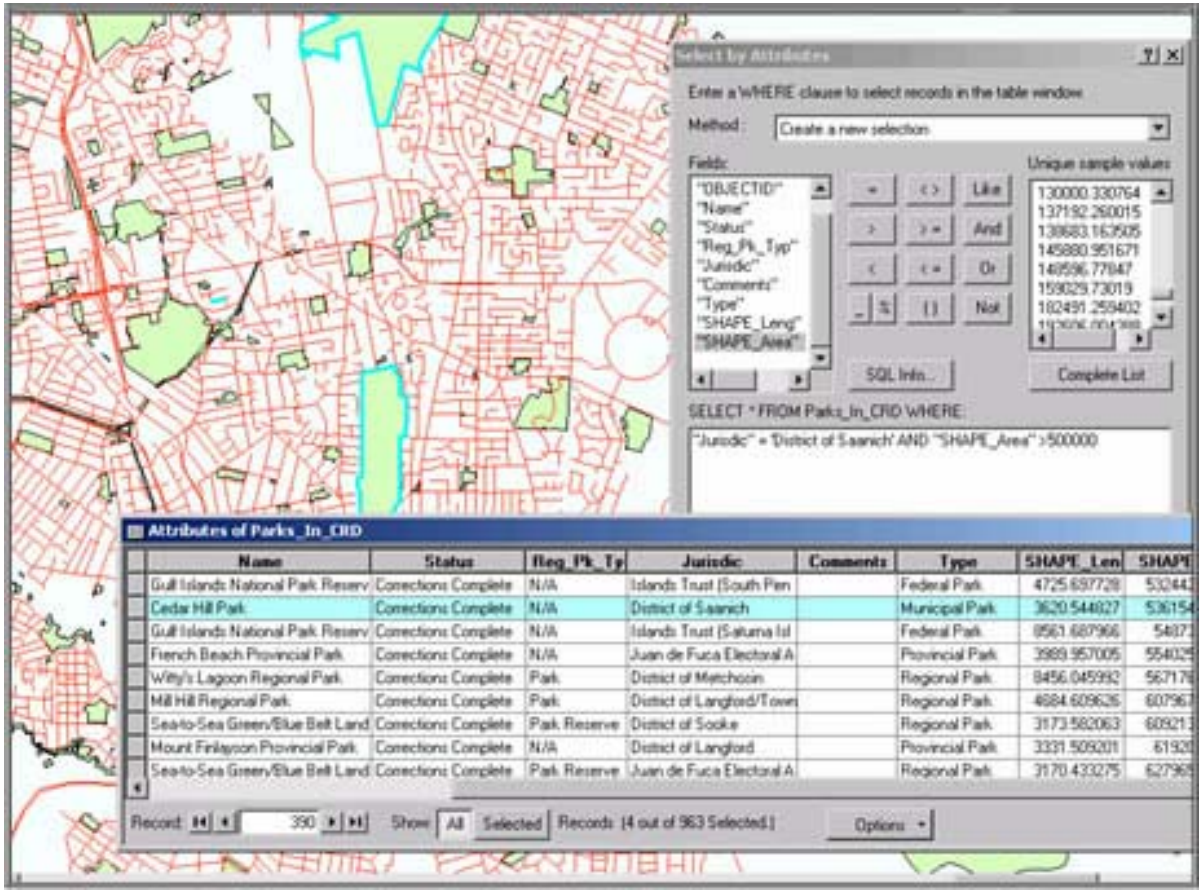
Introduction to GIS Concept

Dr. Tarendra Lakhankar



What is a GIS?

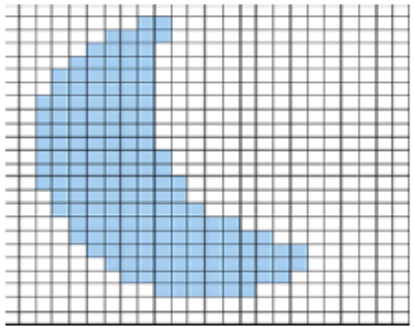
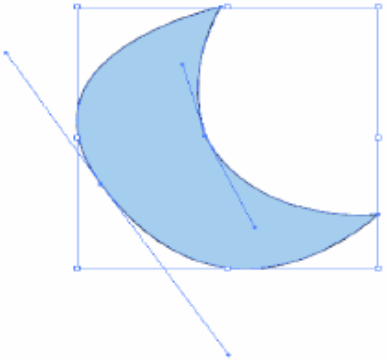
A **GIS** is a computer system that allows you to map, model, query, and analyze large quantities of data within a single database according to their location.

Geographic primitive

$$G = f(x, y, z, t, F)$$

Raster vs Vector

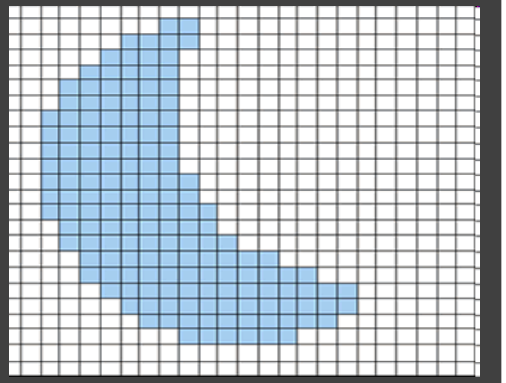
Raster and Vector Graphics

Raster	Vector
	
Made up of a ***grid*** of pixels	Geometric shapes and lines that are defined ***mathematically***
Resolution dependent	Resolution ***independent***
When scaled, visual quality and sharpness is degraded	When scaled, visual quality and sharpness is ***unaffected***
File size is relatively ***large***	File size is relatively ***small***
File Formats: ***GIF, TIF, BMP, PSD***	File Formats: ***EPS, WMF, AI***
Pixel-oriented	***Object*** -oriented

Representing Data with *Raster* and *Vector* Models

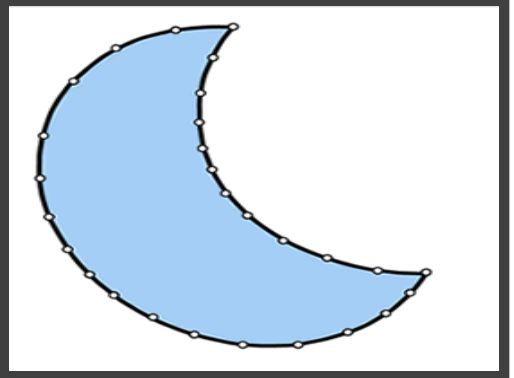
Raster Model

- area is covered by grid with (usually) equal-sized, square cells
- attributes are recorded by assigning each cell a single value based on the majority feature (attribute) in the cell, such as land use type.
- *Image* data is a special case of raster data in which the "attribute" is a reflectance value from the geomagnetic spectrum
- cells in image data often called *pixels* (picture elements)



Vector Model

- The fundamental concept of vector GIS is that all geographic features in the real work can be represented either as:
  - **points or dots (*nodes)***: trees, poles, fire plugs, airports, cities
  - **lines (*arcs)*:** streams, streets, sewers,
  - **areas (*polygons)*:** land parcels, cities, counties, forest, rock type
- Because representation depends on shape, ArcView refers to files containing vector data as *shapefiles*



The Feature Model

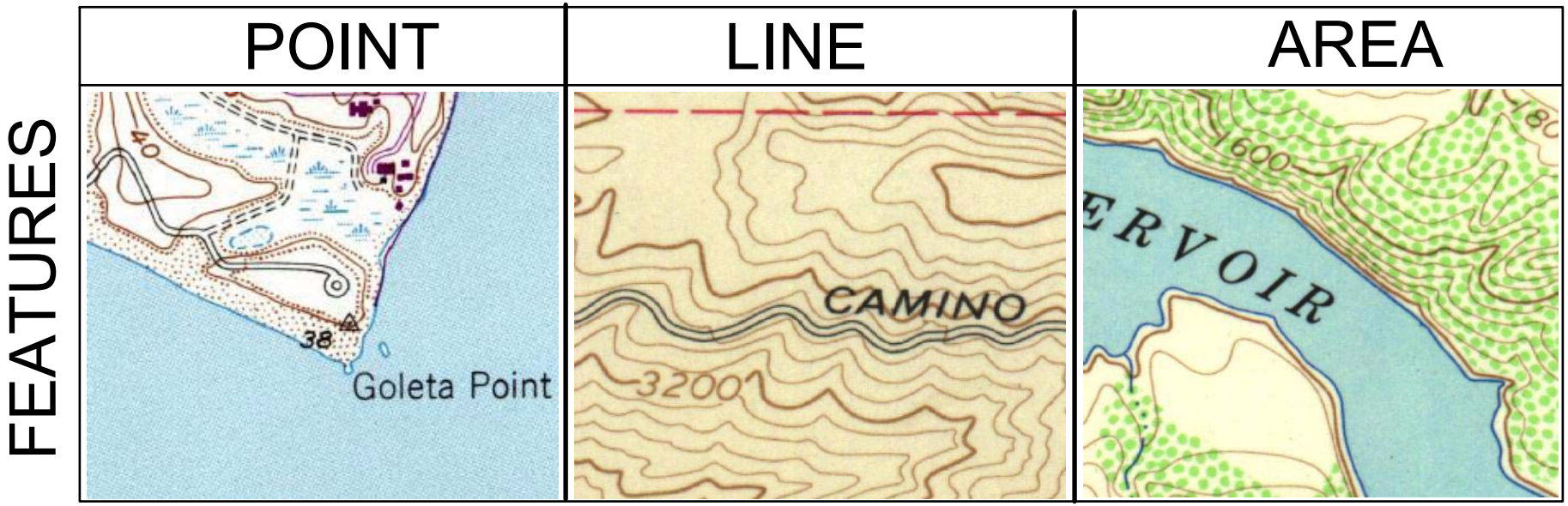
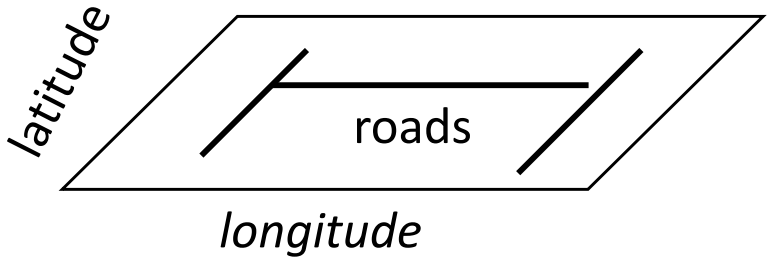
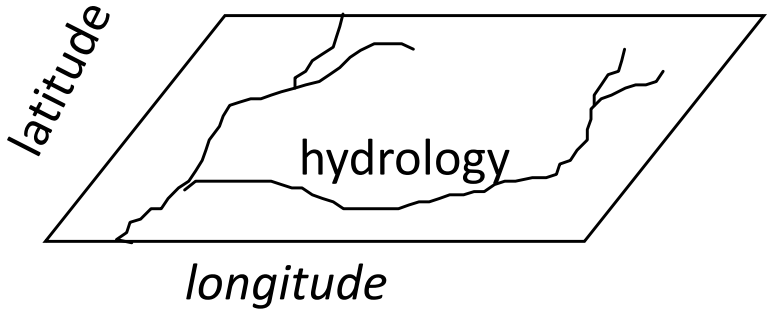


Figure 1.2 The Feature Model: Examples of a point feature (38 foot elevation bench mark), a line fea ture (road, contours) and area features (reservoir, vegetation).

The GIS Model: example



Here we have three layers or themes:
--roads,
--hydrology (water),
--topography (land elevation)
They can be related because precise geographic coordinates are recorded for each theme.

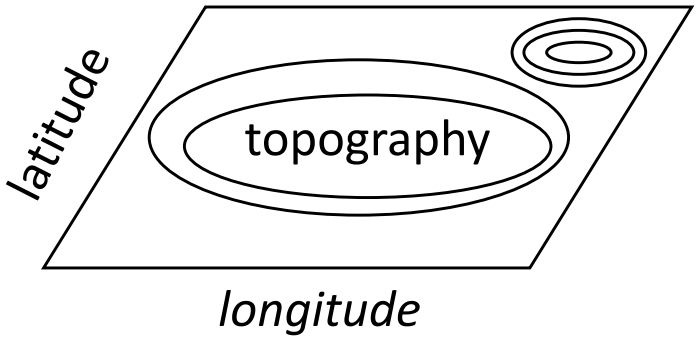


Layers are comprised of two data types:

- ***Spatial data*** which describes location (where)

(Specifies location, stored in a shape file, geodatabase or similar geographic file)

- ***Attribute data*** specifying what, how much, when. (specifies characteristics at that location, natural or human-created, stored in a data base table)



Layers may be represented in two ways:
- in *vector* format as points and lines
- in *raster(or image)* format as pixels

All geographic data has 4 properties:
projection, scale, accuracy and resolution

Projection, Scale, Accuracy and Resolution

- **Projection:** the method by which the curved 3-D surface of the earth is represented by X,Y coordinates on a 2-D flat map/screen
- **Scale:** the ratio of distance on a map to the equivalent distance on the ground
- **Accuracy:** how well does the database info match the real world
  - *Positional*: how close are features to their real-world location?
  - *Consistency*: do feature characteristics in database match those in real world
  - *Completeness*: are all real-world instances of features present in the database?
- **Resolution:** the size of the smallest feature able to be recognized (for raster data, it is the *pixel* size)

Zoom raster image

100%

(900 x 983 dpi)

200%

500%

Raster (image)

- A raster image is defined by pixels. A pixel is the smallest display element that makes up the images seen on a computer monitor or television.
- In raster images, the more pixels an image contains, the higher its resolution.
- For example, in a raster image a square is drawn as a grid of pixels (dots) and each of those pixels will have a specific color value.
- A line is made up of a row of pixels with each pixel having a color value, and you work with this line by working with the group of pixels that makes up the line, not as a single object as you would in a vector file.
- Programs such as Photoshop, PaintShop, and PhotoPaint all work with pixels (raster images).
- Raster images are the best choice for creating subtle gradations of shades and color, such as in a photograph.
- A raster image is resolution-dependant because it contains a fixed number of pixels that are used to create the image. Since there is a fixed and limited number of pixels, a raster image will lose quality if enlarged beyond that number of pixels as the computer will have to 'make up' the missing information. This is usually the cause of the image becoming fuzzy or "steppy".

Vector Image

100%

200%

500%

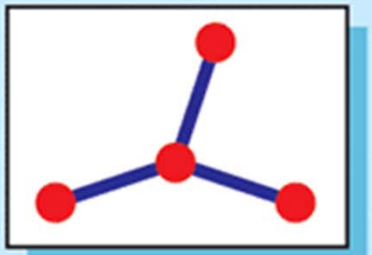
Vector (Lines/Points/Polygons)

- A vector image is defined by objects which are made of lines and curves that are defined mathematically in the computer. Vectors can have various attributes such as line thickness, length and color. For example, in a vector image, a square is drawn as four lines connected at the corners. Those lines can be set to different thickness and colors. The square can be empty or filled.
- A line is one object with attributes, and you work with this line as a single object, not as a group of pixels as you would in a raster image. Programs such as PowerPoint, Illustrator and Freehand all work with vectors.
- Vector graphics are resolution-independent because the vector objects are drawn mathematically in the computer. They can be made larger or smaller without any loss of quality to the image.
- Vectors can be printed at any size, on any output device, at any resolution, without losing detail and without altering the resolution of the image.
- Vector images are the best choice for typefaces, charts and graphs, drawings, and other graphics that must have sharp lines when scaled to various sizes.
- Even though word processing programs such as MSWord are not drawing programs, they too use vectors because they use fonts and fonts are vector files! Fireworks is unique in that it works with both raster and vector art.

Types of Topology

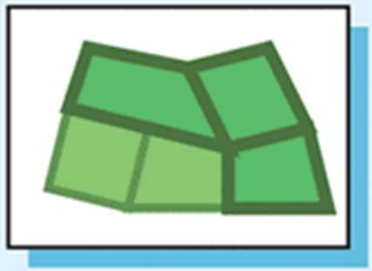
Line features can share endpoints

arc-node topology



Area features can overlap with other area features

region topology

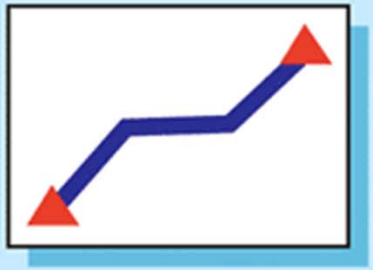


Area features can share boundaries

polygon topology

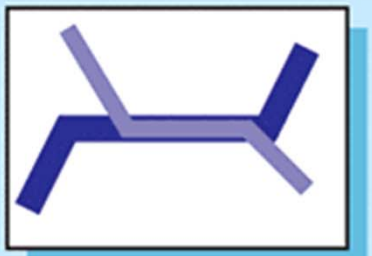
Line features can share endpoint vertices with point features

node topology



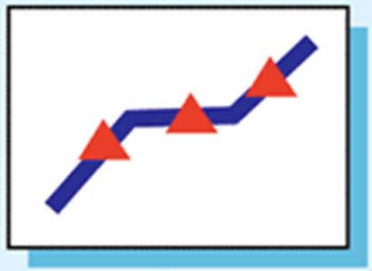
Line features can share segments with other line features

route topology



Point features can share vertices with line features

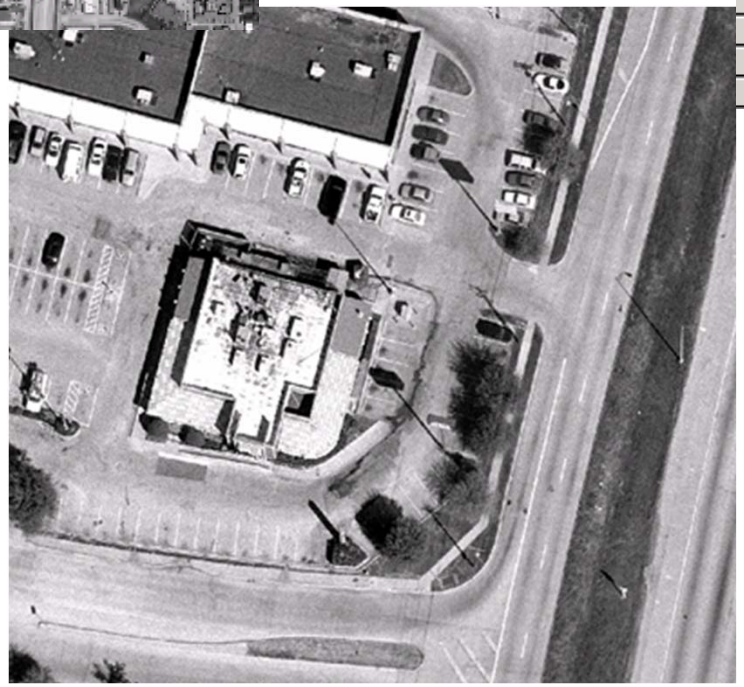
point events



Spatial and non-spatial data

Part Number	Quantity	Description
1034161	5	Wheel spoke
1051671	1	Ball bearing
1047623	6	Wheel rim
1021413	2	Tire
1011210	3	Handlebars
Crimes during 2003		
Date	**Location**	**Type**
22-Jan	123 James St.	Robbery
24-Jan	22 Smith St.	Burglary
10-Feb	9 Elm St. #4A	Assault
13-Feb	12 Fifth Avenue	Breaking and Entering
14-Feb	17 Del Playa	Drunk and Disorderly

Dumb Images & Smart GIS Data



Images—dumb Raster
(although they look good!)

OID	FEA_C	Area_PC	FEA_CODE_1
0	165	2.557509	Paved Driveway
1	165	0.177594	Paved Driveway
2	161	404.664113	Paved Road
3	165	25.081809	Paved Driveway
4	169	11.185954	Grass or Planted Med
5	165	85.809233	Paved Driveway
6	163	27.941142	Public Sidewalk
7	165	104.295646	Paved Driveway
8	165	85.484622	Paved Driveway
9	165	80.315827	Paved Driveway
10	163	17.667767	Public Sidewalk
11	165	147.556552	Paved Driveway
12	165	75.181746	Paved Driveway
13	165	199.456888	Paved Driveway

Smart Vector—Pavement polygons

Smart Raster—5 feet grids

Value	Count	FEA_CODE	Prct_tran	Prct_land
160	62306	Paved Alley	5.056571	1.622552
161	441326	Paved Road	35.816712	11.492865
162	350	Unpaved Road	0.028405	0.009115
163	70285	Public Sidewalk	5.704123	1.830339
164	532582	Paved Parking Lot	43.222779	13.869323
165	96854	Paved Driveway	7.860384	2.522240
166	6119	Paved Trail	0.496600	0.159349
167	6513	Bridge	0.528576	0.169609
168	11518	Paved Median	0.934767	0.299948
169	4326	Grass or Planted Median	0.351085	0.112656

Resolution

Resolution refers to the number of pixels in an image and is a measurement of the output quality of an image, usually in terms of samples, pixels, dots, or lines per inch.

Images are displayed on your computer screen at display (or screen) resolution (72 or 96 ppi). However, a printer needs much more image data than a monitor. In order print a clean image, the typical desktop printer needs 150-300 dpi. If you've ever wondered why your 300 dpi scanned image appears so much larger on your monitor, this is why.

Resolution terminology varies according to the intended output device. PPI (pixels per inch) refers to screen resolution (monitors), DPI (dots per inch) refers to print resolution, SPI (samples per inch) refers to scanning resolution, and LPI (lines per inch) refers to halftone (often newspapers) resolution.

Scanner and digital camera manufacturers often refer to two different types of resolution when listing product specs: optical resolution and interpolated (or digital) resolution.

The optical resolution is the true measurement of resolution that the output device can capture. Interpolated, or digital, resolution is acquired artificially.