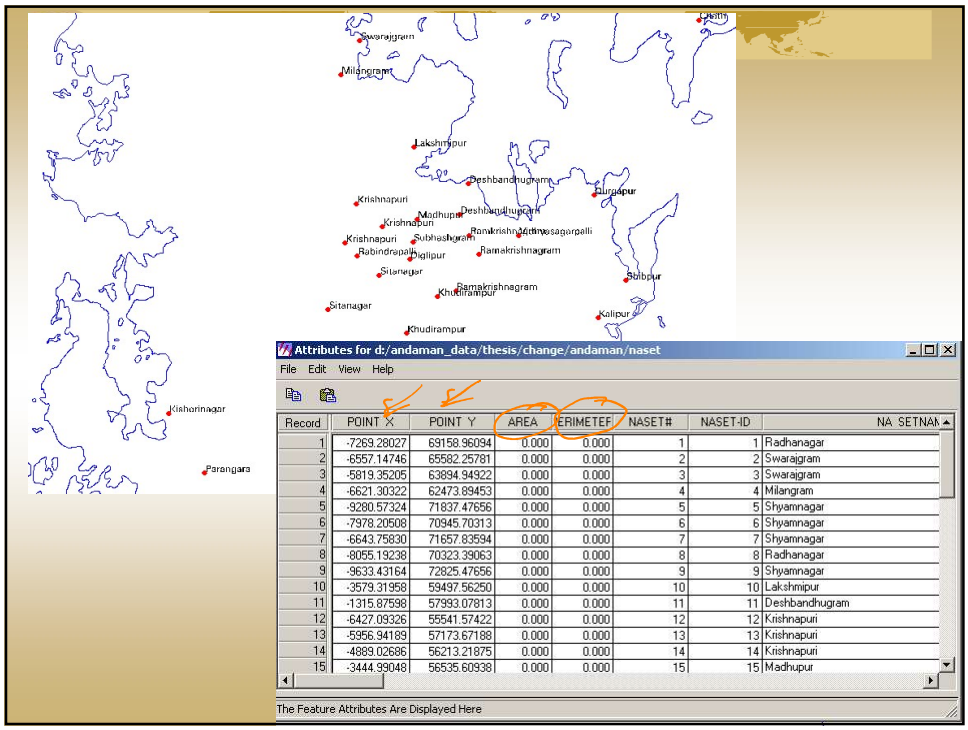
Points

- Represent geographic features too small to be depicted as lines or areas
- A point is recorded as a pair of x, y coordinates.
- **Node** is a topological point at which two or more arcs connect each other
- Each point is stored by its location (X, Y) together with the table attribute of this point.
- For example, 4 points below has their coordinate location in (X, Y) and each point has attributes of deep and amount of water contamination.



Point No	X	Y
1	7	6
2	1	4
3	5	4
4	4	1

Point No	deep	amount
1	12.35	63.45
2	25.79	79.83
3	3.50	24.34
4	17.52	69.23

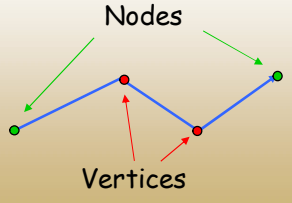


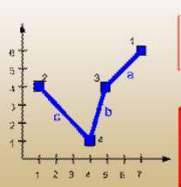
Record	POINT X	POINT Y	AREA	PERIMETER	NASET#	NASET-ID	NA_SETNAM
1	-7269.28027	69158.96094	0.000	0.000	1	1	Radhanagar
2	-6557.14746	65582.25781	0.000	0.000	2	2	Swarajgram
3	-5819.35205	63894.94922	0.000	0.000	3	3	Swarajgram
4	-6621.30322	62473.89453	0.000	0.000	4	4	Milangram
5	-9280.57324	71837.47656	0.000	0.000	5	5	Shyamnagar
6	-7978.20508	70945.70313	0.000	0.000	6	6	Shyamnagar
7	-6643.75830	71657.83594	0.000	0.000	7	7	Shyamnagar
8	-8055.19238	70323.39063	0.000	0.000	8	8	Radhanagar
9	-9633.43164	72825.47656	0.000	0.000	9	9	Shyamnagar
10	-3579.31958	59497.56250	0.000	0.000	10	10	Lakshmipur
11	-1315.87598	57993.07813	0.000	0.000	11	11	Deshbandhugram
12	-6427.09326	55541.57422	0.000	0.000	12	12	Krishnapuri
13	-5956.94189	57173.67188	0.000	0.000	13	13	Krishnapuri
14	-4889.02686	56213.21875	0.000	0.000	14	14	Krishnapuri
15	-3444.99048	56535.60938	0.000	0.000	15	15	Madhupur

30

Line

- Represent geographic features too narrow to depict as areas;
- Lines are recorded as a series of ordered x,y coordinates.
- Each line is stored by the sequence of first and last point together with the associated table attribute of this line.
- For example, three lines below (a, b and c) have their first and last node to distinguish their location and each line has attributes of flow and capacity of the sewerage pipe. Notice that each node has coordinate (X, Y) that is stored in another table.



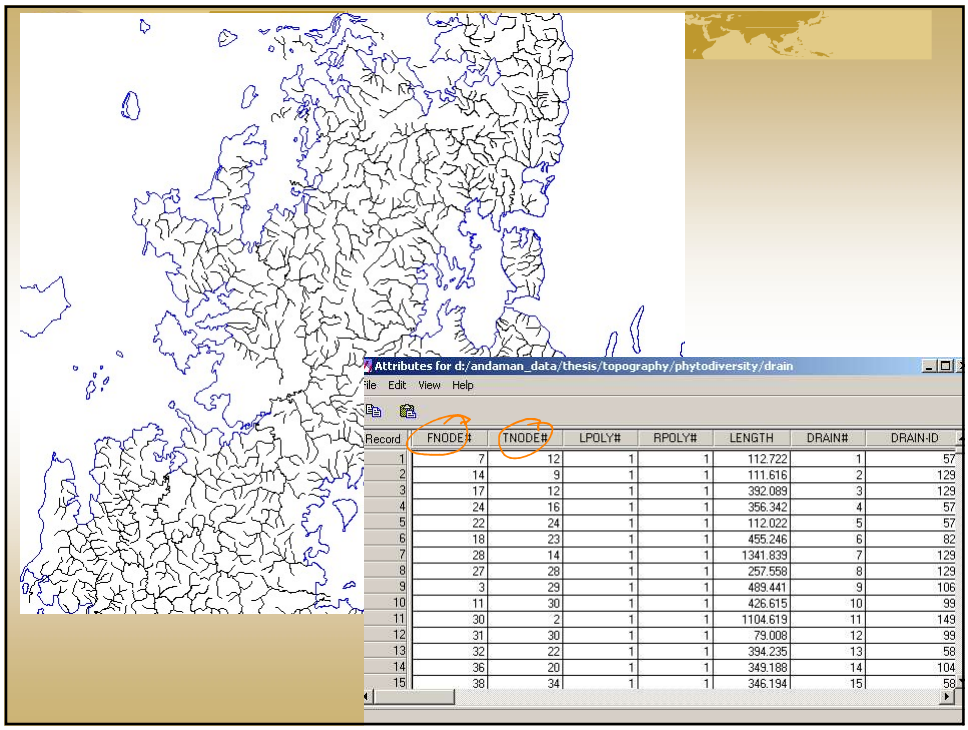


Node No	X	Y
1	7	6
2	1	4
3	5	4
4	4	1

Line	First Node	Last Node
a	1	3
b	3	4
c	4	2

Line	Flow	Capacity
a	960	2200
b	1250	2000
c	1100	2000

31

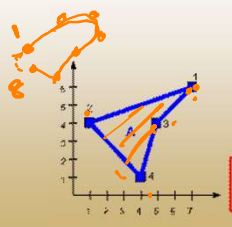


Record	FNODE#	TNODE#	LPOLY#	RPOLY#	LENGTH	DRAIN#	DRAIN-ID
1	7	12	1	1	112.722	1	57
2	14	9	1	1	111.616	2	129
3	17	12	1	1	392.089	3	129
4	24	16	1	1	356.342	4	57
5	22	24	1	1	112.022	5	57
6	18	23	1	1	455.246	6	82
7	28	14	1	1	1341.839	7	129
8	27	28	1	1	257.558	8	129
9	3	29	1	1	489.441	9	106
10	11	30	1	1	426.615	10	99
11	30	2	1	1	1104.619	11	149
12	31	30	1	1	79.008	12	99
13	32	22	1	1	394.235	13	58
14	36	20	1	1	349.188	14	104
15	38	34	1	1	346.194	15	58

Polygon

- Represent homogeneous geographic features. The term polygon, meaning 'many-sided.
- Polygon is represented by a *closed* sequence of lines. Unlike line or poly-line (sequence of line), polygon always closed. That is, *the first point is equal to the last point*.
- A polygon can be represented by a sequence of nodes where the last node is equal to the first node.
- For example, polygon A below has its first and last node in node number 1 to settle its location. Aside from location attributes, the polygon has associated attributes of area and bacterial population.

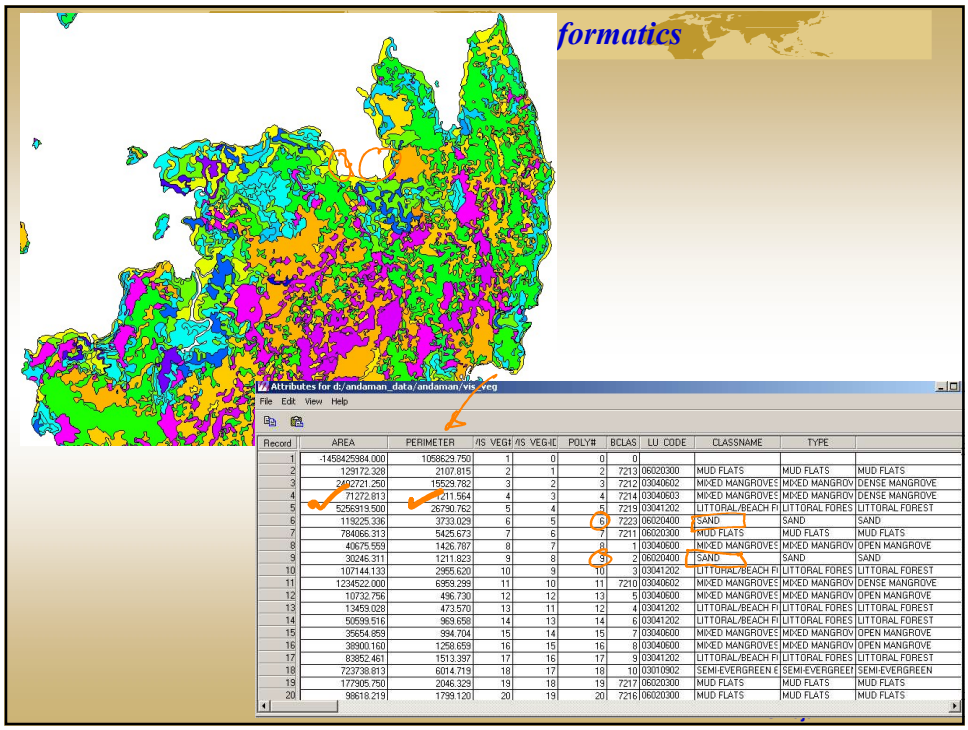
Using polygon, several geometric attributes such as area and perimeters can be derived easily



Node No	X	Y
1	7	6
2	1	4
3	5	4
4	4	1

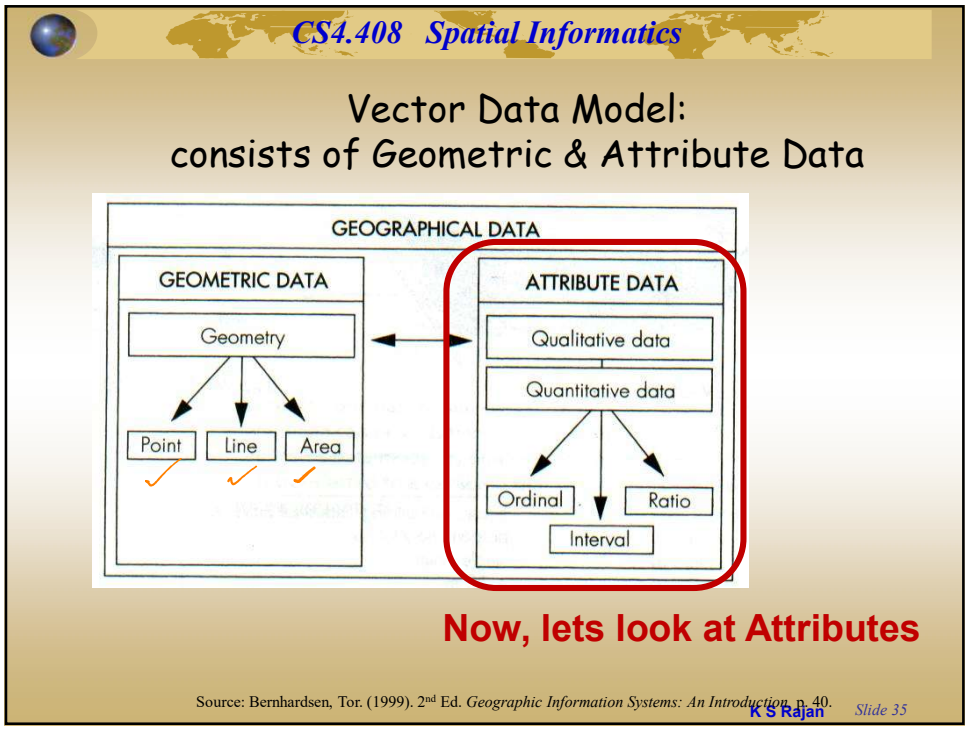
Polygon	Node sequence
A	3,4,2,1

Polygon	Area	Population
A	15.23	12.35



Record	AREA	PERIMETER	/IS_VEGI	/IS_VEGI-ID	POLY#	BCLAS	LU_CODE	CLASSNAME	TYPE	
1	-1458425984.000	1058629.750	1	0	0	0				
2	129172.328	2107.815	2	1	2	7213	06020300	MUD FLATS	MUD FLATS	MUD FLATS
3	2492721.250	15529.782	3	2	3	7212	03040602	MIXED MANGROVES	MIXED MANGROV	DENSE MANGROVE
4	71272.813	1211.564	4	3	4	7214	03040603	MIXED MANGROVES	MIXED MANGROV	DENSE MANGROVE
5	5256913.500	26790.762	5	4	5	7219	03041202	LITTORAL/BEACH FI	LITTORAL FORES	LITTORAL FOREST
6	119225.336	3733.029	6	5	6	7223	06020400	SAND	SAND	SAND
7	784066.313	5425.673	7	6	7	7211	06020300	MUD FLATS	MUD FLATS	MUD FLATS
8	40675.559	1426.787	8	7	8	1	03040600	MIXED MANGROVES	MIXED MANGROV	OPEN MANGROVE
9	30246.311	1211.823	9	8	9	2	06020400	SAND	SAND	SAND
10	107144.133	2955.620	10	9	10	3	03041202	LITTORAL/BEACH FI	LITTORAL FORES	LITTORAL FOREST
11	1234522.000	6959.299	11	10	11	7210	03040602	MIXED MANGROVES	MIXED MANGROV	DENSE MANGROVE
12	10732.756	496.730	12	12	13	5	03040600	MIXED MANGROVES	MIXED MANGROV	OPEN MANGROVE
13	13459.028	473.570	13	11	12	4	03041202	LITTORAL/BEACH FI	LITTORAL FORES	LITTORAL FOREST
14	50599.516	969.658	14	13	14	6	03041202	LITTORAL/BEACH FI	LITTORAL FORES	LITTORAL FOREST
15	35654.859	994.704	15	14	15	7	03040600	MIXED MANGROVES	MIXED MANGROV	OPEN MANGROVE
16	38900.160	1258.659	16	15	16	8	03040600	MIXED MANGROVES	MIXED MANGROV	OPEN MANGROVE
17	83852.461	1513.397	17	16	17	9	03041202	LITTORAL/BEACH FI	LITTORAL FORES	LITTORAL FOREST
18	723738.813	6014.719	18	17	18	10	03010902	SEMI-EVERGREEN E	SEMI-EVERGREE	SEMI-EVERGREEN
19	177905.750	2046.329	19	18	19	7217	06020300	MUD FLATS	MUD FLATS	MUD FLATS
20	98618.219	1799.120	20	19	20	7216	06020300	MUD FLATS	MUD FLATS	MUD FLATS

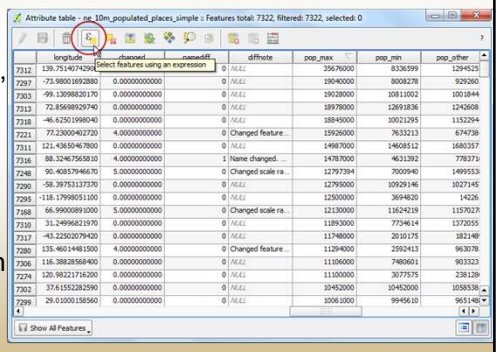
34



35

Attribute – what is it?

- It is the Information about the Spatial Object – it characterizes the Object
- Referred to as Non-Spatial data, Tabular data
  - What it is
  - What it may contain, consist of, etc
  - Where
  - Class / Label / Category
  - Measurable parameters like Area, length, size, etc
  - Any Info related to the Location



Source of table: https://www.qgistutorials.com/en/docs/working_with_attributes.html

Types of Attributes (1/2)

Qualitative Information –Categorical(names)

– Nominal
  - no inherent ordering
  - Eg: land use types, district names

– Ordinal
  - inherent order
  - Eg: road class (NH, SH, etc); stream class

- Cyclical
  - Calendar months, Wind direction, etc

May be represented as Numbers (codes; Class value)
but can't do arithmetic operations on them

Types of Attributes (2/2)

Quantitative Information - Numerical

– Interval
  - Range values, histogram bins
  - Eg: temperature of cities

– Ratio
  - ratios make sense (e.g. twice as much)
  - Eg: income, age, rainfall

They have known differences in values and hence can apply statistics on them (most of the time).

Sources of Attribute Data

- Collated from Reports, Statistical tables
- Field information
- ...

What to do with the Spatial Data?

- Primarily to V**isualize** **(Depict or Draw)** the data

Location of a Spatial Object
- Points like IIITH
- Linear Objects like River Ganges
- District of Hyderabad

- How to **locate or identify** the interested Spatial Object/Objects ?

Query the data to highlight
Extract the data (for use here or elsewhere)

- What about **Analysis** ?

Spatial relationships ?
Cause-effect relations; Co-occurence; mathematical

- Spatial **Modelling** – what and how?