UNIT III

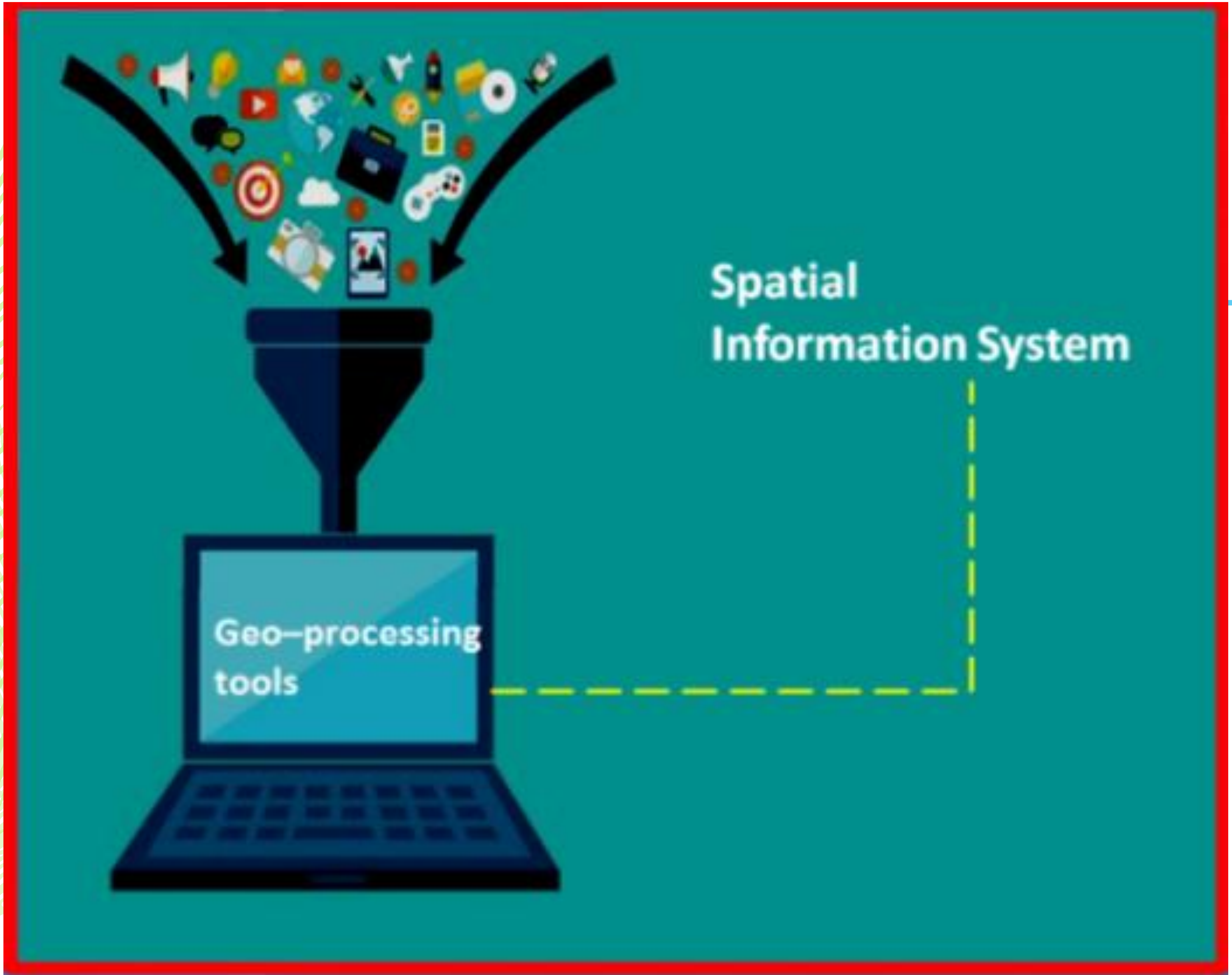
INTRODUCTION TO GIS

Compiled by Dr. Aparna S. Bhaskar

CONTENT

- Introduction to GIS
- Elements of GIS
- Cartography
- Maps and types
- Co-ordinate system
- Projection
- Datum
- GIS data type
- Spatial and non spatial data

- Data input method
- Digitization
- Errors in digitization
- Data output method
- Software modules
- Vector data structure
- Topography
- Raster data structure
- Merits and demerits


Spatial
Information System
Geo–processing
tools

Geographic Information Systems (GIS)

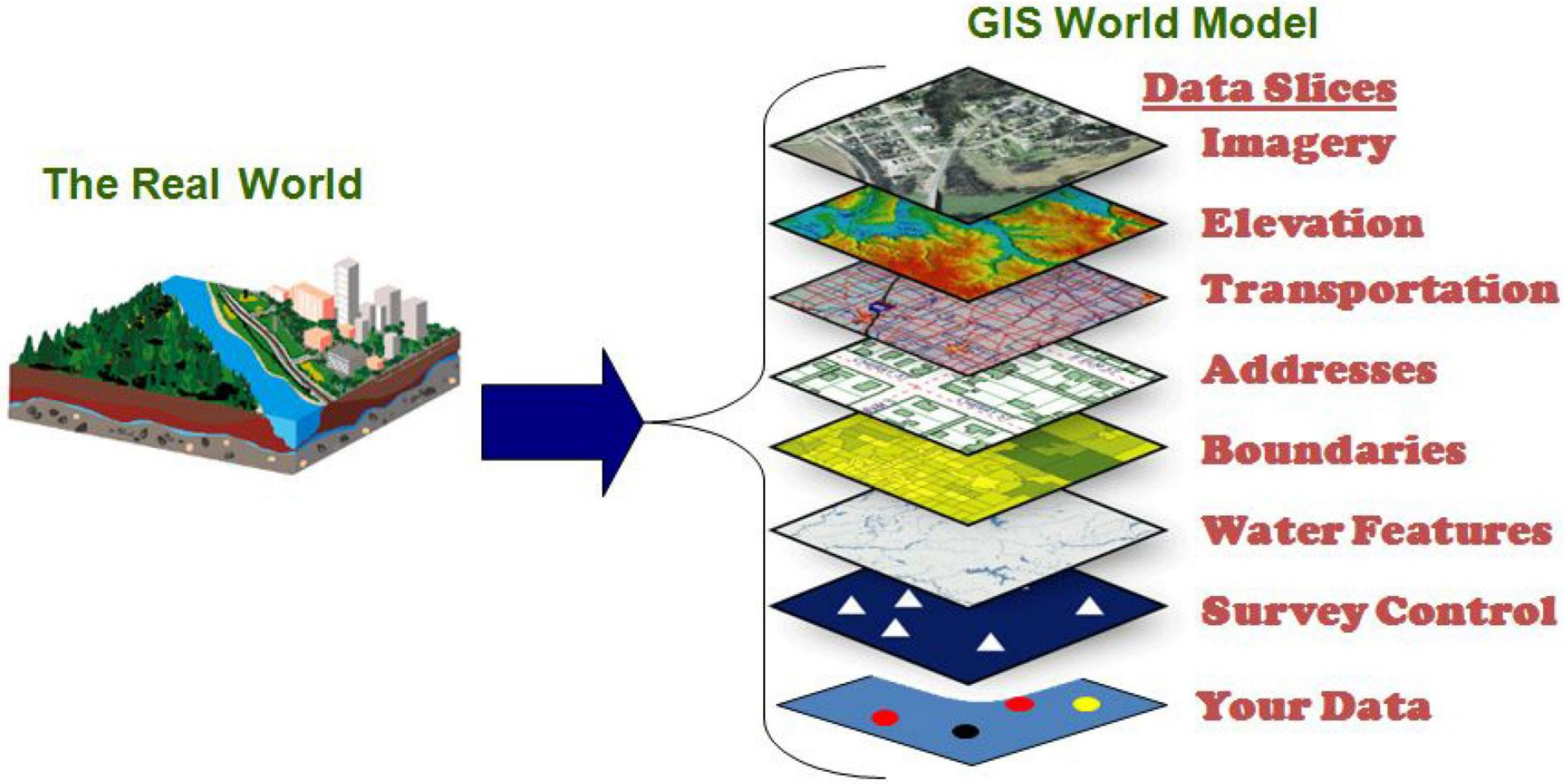
- **GIS** is an acronym that stands for Geographic Information Systems.
- **GIS** is a system or tool for displaying and analyzing data related to positions on Earth's surface.

The phrase, "Geographic Information System", was coined by **Roger Tomlinson** in 1968 - father of GIS“

Tomlinson created a framework for a database that was capable of storing and analysing huge amounts of data; leading to the Canadian government being able to implement its National Land-Use Management Program.

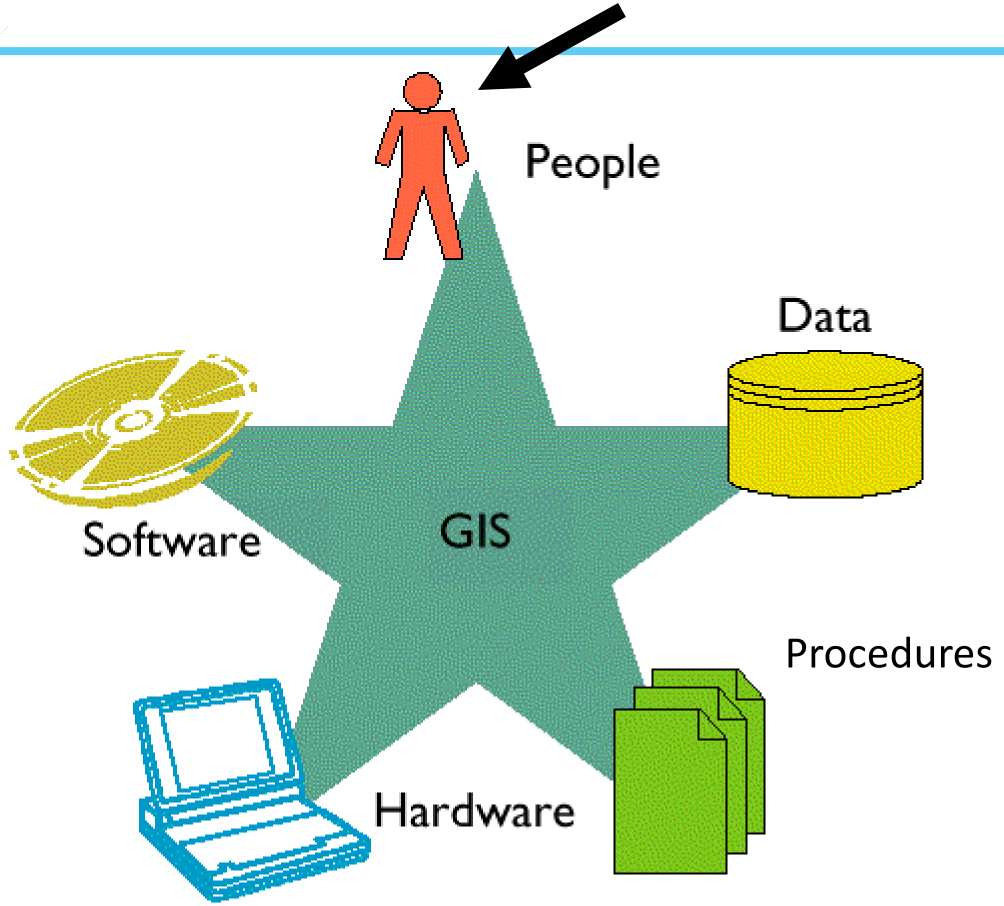
A GIS is a system, consisting of hardware, software data, procedures and a proper organizational context which compiles, stores, manipulates, analyses, models and visualizes spatial data, to solve planning and management problems.

Compilation of definition components given by GOODCHILDS & KEMP 1990; BURROUGH 1986; BILL&FRITSCH 1999


The Real World
GIS World Model
Data Slices
Imagery
Elevation
Transportation
Addresses
Boundaries
Water Features
Survey Control
Your Data

What is a GIS?

– An integration of five basic components



Components of Information Systems

Five Component Framework

Hardware	Software	Data	Procedure	People

1. **Hardware – Desktops, Laptops.**
2. **Software – Operating System , Application Programs.**
3. **Data – Facts & figures entered into Computers.**
4. **Procedures – How the other four components are used.**
5. **People – User, Technologists, is support.**

Hardware is the computer on which a GIS operates. Today, GIS runs on a wide range of hardware types, from centralized computer servers to desktop computers used in standalone or networked configurations.

COMPUTER HARDWARE:

Various functions of hardware peripherals are listed below.

1. Disk drive: The CPU is linked to disk drive, storage unit which gives space for storing data and programs.
2. Digitizer: It is used to convert data from maps and documents into digital form and send them to the computer.
3. Plotter: It is used to display the result of the data processing.
4. Tape drive: It is used to communicate with other system.
5. The visual: It is used to control the computer and peripherals. Otherwise known as 'terminals'.

The GIS software includes the programs and the user interface for driving the hardware. GIS software is essential to generate, store, analyze, manipulate and display geographic information or data.

GIS SOFTWARE:

1. Data input and verification 2. Data storage and management 3. Data analysis and modeling 4. Data output and Presentation

DATA INPUT AND VERIFICATION: all aspects of transforming data captured in the form of existing maps, field observations and sensors in to a compatible digital from.

DATA STORAGE AND DATA BASE MANAGEMENT: concerns with the way in which the data about the position and attributes of geographic entities are structured and organized both with respect to the way they must be handled in the computer and they are received by the users.

DATA ANALYSIS AND MODELLING: This embraces two operations

i) Transformation needed to remove errors from the data or to bring them up to date or to data files for analysis.

ii) Analysis through suitable recoding on the sample data, to suit the problem in question and resorting the various modeling techniques.

DATA OUTPUT AND PRESENTATION: creation of special thematic map displays using graphic technique. The displays represent an extremely powerful tool for both summarization and operations within the spatial and non-spatial database. The report may be generated in any format subjected to the device used for display

Data

- Maybe the most important component of a GIS is the data. Geographic data and related tabular data can be collected in-house or bought from a commercial data provider.
- Most GIS employ a DBMS to create and maintain a database to help organize and manage data.
- The data that a GIS operates on consists of any data bearing a definable relationship to space, including any data about things and events that occur in nature.
- At one time this consisted of hard-copy data, like traditional cartographic maps, surveyor's logs, demographic statistics, geographic reports, and descriptions from the field. Advances in spatial data collection, classification, and accuracy have allowed more and more standard digital base-maps to become available at different scales.

Methods

- A successful GIS operates according to a well-designed plan and business rules, which are the models and operating practices unique to each organization.

CARTOGRAPHY

The art and science of graphically representing a geographical area, usually on a flat surface such as a map or chart. It may involve the superimposition of political, cultural, or other non geographical divisions onto the representation of a geographical area.

- **Cartography** is most relevant. Without maps, we would be “spatially blind.” Knowledge about spatial relations and location of objects are most **important** to learn about space, to act in space, to be aware of what is where and what is around us, or simply to be able to make good decisions.
- The achievement of **cartographic** understanding has, as a prerequisite, the understanding of the **basic characteristics** of maps. Scale, map projections, generalization, and symbolization are common to every map and are considered as **basic characteristics** of maps.

Map

A **map** is a representation of a place. ... It is a symbolic interpretation of place and highlights the relationships between elements in space, either perceived or actual.

Imago Mundi, commonly known as the Babylonian Map of the World, is considered the oldest surviving world map. It is currently on display at the British Museum in London. It dates back to between 700 and 500 BC and was found in a town called Sippar in Iraq.

Types of Map

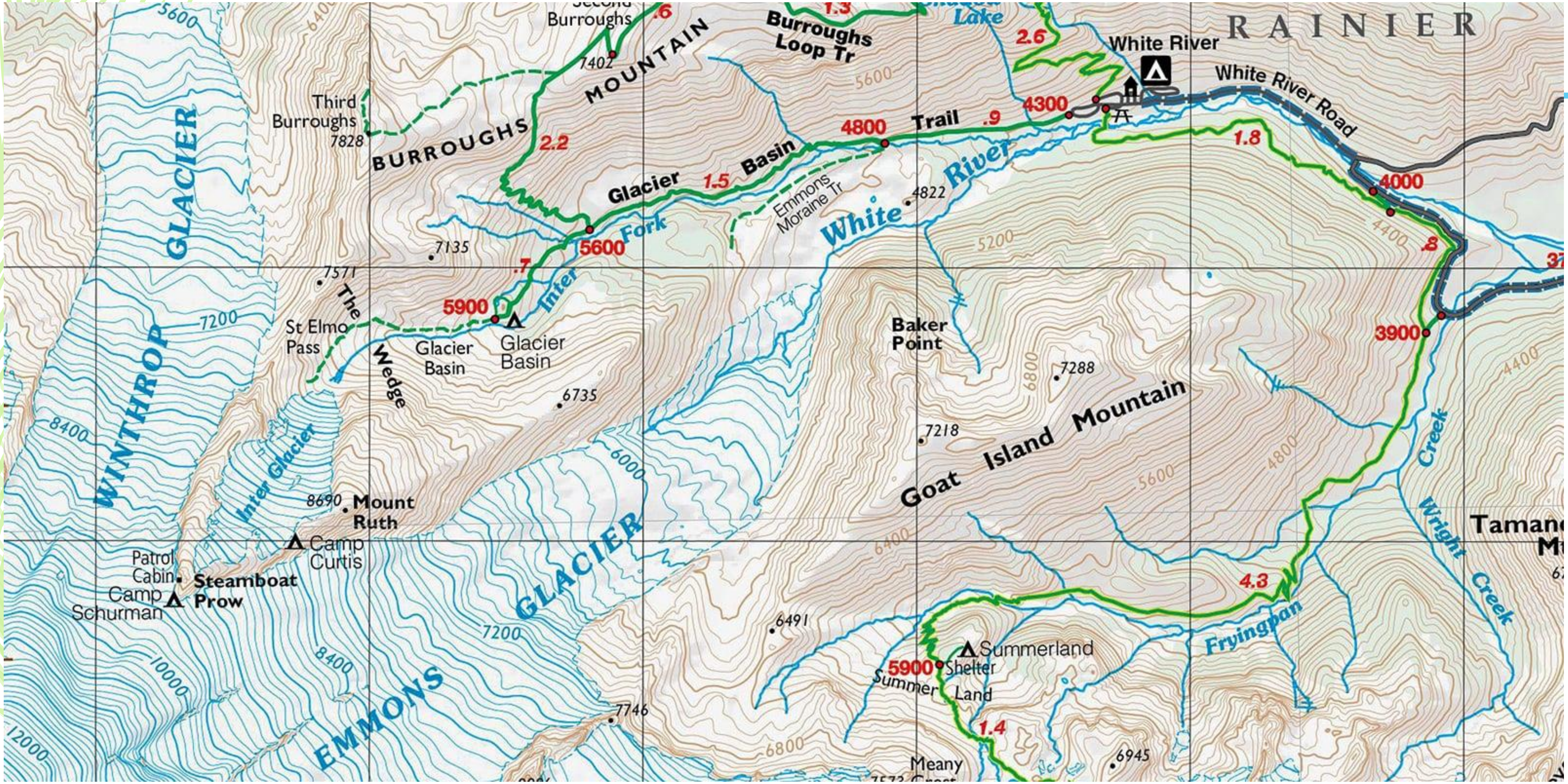
– According to the ICSM (Intergovernmental Committee on Surveying and Mapping), there are five **different types of maps**:

1. General Reference,
2. Topographical,
3. Thematic maps
4. Navigation Charts and
5. Cadastral **Maps** and Plans.

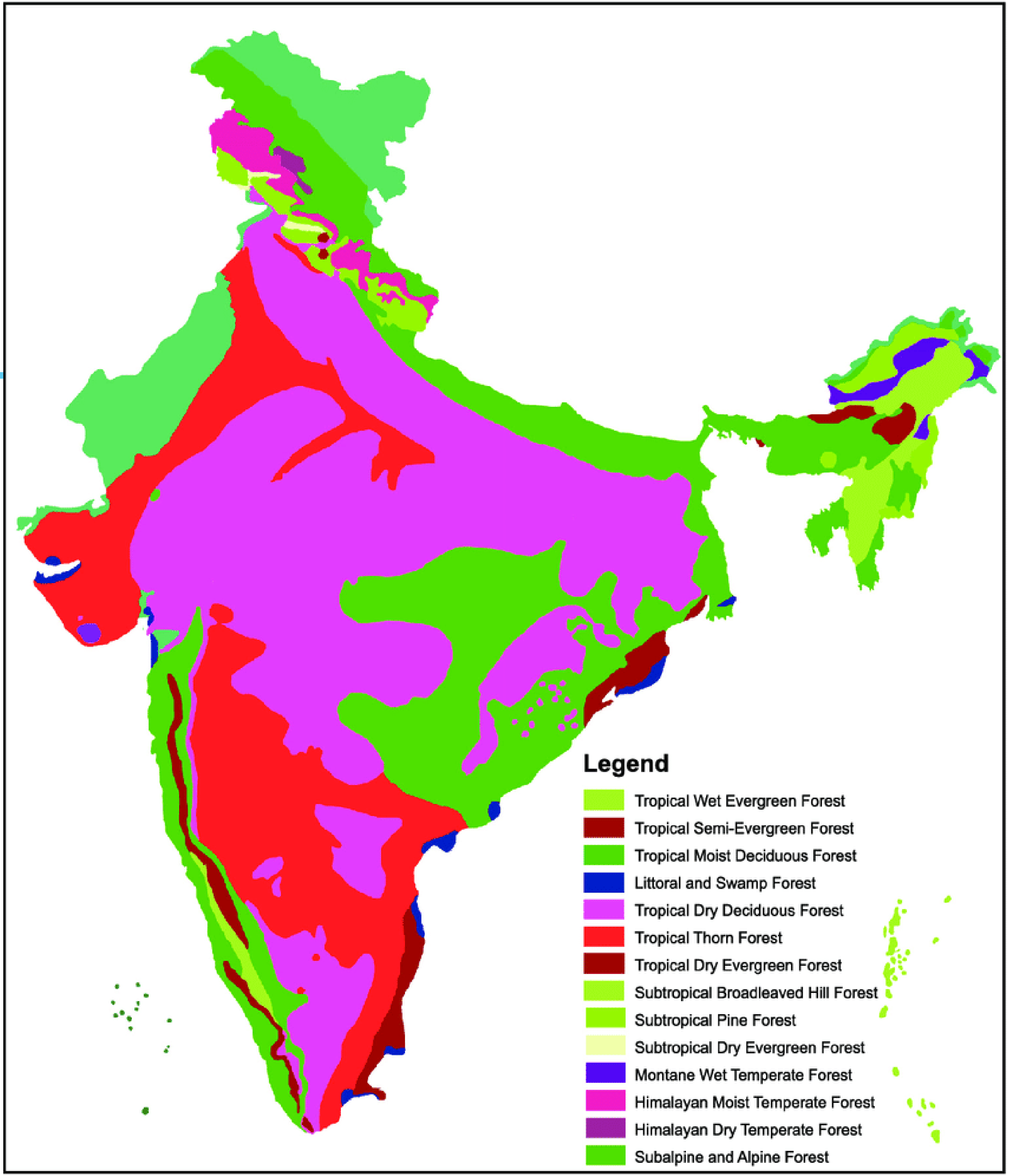
General Reference

– Think of a regular map, where cities and towns are named, major transport routes are included along with natural features like lakes and rivers, and you'll be thinking of a general reference map. These are the maps that are ideal for helping you to get to your destination – they tend to be easy to read, and include street and tourist maps, and we can't think of many better examples than HERE maps.

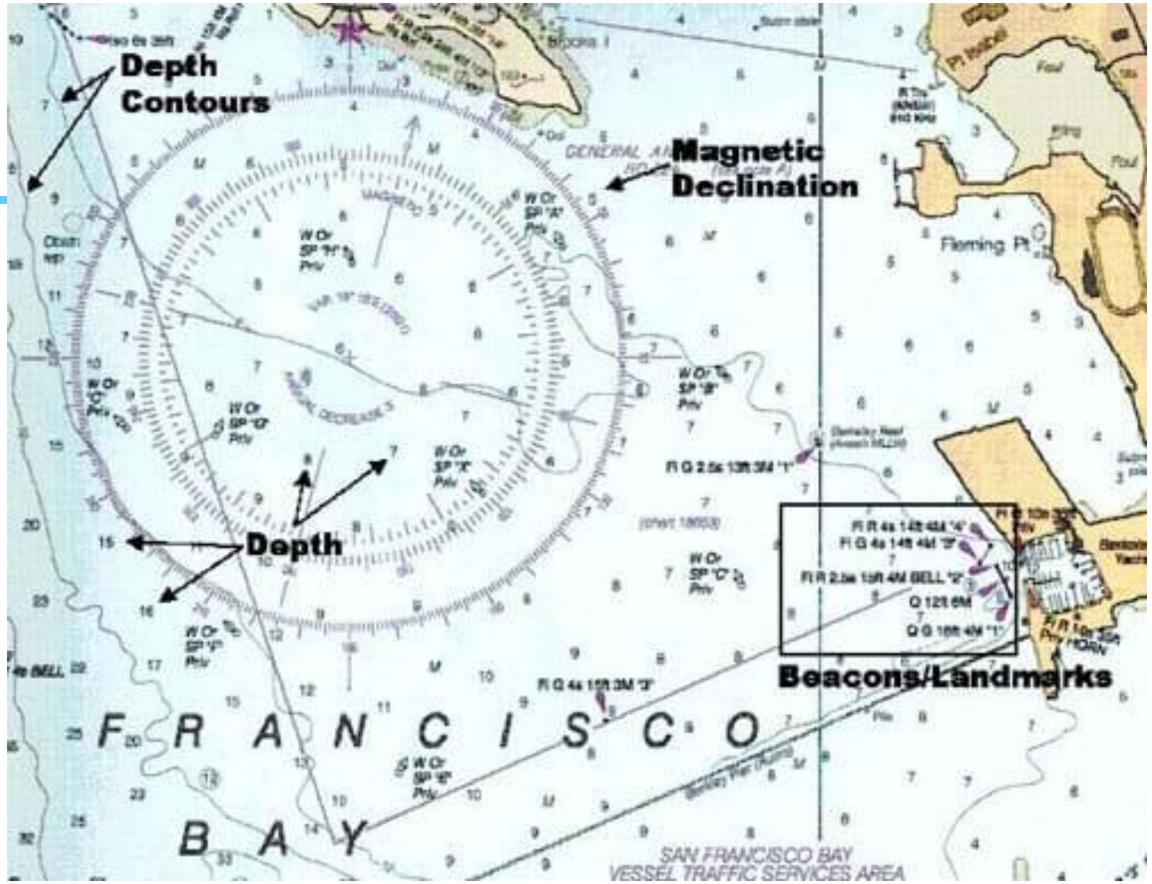
– **Topographic maps** are a detailed record of a land area, giving geographic positions and elevations for both natural and man-made features. They show the shape of the land the mountains, valleys, and plains by **means** of brown contour lines (lines of equal elevation above sea level).



- A thematic map is a type of map specifically designed to show a particular theme connected with a specific geographic area, such as temperature variation, rainfall distribution or population density.

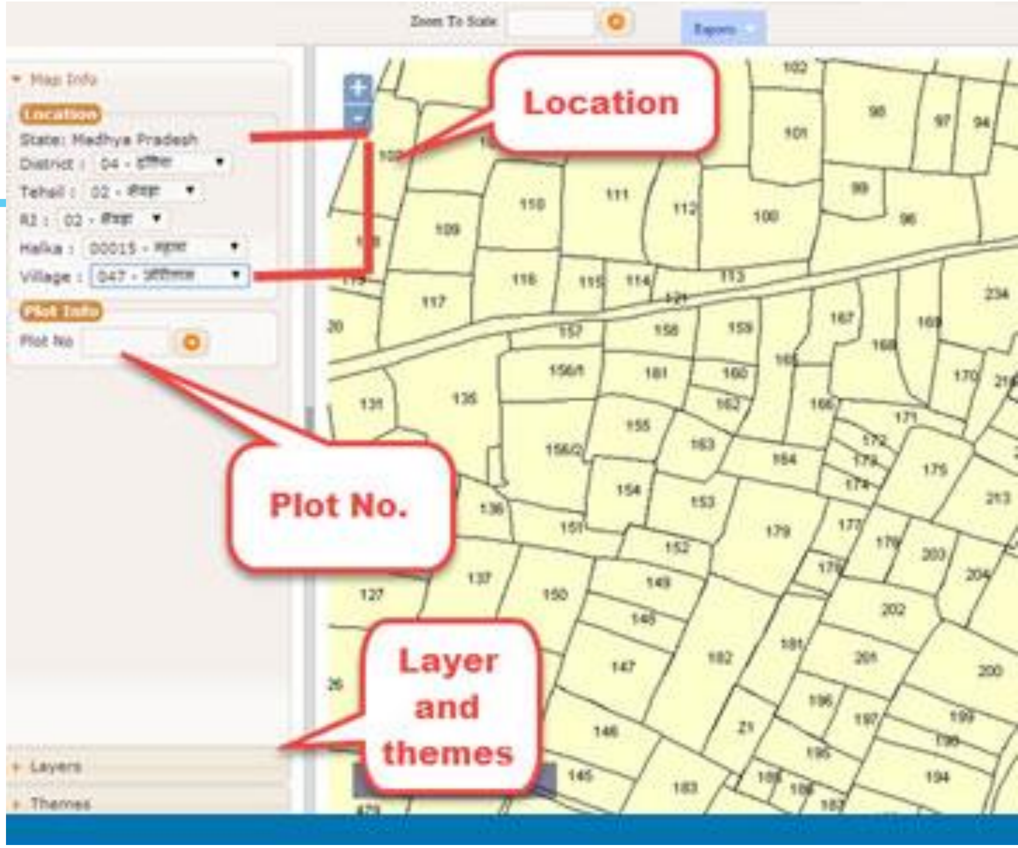


– A **nautical chart** is a graphic representation of a sea area and adjacent coastal regions. Depending on the scale of the chart, it may show depths of water and heights of land (topographic map), natural features of the seabed, details of the coastline, navigational hazards, locations of natural and human-made aids to navigation, information on tides and currents, local details of the Earth's magnetic field, and human-made structures such as harbours, buildings and bridges. Nautical charts are essential tools for marine navigation.



- A cadastre is a comprehensive land recording of the real estate or real property's metes-and-bounds of a country. In most countries, legal systems have developed around the original administrative systems and use the cadastre to define the dimensions and location of land parcels described in legal documentation.

- **A cadastral map** is a map which provides detailed information about **real property** within a specific area. A simple example of a cadastral map might be a map of a village which shows the boundaries of all of the parcels or lots within the village, although cadastral maps can show other types of areas as well. These maps are usually maintained by the government, and they are a matter of public record; anyone who wishes to go to the office which maintains the records can ask to see them.



Map components/Elements

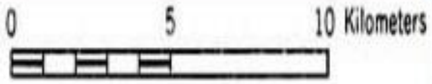
– Maps contain lots of information. Most maps will have the five following elements
– **Title**
– **Legend**
– **Grid**
– Compass to indicate **direction**, and
– **Scale**.

- **Scale is the relationship that the depicted feature on map has to its actual size in the real word (more: *map scale*).**
- All maps are modelled representations of the real world and therefore the features are reduced in size when mapped. In other words, scale is the measurement of the amount of reduction a mapped feature has to its actual counterpart on the ground.
- There are **three** main ways that **scale** is indicated on a **map**:

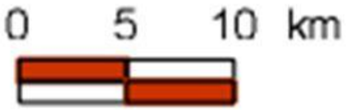
1. **Graphic (or bar),**
2. **Verbal, and**
3. **Representative fraction (RF).**

Three Types of Scale

- **There are three different ways to write scale.**

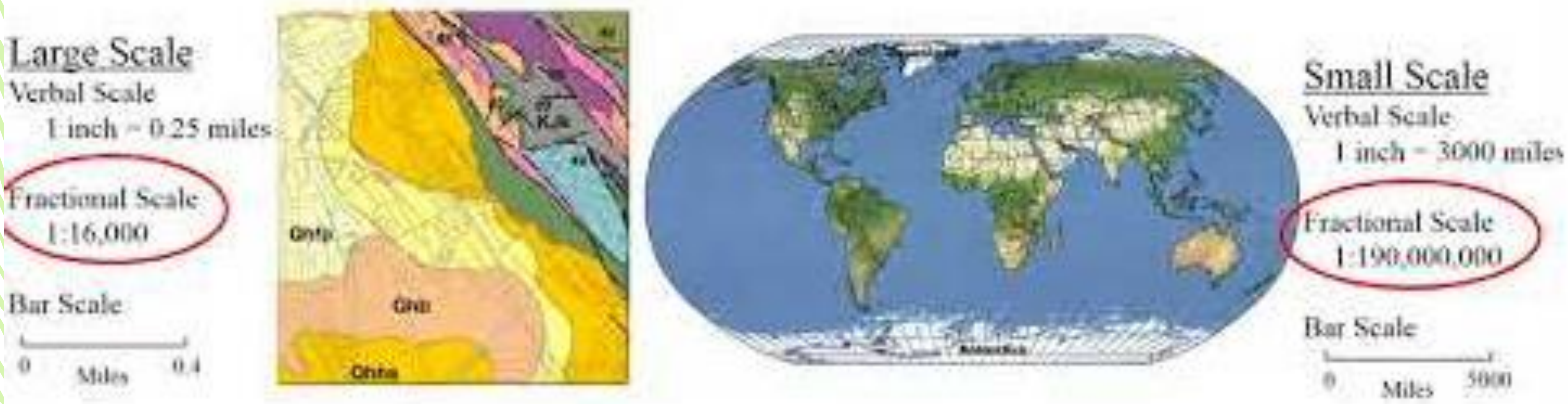
Word Scale	• 1 cm = 250 km
Linear Scale or Bar Scale	
Ratio Scale or Representative Fraction Scale	• 1:25 000 000

<u>3 types of map scales</u>

Graphic Scale:	
Verbal Scale:	1 cm = 1 km
Representative Fraction:	1:100,000

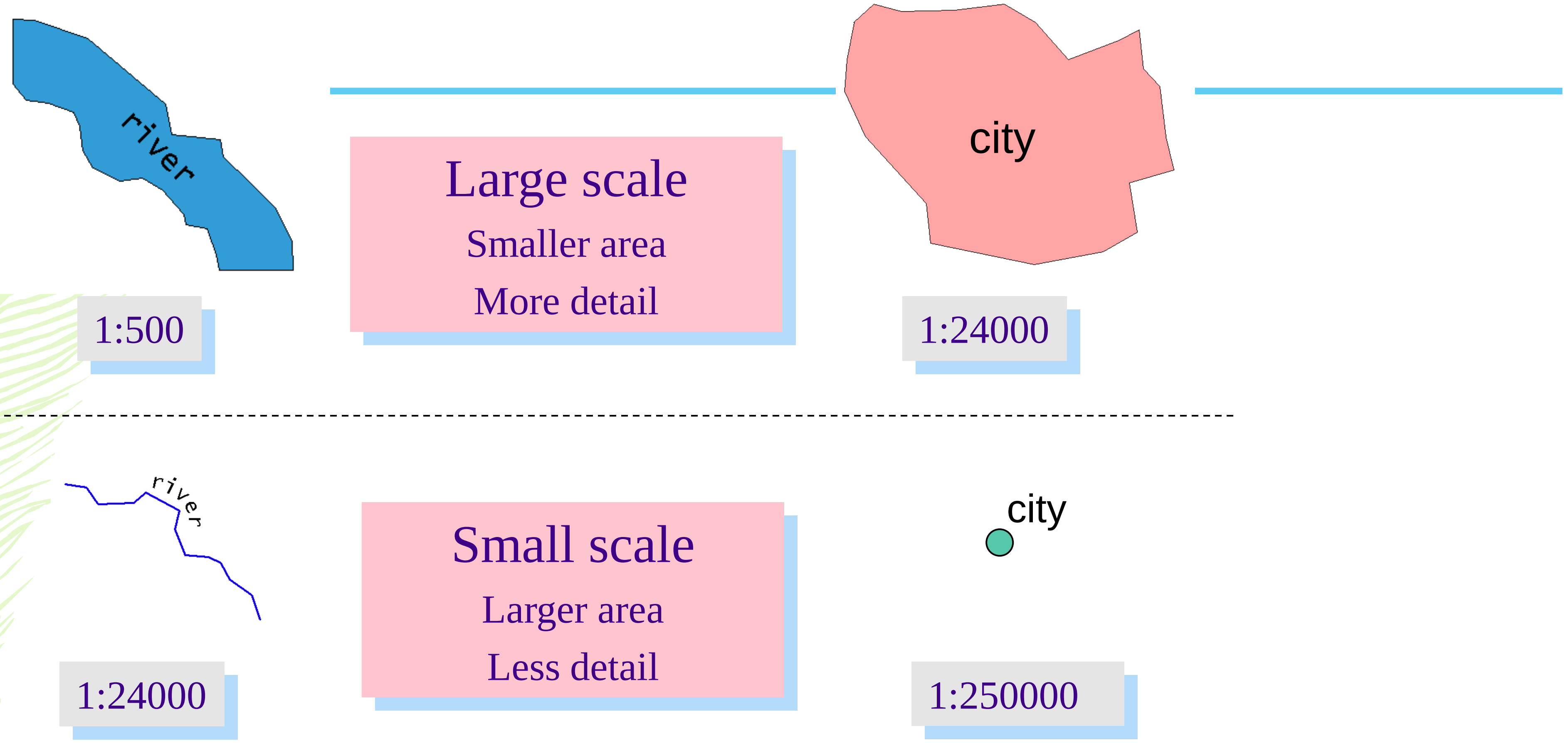
Large Scale vs. Small Scale

- Large scale – large amount of detail; can only show small area
- Small scale – small amount of detail; can show a large area



Map scale

– Map scale determines the size and shape of features



Based on scale: Large , Medium and small

Ways of expressing scale of a map: Representative Fraction (RF), Statement / Verbal, Graphical/Bar

Typical RF	1:1000 1:5,000 1:10,000	1:20,000 1:50,000 1:100,000	1:1,000,000 1:2,500,000
Description	LARGE-SCALE	MEDIUM-SCALE	SMALL-SCALE
Characteristics	• Depict small features • Show geometric shapes	• Small features disappear • Generalize geometric shapes • Good compromise between map detail and extent of map coverage	• Symbolize features, e.g., areas represented by point or line symbols • Show macro features, e.g., climatic zones

COORDINATE SYSTEM

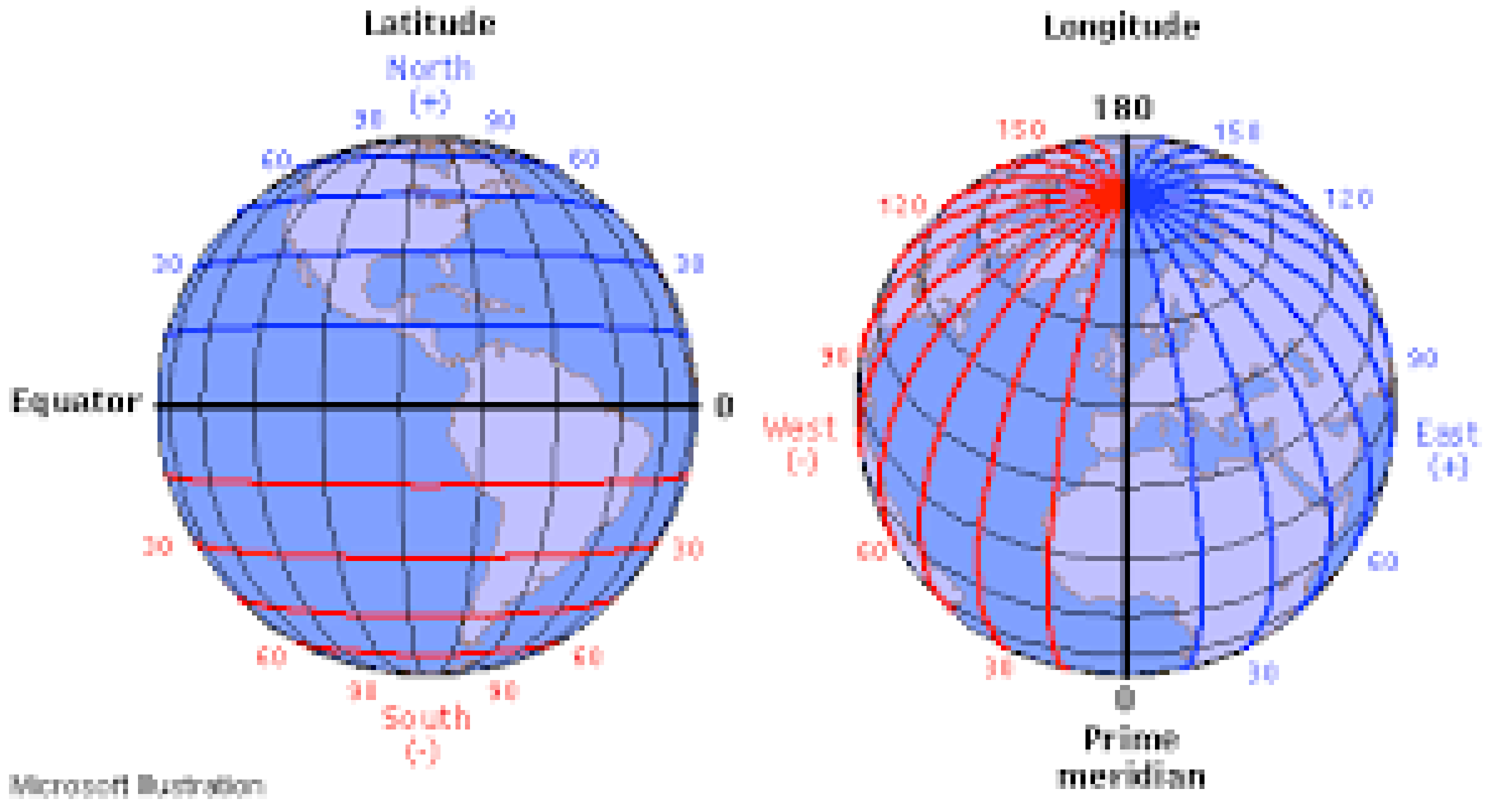
A coordinate system is a reference system used to represent the locations of geographic features, imagery, and observations such as GPS locations within a common geographic framework.

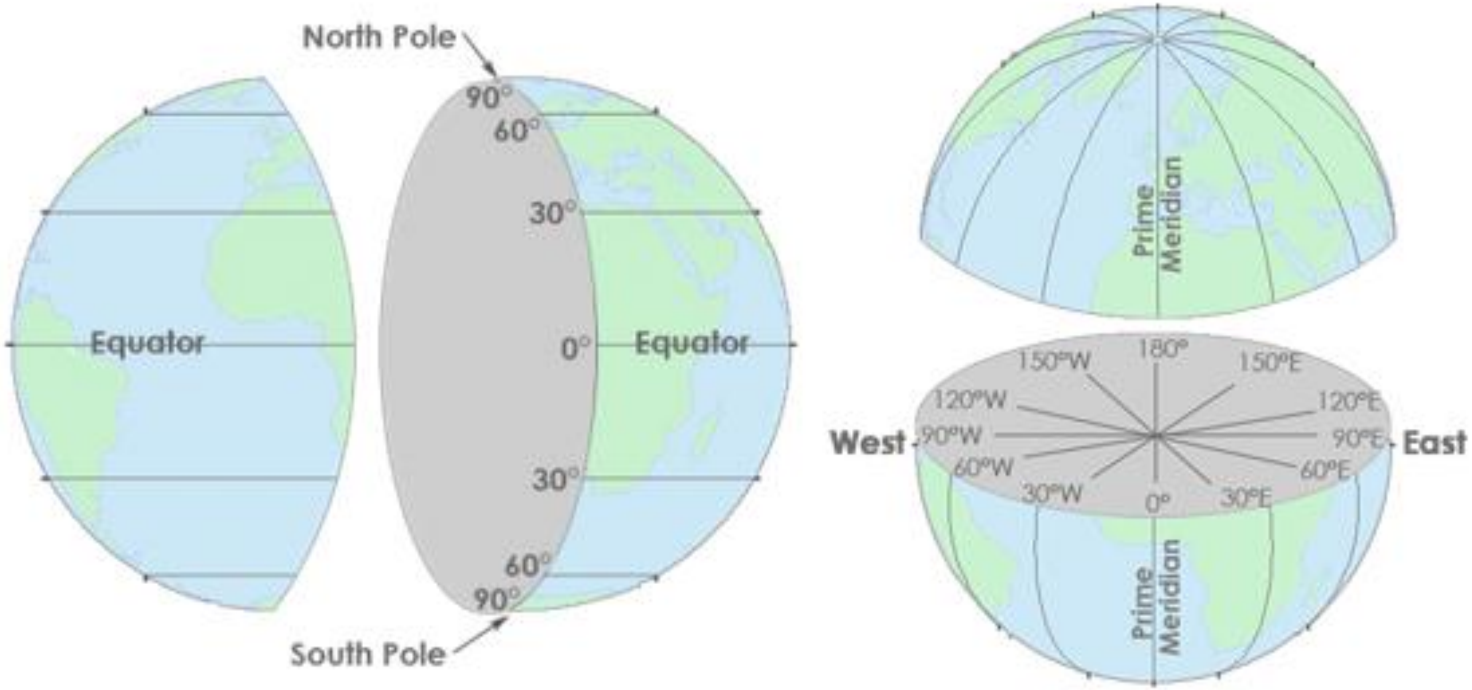
Coordinate systems (either geographic or projected) provide a framework for defining real-world locations.

Coordinate system

- Each coordinate system is defined by the following:
- Its measurement framework, which is either geographic (in which spherical coordinates are measured from the earth's center) or planimetric (in which the earth's coordinates are projected onto a two-dimensional planar surface)
- Units of measurement (typically feet or meters for projected coordinate systems or decimal degrees for latitude-longitude)
- The definition of the map projection for projected coordinate systems
- Other measurement system properties such as a spheroid of reference, a datum, one or more standard parallels, a central meridian, and possible shifts in the x- and y-directions

- **Latitude, Longitude**
- **Latitude** lines run east-west and are parallel to each other. If you go north, latitude values increase. Finally, latitude values (Y-values) range between -90 and +90 degrees
- But **longitude** lines run north-south. They converge at the poles. And its X-coordinates are between -180 and +180 degrees.
- Latitude and longitude coordinates make up our geographic coordinate system.





20.5937° N, 78.9629° E

Types of coordinate systems

Horizontal **coordinate systems** locate data across the surface of the earth, and vertical **coordinate systems** locate the relative height or depth of data. Horizontal **coordinate systems** can be of three **types**: geographic, projected, or local.

- Geographic coordinate systems
- A geographic coordinate system (GCS) uses a three-dimensional spherical surface to define locations on the earth.

Projected coordinate systems

- A projected coordinate system (PCS) is defined on a flat, two-dimensional surface. Unlike a GCS, a PCS has constant lengths, angles, and areas across the two dimensions. A PCS is always based on a GCS that is based on a sphere or spheroid. In addition to the GCS, a PCS includes a map projection, a set of projection parameters that customize the map projection for a particular location, and a linear unit of measure.

- **Vertical coordinate systems** are either gravity-based or ellipsoidal. Gravity-based vertical coordinate systems reference a mean sea level calculation. Ellipsoidal coordinate systems reference a mathematically derived spheroidal or ellipsoidal volumetric surface.

A **geographic coordinate** system (GCS) uses a three-dimensional spherical surface to define locations on the earth. Geographic coordinates based on a spheroid is **geodetic coordinates**

A projected coordinate system based on a map projection to project features of the earth's spherical surface onto a two dimensional Cartesian coordinate plane.

Coordinate system on a plane two types

- Plane Rectangular/Cartesian Coordinate system
- Plane polar coordinate system

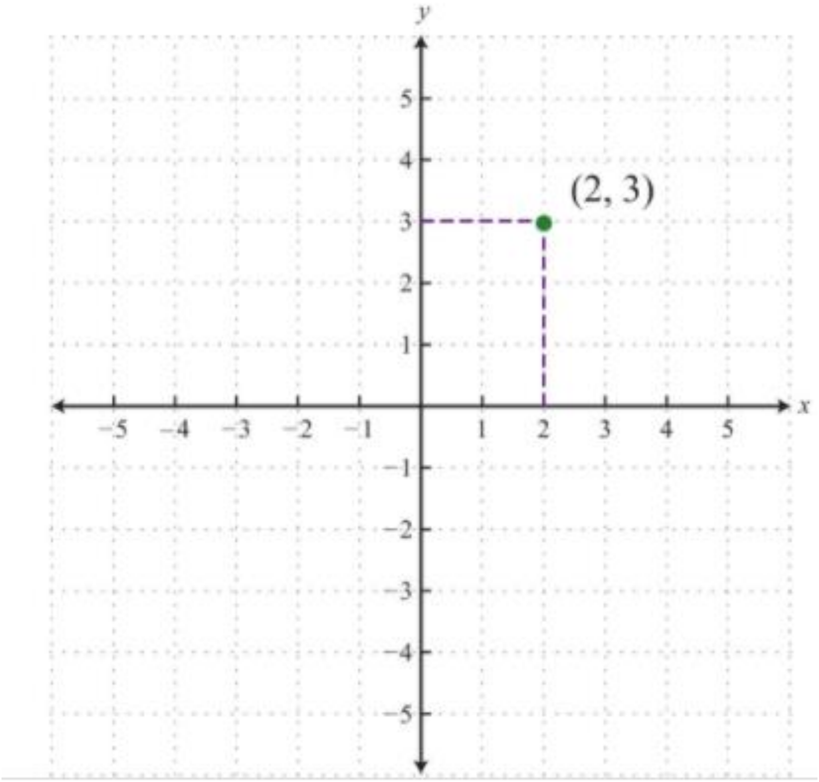
Rectangular/Cartesian Coordinate system

Coordinate system that specifies each point uniquely in a plane by a pair of numerical**coordinates**, which are the distances to the point from two fixed perpendicular directed lines, measured in the same unit of length. Each reference line is called a *coordinate axis (X axis/easting, Y axis/northing)* and the point where they meet is its *origin*, at ordered pair (0, 0).

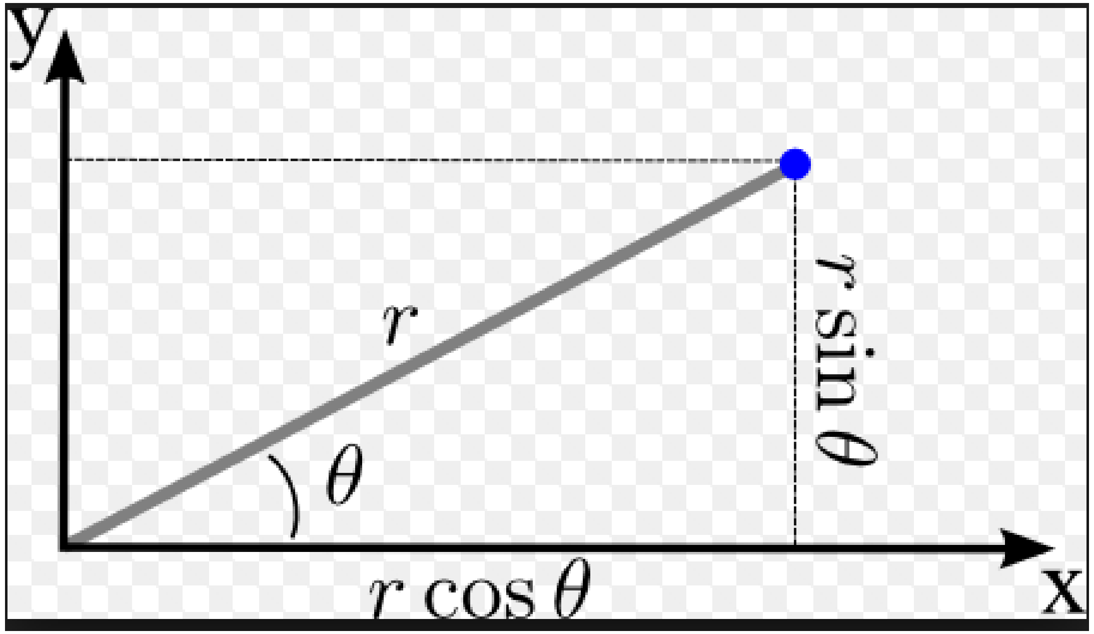
Simple system to locate objects in space

Two straight line intersecting at right angle to define the geographic space.

Divides geographic space in to 4 quadrants.



The **polar coordinate system** is a two-dimensional **coordinate system** in which each point on a **plane** is determined by a distance from a reference point and an angle from a reference direction.



For **geographic coordinate systems**, the prime meridian is the longitude that passes through Greenwich, England. The origin of the graticule (0,0) is defined by where the equator and prime meridian intersect.

The line of latitude midway between the poles is called the equator. It defines the line of zero latitude. The line of zero longitude is called the prime meridian.

Projection bridges geographical and projected coordinate system.

DATUM

Mathematical model of the earth which serves as the reference or base for calculating the geographic coordinates of a location.

It consists of a origin, parameter of the spheroid, and separation of spheroid and earth at the origin.

Clarke 1866, ground measured spheroid flattening 1/294.979
NAD27, North American Datum based on Clark 1866
GRS1980,Geodetic Reference System, satellite observation, 1/298.257
NAD83 geocentred datum, based on GRS80
WGS84 by national imagery and mapping agency, NIMA.

WGS agrees with GRS in axis, primary parameter define size and shape, secondary parameter refer to local datum refer in different countries.

A **map projection** is a systematic transformation of the latitudes and longitudes / all or part of surface of a sphere into locations on a plane.

Perspective	**Non perspective**
Geometric in nature	Perspective with modification to maintain area, distance, direction and shape

Many properties can be measured on the Earth's surface independent of its geography. Some of these properties are:

Area, Shape, Direction, Bearing, **Distance,** Scale

A. By preserved property

Conformal/ orthomorphic: Shape and angle

Equal area/equivalent : Area

Equidistant: Distance along certain lines

Azimuthal: Direction

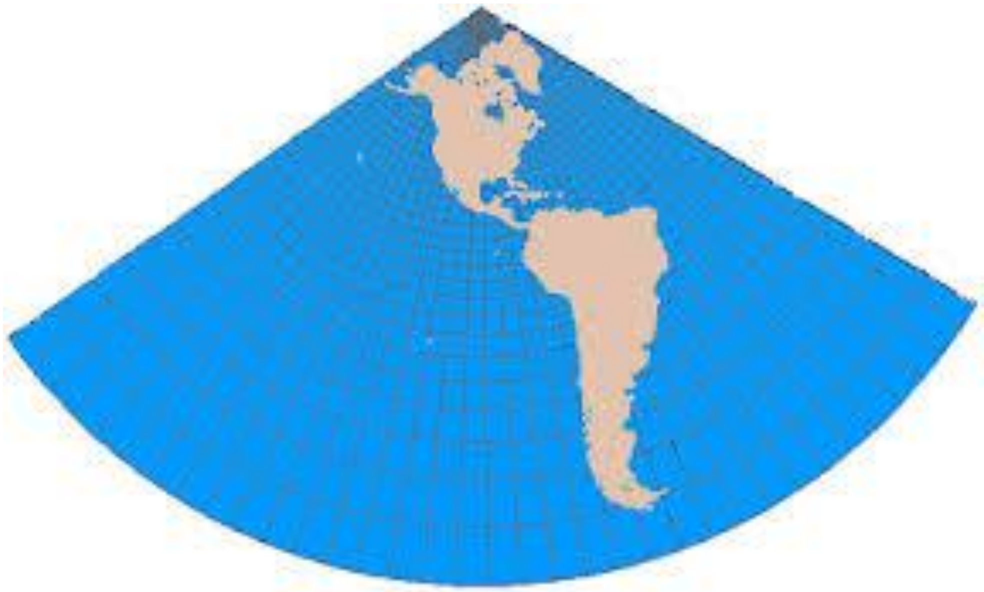
B. Projections by surface: case and aspect

The three developable surfaces (plane, cylinder, cone) provide useful models for understanding, describing, and developing map projections. However, these models are limited in two fundamental ways.

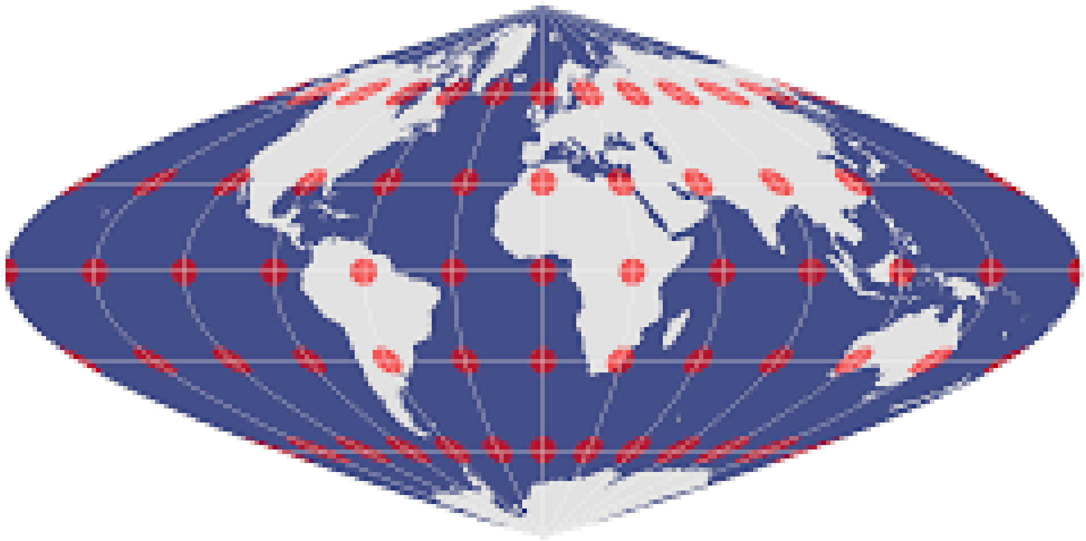
TYPES OF PROJECTIONS

1. Conformal projections
2. Equal area projections
3. Equidistant projections
4. True-direction projections

- A **conformal projection** is a map **projection** that favors **preserving** the shape of features on the map but may greatly **distort** the size of features.
- reserve Shape (only for small areas)
- Preserve Angles
- Used for Large Scale Mapping
- Distortion increases outwards from the central meridian and standard parallel Fig. Lambert conformal conic projection



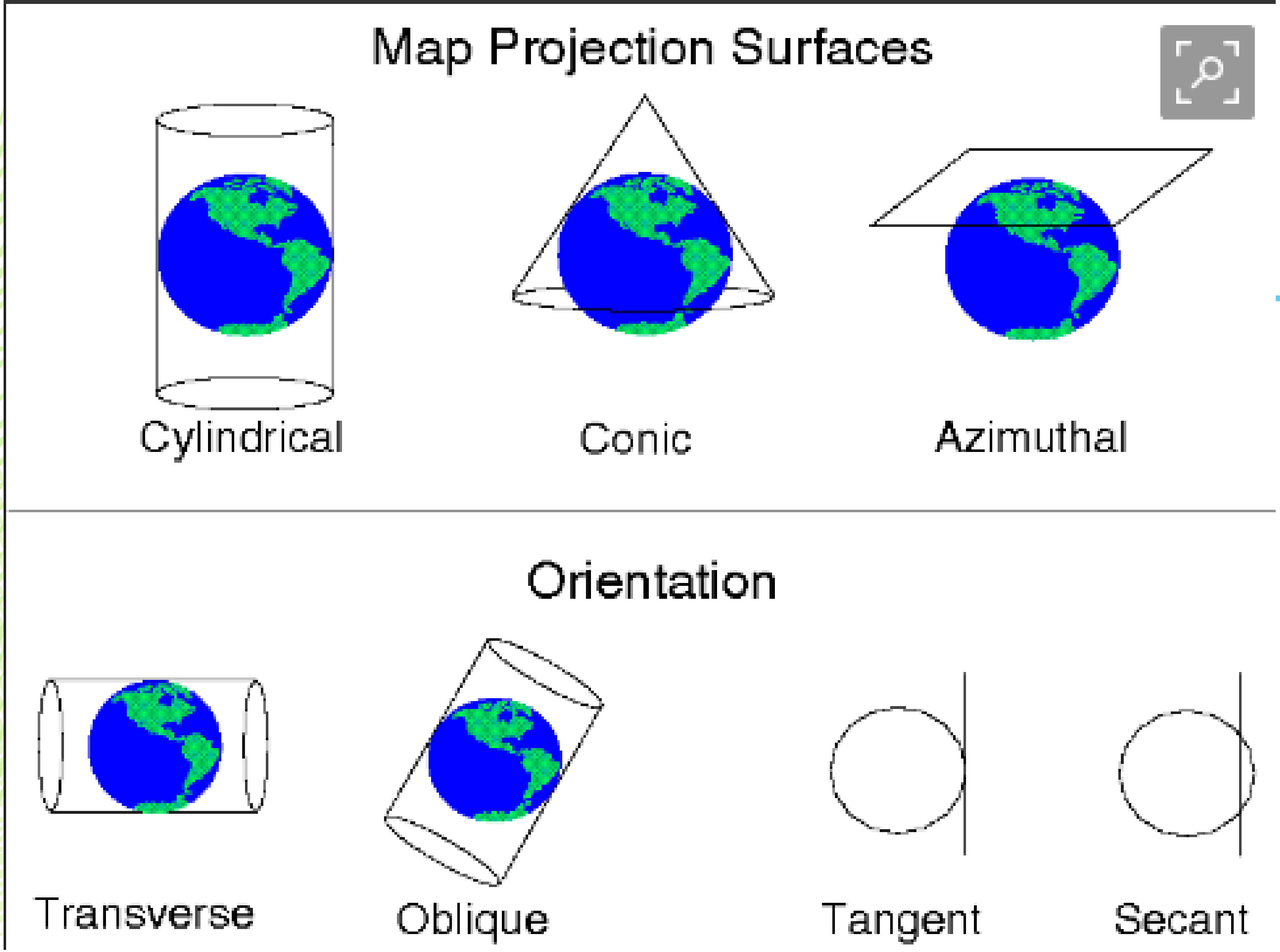
– An **equal area projection** is a map projection that shows regions that are the same size on the Earth the same size on the map but may distort the shape, angle.

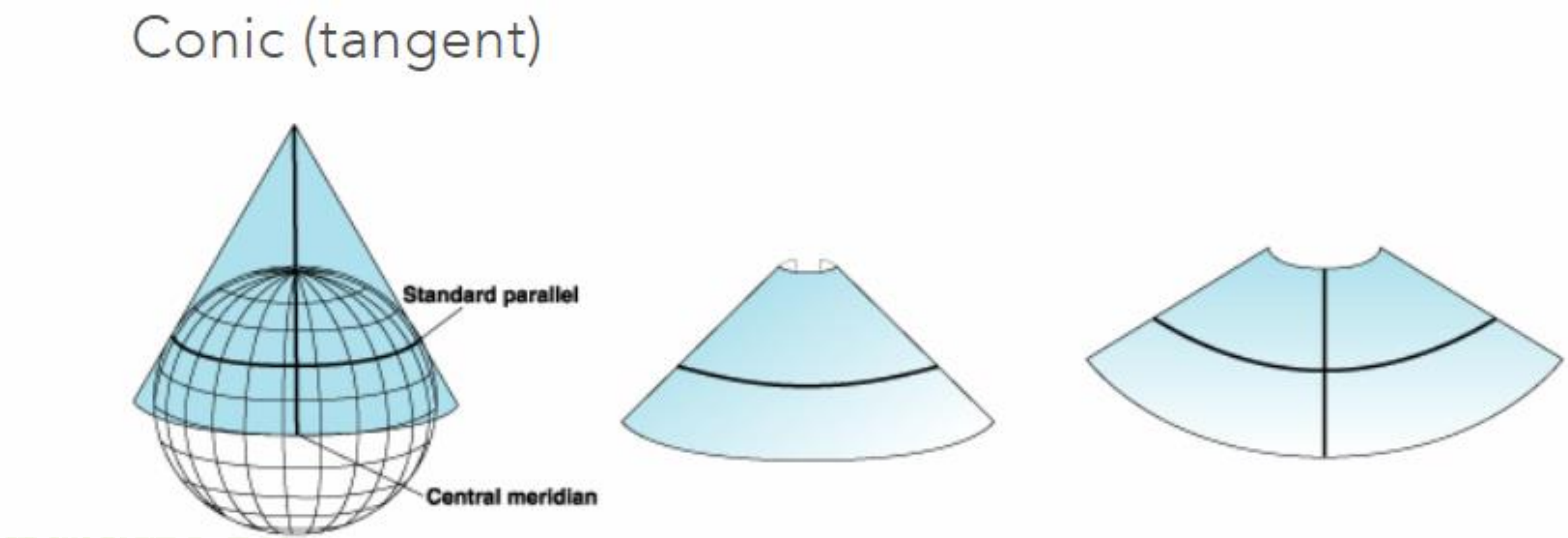


- Equidistance projections A **projection** that maintains scale along one or more lines, or from one or two points to all other points on the **map**. Lines along which scale (distance) is correct are the same proportional length as the lines they reference on the globe.

True directional projections

– Directions from a central point to all other points are maintained accurately in *azimuthal projections* (also known as *zenithal or true-direction projections*). These projections can also be equal area, conformal or equidistant.

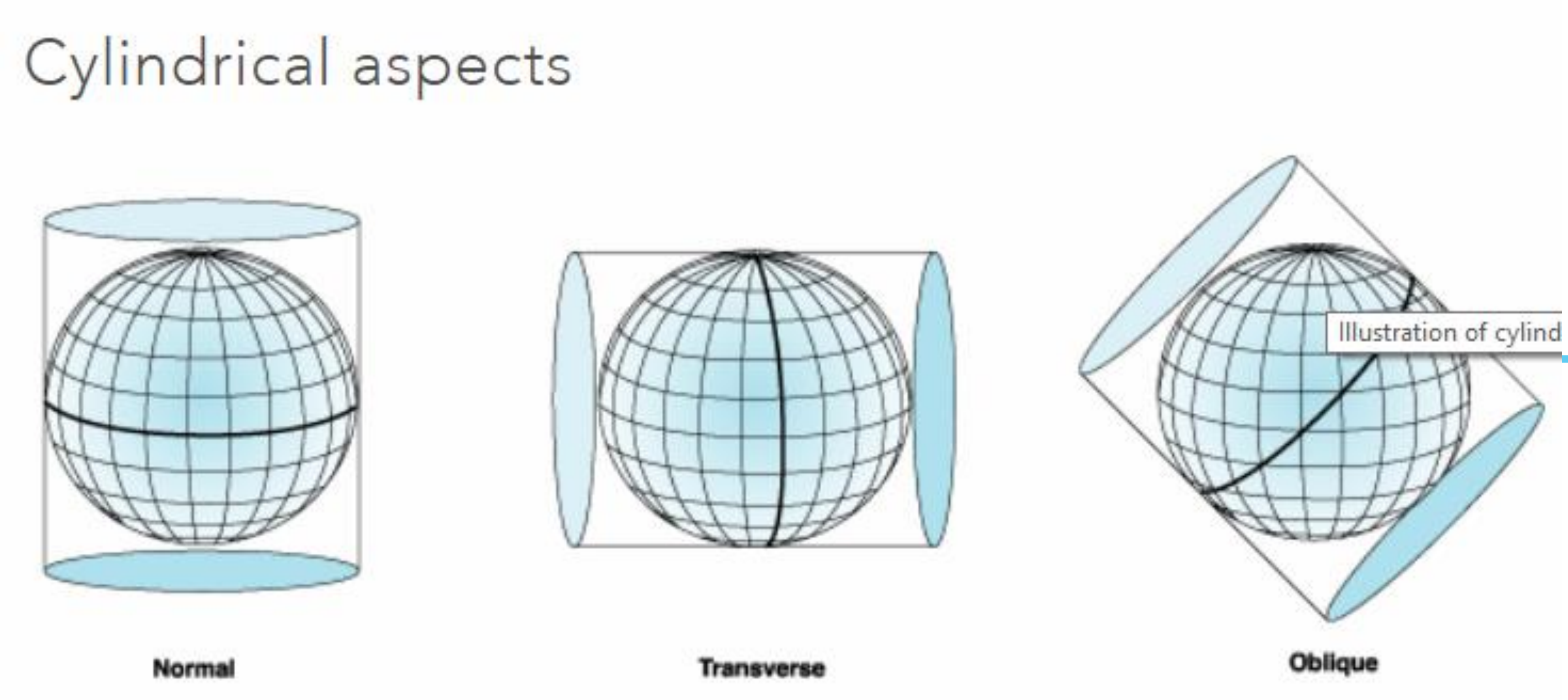

Map Projection Surfaces
Cylindrical
Conic
Azimuthal
Orientation
Transverse
Oblique
Tangent
Secant



A cone is placed over a globe. The cone and globe meet along a latitude line. This is the standard parallel. The cone is cut along the line of longitude that is opposite the central meridian and flattened into a plane

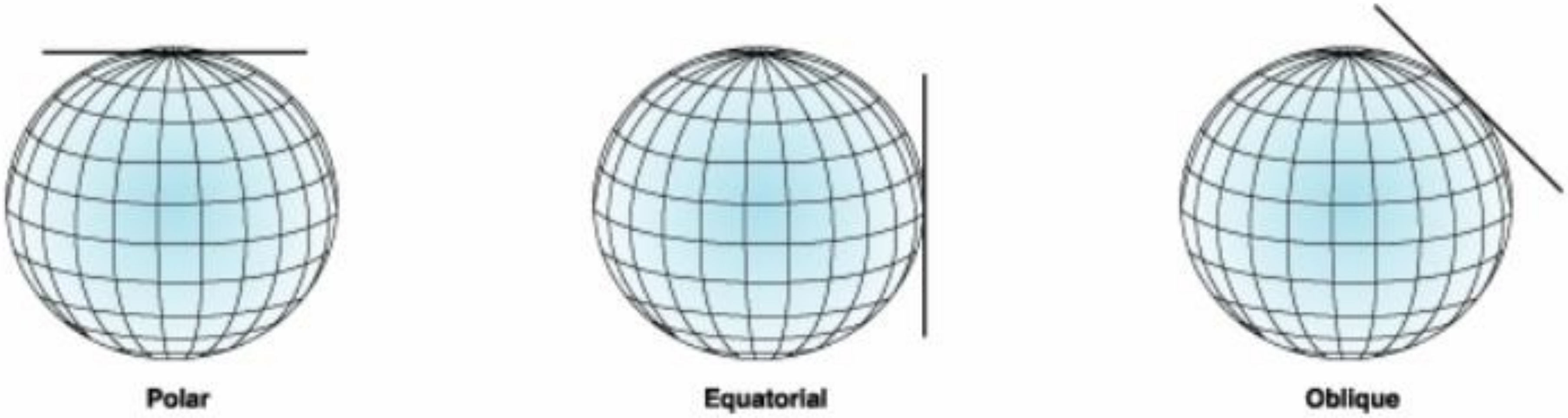
Conic (secant)

A cone is placed over a globe but cuts through the surface. The cone and globe meet along two latitude lines. These are the standard parallels. The cone is cut along the line of longitude that is opposite the central meridian and flattened into a plane.

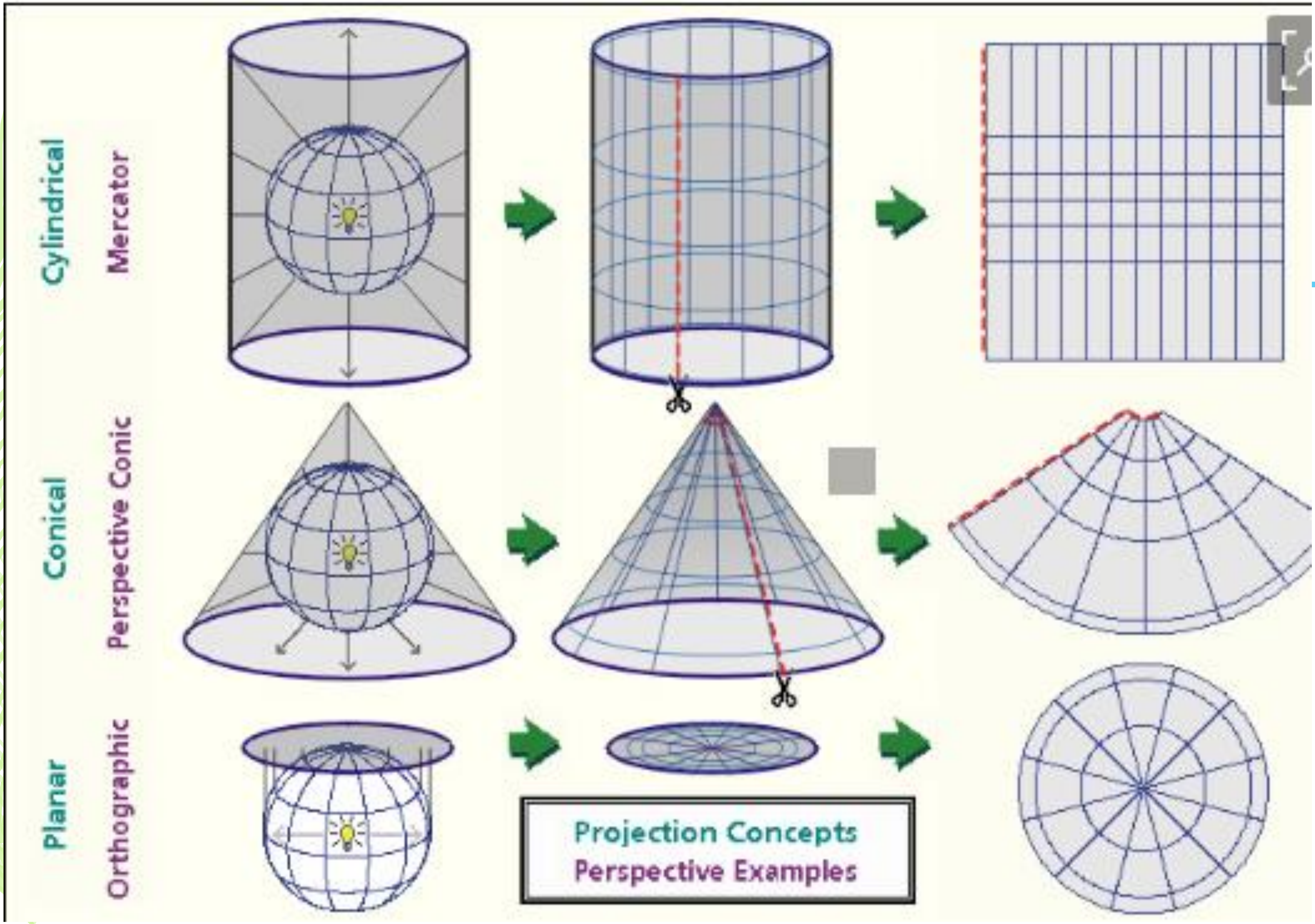


A cylinder is placed over a globe. The cylinder can touch the globe along a line of latitude (normal), a line of longitude (transverse), or another line (oblique)

Planar aspects



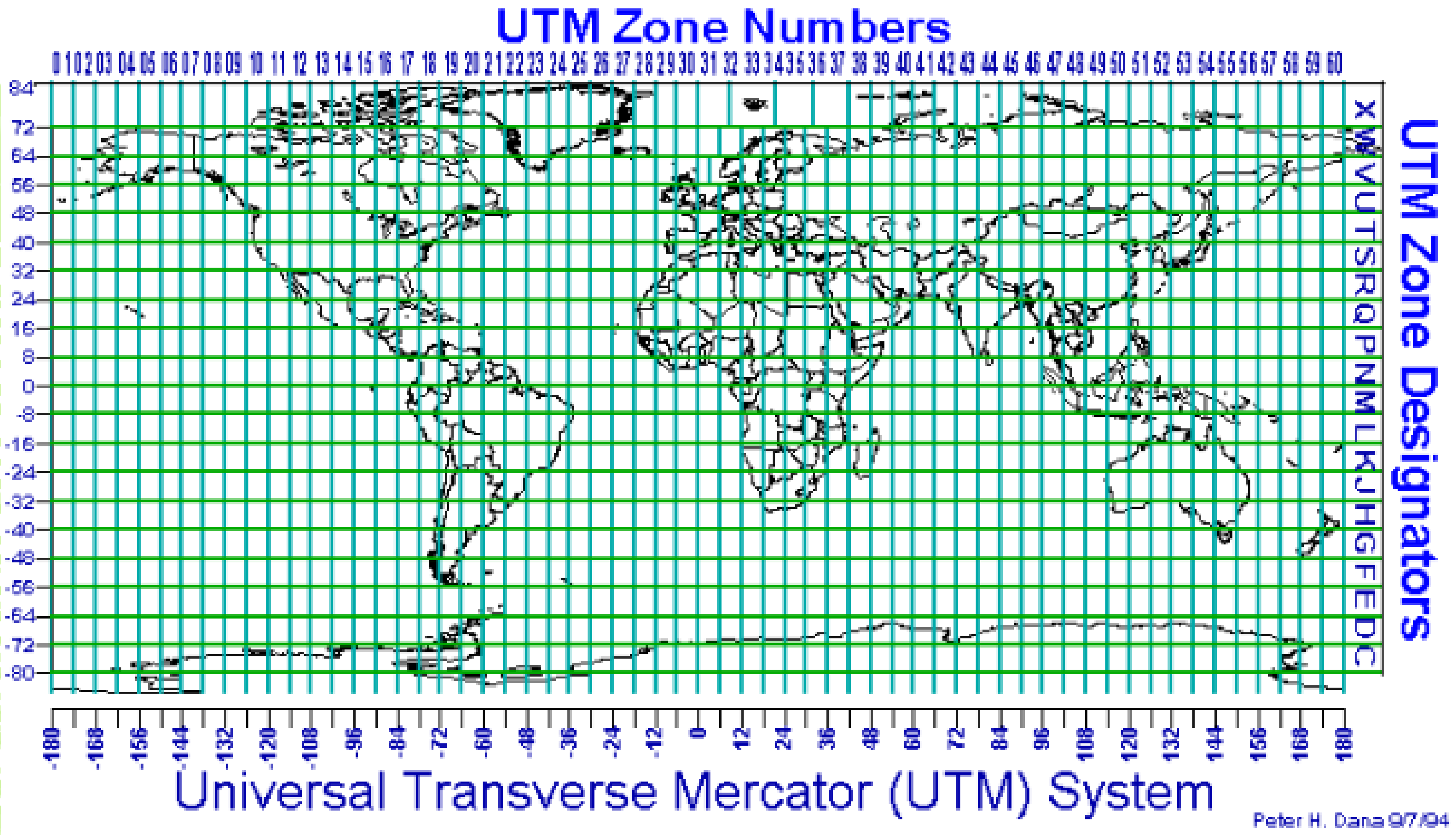
A plane is placed over a globe. The plane can touch the globe at the pole (polar), the equator (equatorial), or another line (oblique)


Cylindrical
Mercator
Conical
Perspective Conic
Planar
Orthographic
Projection Concepts
Perspective Examples

Universal Transverse Mercator (UTM) projection

UTM projection uses a 2-dimensional Cartesian coordinate system to give locations on the surface of the Earth.

UTM system is not a single map projection. The system instead divides the Earth into sixty zones, each being a six-degree band of longitude, and uses a secant transverse Mercator projection in each zone.



LONGITUDE 6 DEGREE INTERVAL

LATITUDE 8 DEGREET EXCEPT AT 72-84 IN 12 DEGREE

60 ZONES (EXCEPT I AND O)

GIS DATA TYPES and REPRESENTATION

- The basic data type in a GIS reflects traditional data found on a map. Accordingly, GIS technology utilizes two basic types of data. These are:
- **Spatial data** describes the absolute and relative location of geographic features.
- **Non-spatial or Attribute data** describes characteristics of the spatial features. These characteristics can be quantitative and/or qualitative in nature. Attribute data is often referred to as tabular data.
- The coordinate location of a forestry stand would be spatial data, while the characteristics of that forestry stand, e.g. cover group, dominant species, crown closure, height, etc., would be attribute data. Other data types, in particular image and multimedia data, are becoming more prevalent with changing technology.

VECTOR DATA format

Vector features (geographic objects with vector geometry) are a versatile and frequently used geographic data representation, well suited for representing features with discrete boundaries, such as wells, streets, rivers, states, and parcels.

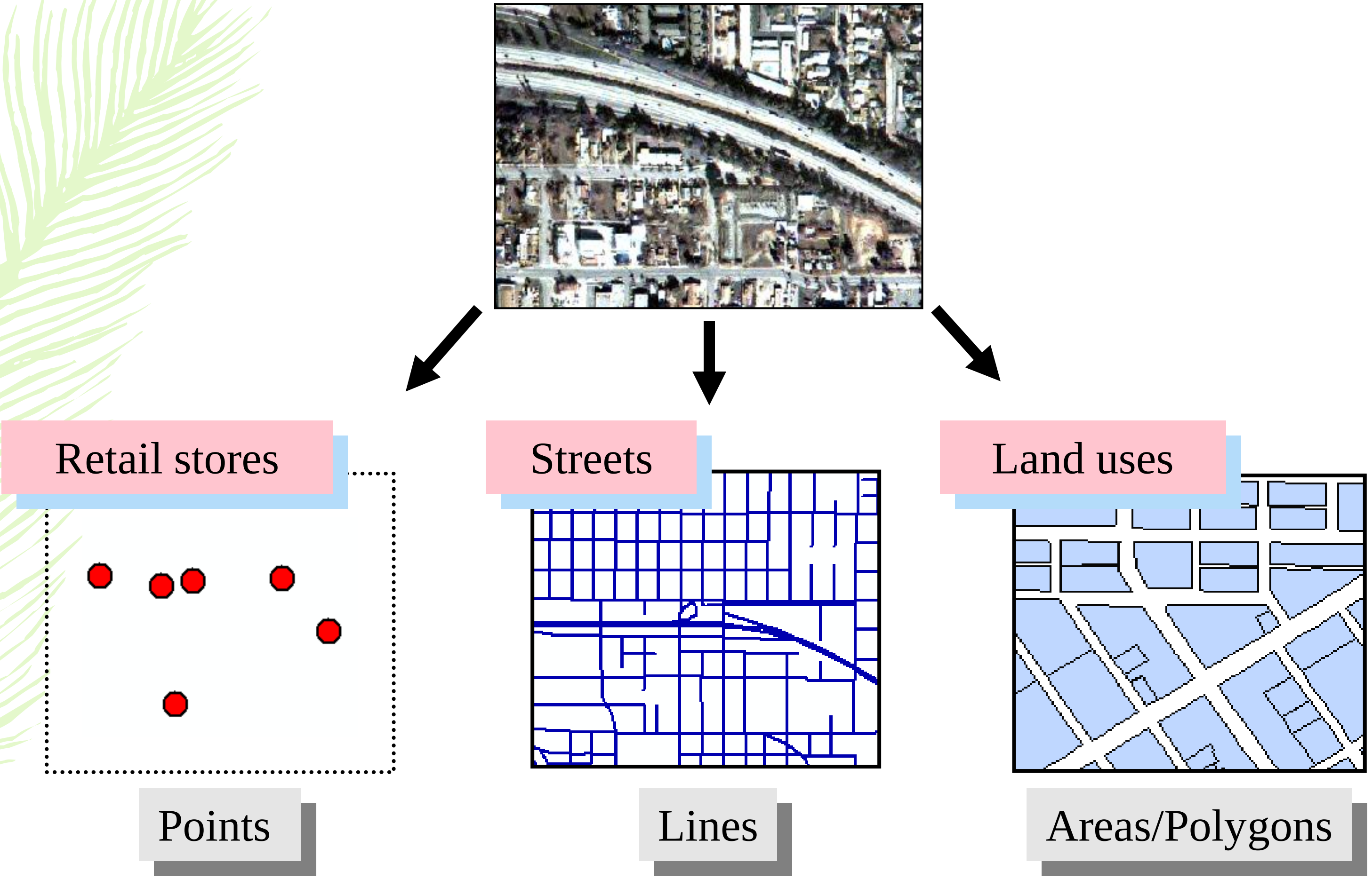
A feature is simply an object that has a location stored as one of its properties (or fields) in the row.

Features are spatially represented as points, lines, polygons, or annotation and are organized into feature classes.

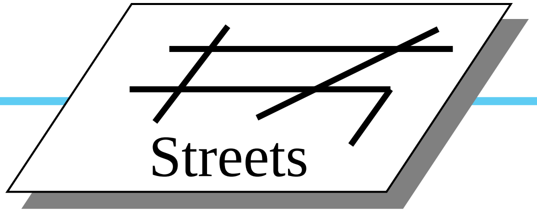
Feature classes are collections of features of the same type with a common spatial representation and set of attributes (Eg. Line feature class for Roads).

Representing features in vector data

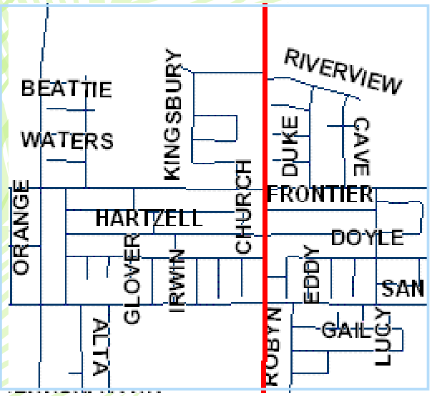
– Real-world entities are abstracted into three basic shapes



– Three general components to geographic information



Geometry



Attributes

Attributes of street

STR_NAME	STR_TYPE
CONE CAMP	RD
CHURCH	ST
OPAL	RD
CHURCH	ST
DISHONG	ST
STATE 30	HWY
STATE 30	HWY
STATE 30	HWY
STATE 30	HWY
OPAL	AV
OPAL	AV

Record: 1

Behavior

Rules:

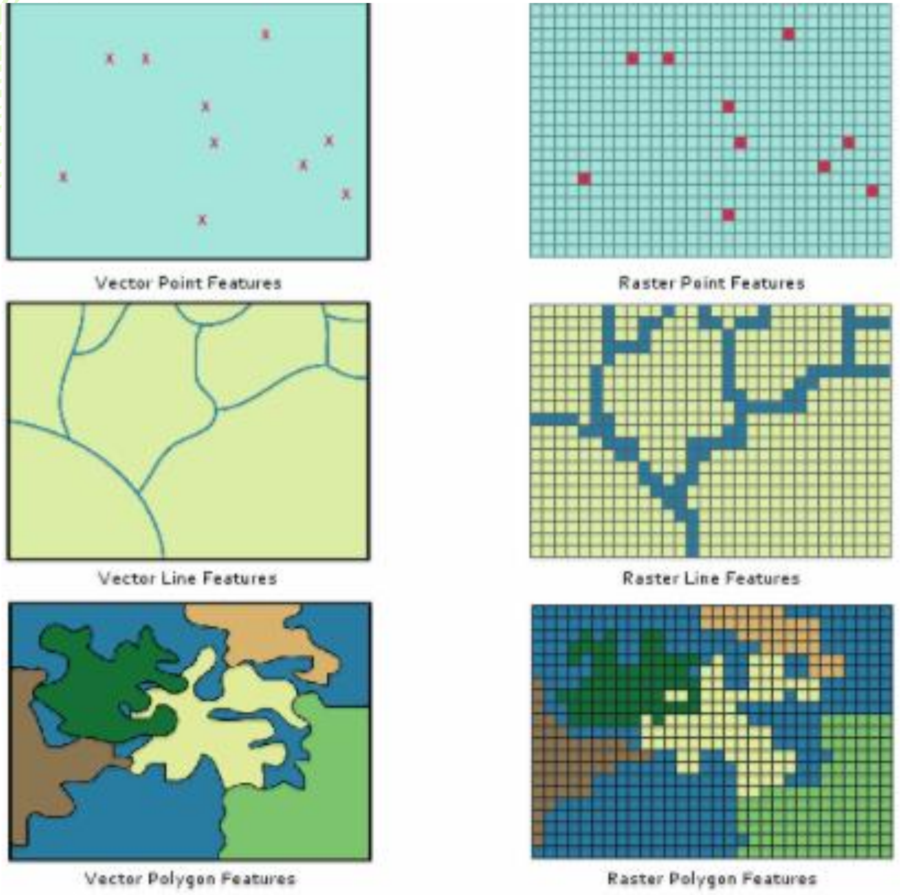
Streets and highways may not intersect

– Each feature corresponds to a record in the attribute table

RASTER DATA

Rasters are used to represent continuous layers, such as elevation, slope and aspect, vegetation, temperature, rainfall, and plume dispersion.

Rasters are most commonly used for aerial photographs and imagery of various kinds.



ATTRIBUTE DATA
VERSUS
SPATIAL DATA

ATTRIBUTE DATA	SPATIAL DATA
Characteristics of geographical features that are quantitative and/or qualitative in nature	All types of data objects or elements that are present in a geographical space or horizon
Town planning and management departments, fire departments, environmental groups and online media help to obtain attribute data	Satellite images and scanned maps help to obtain spatial data
Describes the characteristics of a geographical feature	Describes the absolute and relative location of a geographic feature

Visit www.PEDIAA.com

INPUT DATA FOR GIS

Cover all aspect of capturing spatial and attribute data

Spatial data sources:

Existing map: Digitization manual or digital, DEM creation etc

Aerial photos: analog and digital photogrammetry

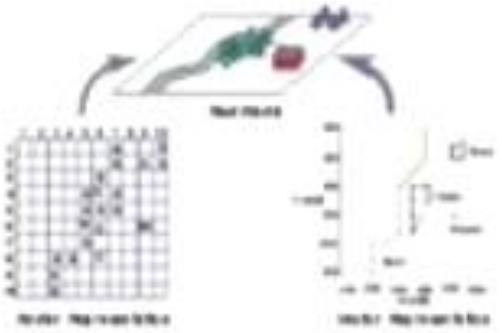
Satellite imageries:

Field Observations: GPS

Other Sources

SPATIAL DATA SOURCES

- **Primary Data Sources** (first-hand collection)
  - Digitizing
    - Tablet digitizing
    - Heads up digitizing
    - Automatic digitizing



  - Scanning
  - Other point measurements
  - Census data
  - GPS collections
  - Aerial photographs
  - Remote sensing data

Secondary Data Sources (from others)

- Published or released data (originally primary data)
- All primary data from others are secondary data for you and me
- Large amount of data is now available
- Several groups of data exist
  - Free data from the government
  - Government data available for a fee
  - Internet map servers
  - Commercial data
  - Data from other GIS users

DATA INPUT METHODS

There are several ways to turn analogue data into digital data:

1. Keyboard Entry – Entering data (into a file) by way of a computer terminal. Example: entering names from a sign-in sheet
2. Manual Digitizing – Requires a digitizing table connected to a computer station. Example: digitizing mountain tops from an aerial photo
3. Automatic Digitizing
4. Scanning is the most common. Example: scanning an old map or photo into digital format

(a) Manual digitizing – digitizing tablet:

- A digitizer is a special table with a grid of fine wiring behind the face.
- You place your map on the digitizer and then secure it with removable sticking tape.
- In your digitizing software, you then specify a coordinate system
- Then you digitize the bounding coordinates of the map.
- After the setup phase, you trace the map features using a magnetic pen (puck). The map features get sent to your computer as a GIS map.
- Digitizing using a digitizing tablet can be carried out in the following modes:
- Point mode: single points are recorded at one time.
- Stream mode: one point each is collected at regular breaks of time and distance.

1. (b) Manual digitizing – onscreen heads-up

- **Scan a map to create a digital version of it.** For small maps, the flatbed scanner sitting on the desk next to you is good enough. You will need to scan bigger maps at an office centre.
- **Place the scan into a coordinate system**. You do this using a GIS process called georectification. Simplistically this involves matching features on the scanned map to identical features on a GIS map.
- **On-screen digitize using the georectified map as a backdrop.** Use the georectified scanned map as a backdrop on your computer screen and then trace the features on the scan with your mouse. The features will automatically be in a coordinate system, and so, once attributed, a GIS map.

OHT 18

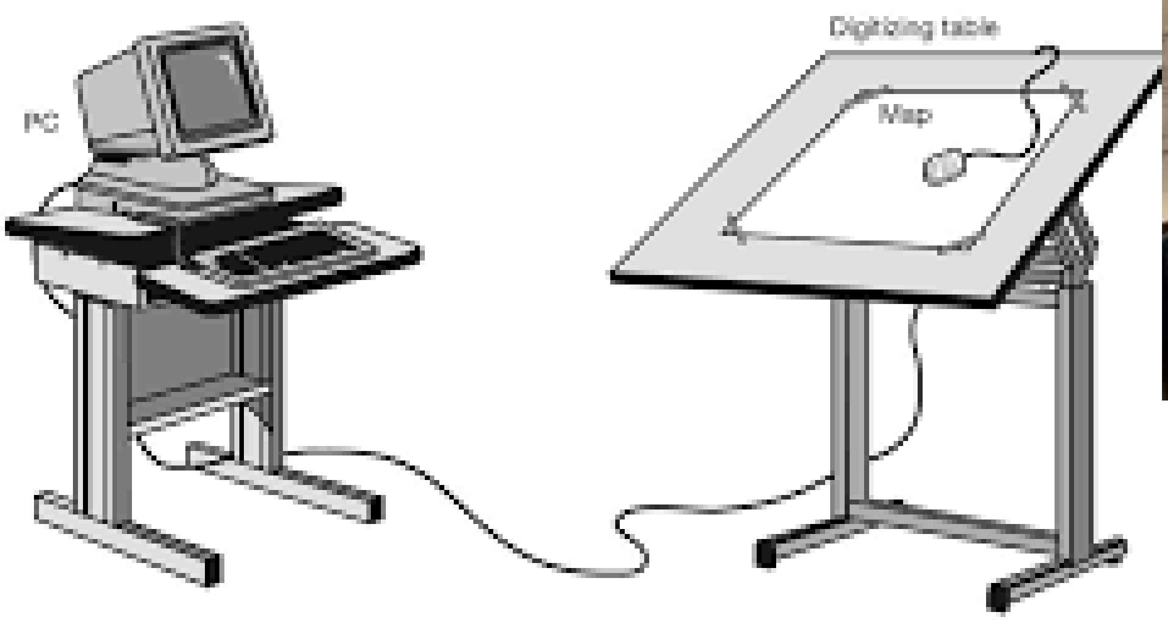


Figure 5.2 Digitizing table and PC workstation

Heywood et al. An Introduction to Geographical Information Systems, 2/e

Automatic Digitizing

Automatic Digitizing is the process of converting raster to vector in an automated method using pattern recognition and image processing techniques. In this technique, the computer traces all the features on the map; it gives high accuracy with low time consumption. It allows customization and improved quality of images. This process is also known as Vectorisation.

Automatic scanning:

- This is where a scanner captures the spatial data automatically. Scanning devices can capture spatial data at a very high speed. However, these scanners are often quite expensive and are rarely able to recognize text and symbols. Scanned data also requires a large amount of editing, which is done manually, therefore it can be more time consuming.

- Automatic scanning can also be very quick when the map being scanned is a single map theme eg. just roads or just watercourses (but not a map with a mix of themes such as roads, rail, watercourses, etc.)

- Scanners are expensive and its better to outsource this task. The final scans are only as good as the original map. Low quality scans might need to be on-screen digitized rather than going through feature auto-recognition.

Types of Errors in Digitization

1. Geodetic Errors

– Geodetic errors are due to the odd choice of a projection system. Generally, earth features are in 3 Dimensions. But the features on the map are in 2 Dimensions. So the projection system leads to inaccuracy. Improper projection leads to inaccurate placement of elements on the map. Therefore making the map and the digitized features do not overlap each other appropriately.

2. Machine Error

– Machine error occurs due to the digitizing tablet or the software used to digitizing the elements. It is an inherent error that cannot be removed but can only be minimized. Sometimes it may occur when converting the maps from analog to the digital formats.

3. Cartographic Errors

– Cartographic errors arise due to existing mistakes that are present in the source map itself, and it can be transferred into the digital map.

– Incorrect interpretations or drafting of the elements in the maps are also one of the reasons for these errors.

- **4. Manuscript Errors**
- Manuscript errors occur based on the quality of the source maps. Hard copy maps shrink with time. Any stretching, warping, wrinkling, or traces of folding of the original map might affect the digitization process. It may lead to irregular shape, area & coordinates of the digitized features. It can’t be completely rectified.
- **5. Positional Errors**
- Positional Error happens when an element is not captured correctly or carelessness of the digitizer, and it can be completely rectified.
- Positional Error is categorized as Dangling Nodes, Switchbacks, knots, and loops, Overshoots, and Undershoots, Silvers, and Overlaps.

Dangling Nodes

– Dangling nodes occur in polygon & polyline. In digitized polyline, which hasn’t met, or there is any gap between the nodes. In a digitized polygon, it happens when a polygon doesn’t connect back to it.

Switchbacks, Knots & Loops

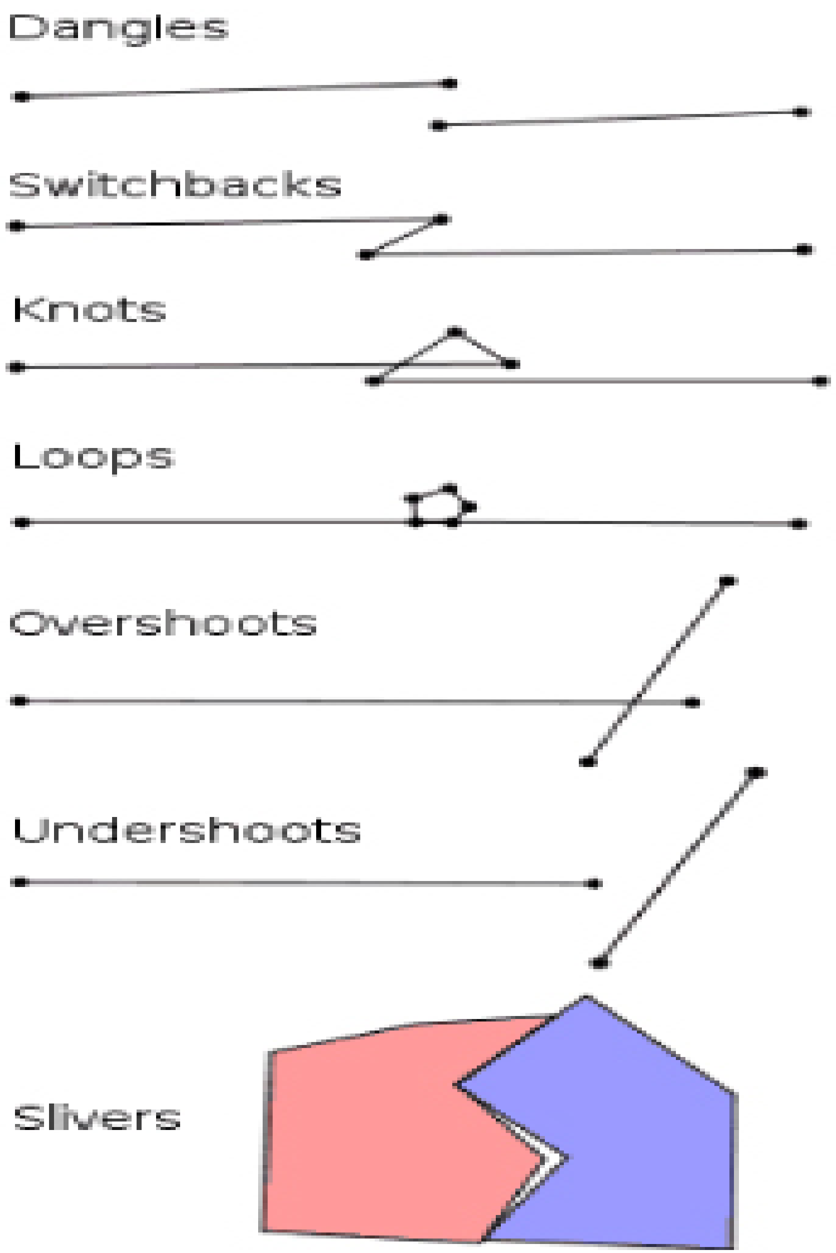
– These errors occur due to unpractised digitizer. With the wrong movement of the cursor/ puck, the line being digitized ends up with extra vertices/nodes. In the case of additional vertices, switchbacks are formed in line with a bend. With knots and loops, the line folds onto itself, creating a polygon. Also known as a weird polygon. It also occurs in both polygon & polyline.

Overshoots and Undershoots

– When a digitized line does not connect properly, overshoots and undershoots errors occur. During Digitization, a snap tolerance or snap distance (snap tolerance or the snap distance is the measurement of the diameter from the point of the cursor) is set by any digitizer.

– The main reason for undershoots and overshoots error occurs is due to the snap distance being set too low, or the digitizer does not set the snap tolerance. It occurred both in lines & polygons.

- Silvers errors occur due to improper snap tolerance values. Because of this, gaps formed between two adjoining polygons. Setting the snap tolerance value properly it can be avoided entirely. It occurs only in polygons.
- Overlap errors are formed between two adjoining polygons, which overlap each other. It is due to the improper snap tolerance value, and it can be avoided by setting the proper tolerance value. It occurs only in polygons.


Dangles
Switchbacks
Knots
Loops
Overshoots
Undershoots
Slivers

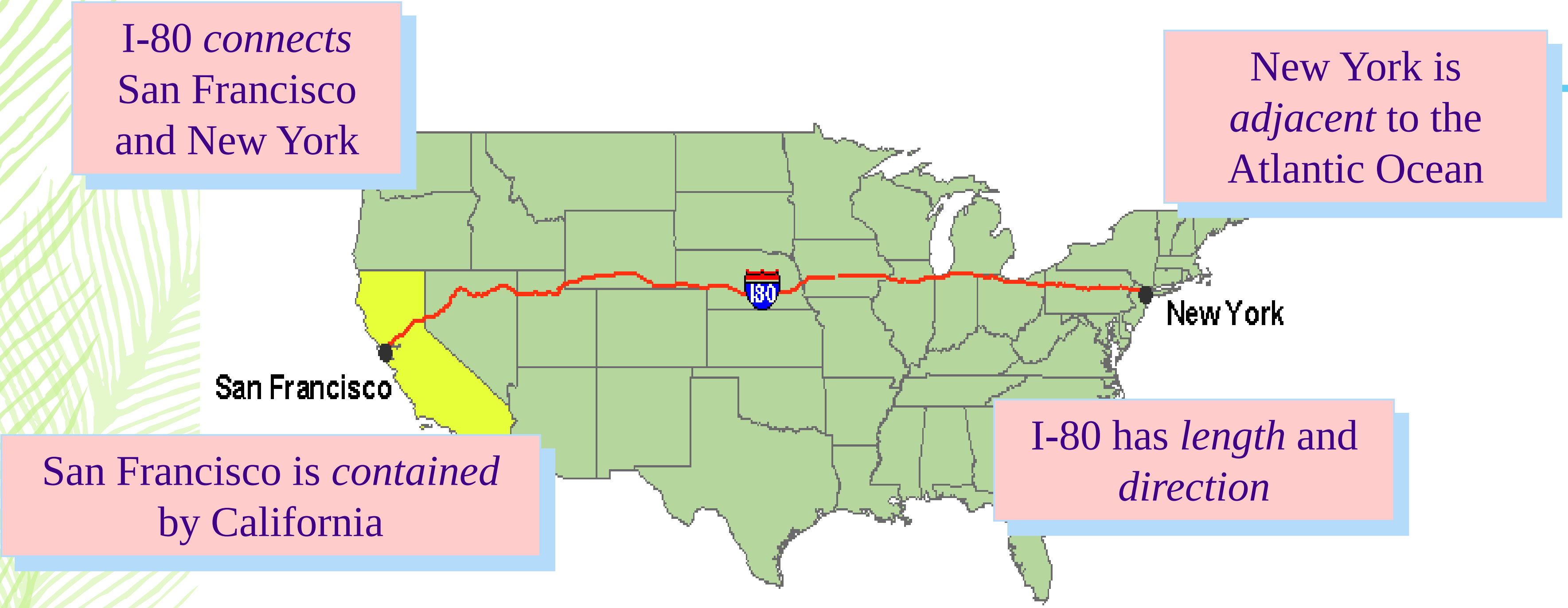
- **6. Attribute Errors**
- These errors occur while entering the attribute information about the feature. Attributes of the features are updated after digitizing the elements. These errors may lead to change characteristics of the features or digitized data.

TOPOLOGY

- **Topology** expresses explicitly the spatial relationships between features.
- **Topology** is the study of those properties of geometric objects that remain invariant under certain transformations such as bending or stretching.

– The relative position of features determines relationships

Using spatial relationships



Topology mathematically models connectivity, adjacency, and coincidence

- The coverage model supports three basic topological relationships:
  - **Connectivity:** Arcs connect to each other at nodes.
  - **Area definition (Containment):** An area is defined by a series of connected arcs.
  - **Contiguity (Adjacency):** Arcs have directions and left and right polygons.

Importance of Topology

- Topology-based data sets require additional data files to store the spatial relationships between features.
- Topology has at least two main advantages.
  - The first is the assurance of data quality.
    - It enable us to detect lines that do not meet correctly
    - polygons that are not closed properly.
    - avoid incomplete features and ensure data integrity.
  - Second, topology can enhance GIS analysis.

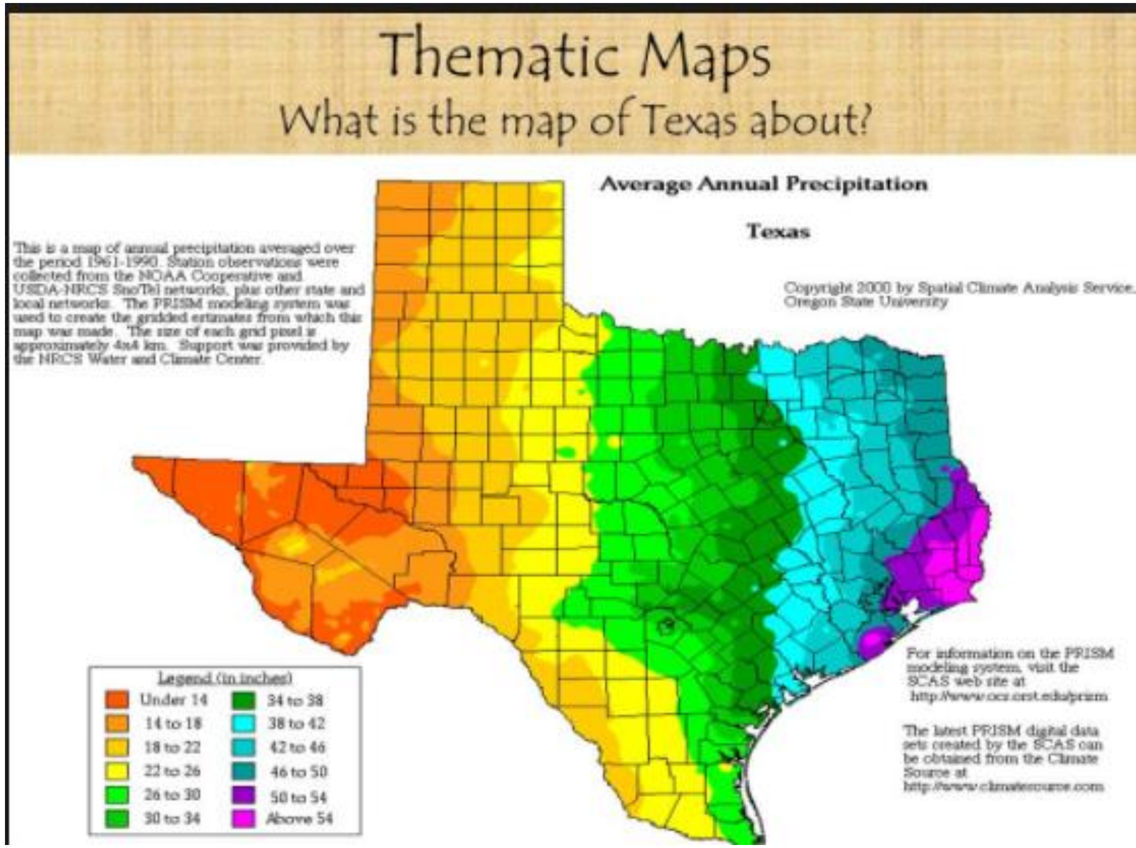
TYPE OF OUTPUT PRODUCTS

Computer screen	Printer	Paper / film plotters	optical and magnetic media
Maps	Tables	Graphs & Chart	video

Thematic maps: a type of **map** especially designed to show a particular **theme** connected with a specific geographic area.

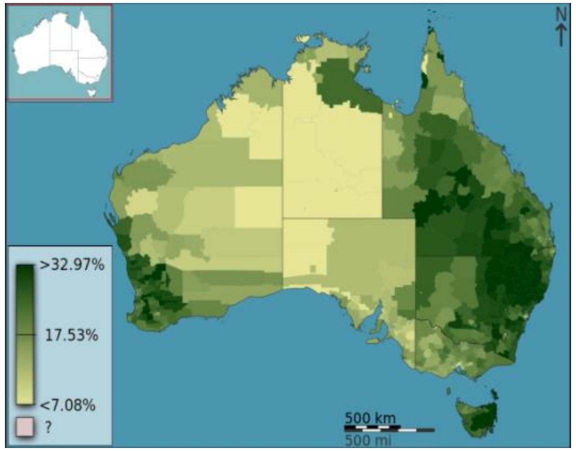
- Represents one attribute or relation between them
- visual representation of characteristics of a given geographic location.

Eg: landcover, population.



Choropleth map: is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable.
Geographical areas or regions that are coloured, shaded or patterned in relation to a data variable

Eg. Population density, per capita income



Contour map: Line connecting points of equal elevation

Other type of maps:

Dot map: spatial distribution of uniform dots, eg. Population

Line map: Show direction/attribute of actual flow

Animated map: to depict sequence through time, eg. Growth of city with increase in population over time

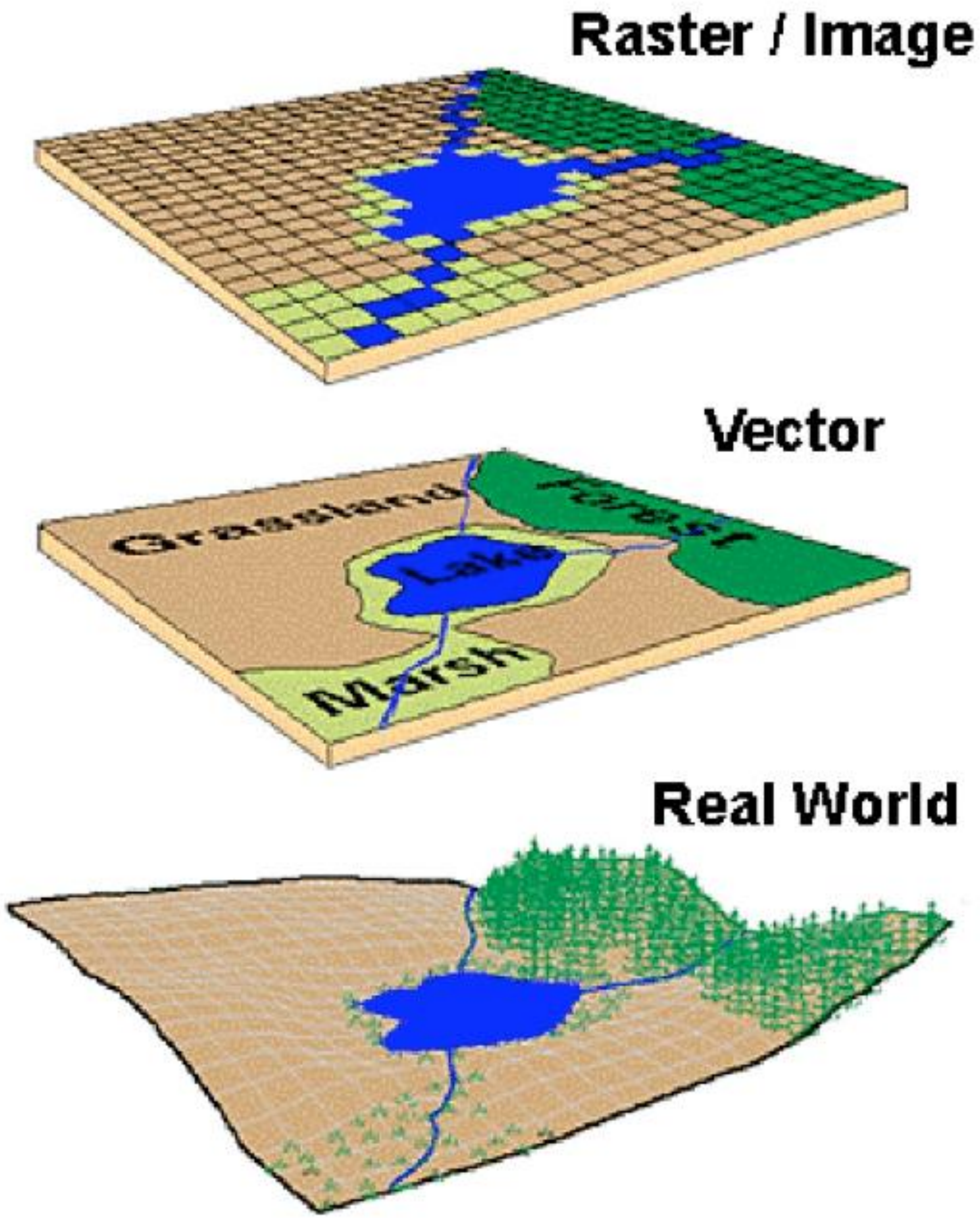
Charts

Bar chart: Illustrate difference in attribute between category (landuse over time)

Pie chart: Displaying information by dividing the circle

Scatter plot: Behavior of 1 attribute *vs other*, *eg*. *Crop vs Fertiliser*

Histogram: Exam the way a attribute change with different possible value, eg. Education / primary, secondary etc


Raster / Image
Vector
Grassland
Forest
Lake
Marsh
Real World

Vector data model **use point and their x, y coordinates** to construct spatial features of point, line and area. Vector data ideal for representing discrete feature and raster ideal for representing continuous feature

Vector data model :

Georelational – Split system (Simple /Compound – TIN, Region, Route/dynamic segmentation models)

Object based data model – single system

	Geo-relational	**Object based**
Topological	Coverage	Geodatabase
Non Topological	Shape file	Geodatabase

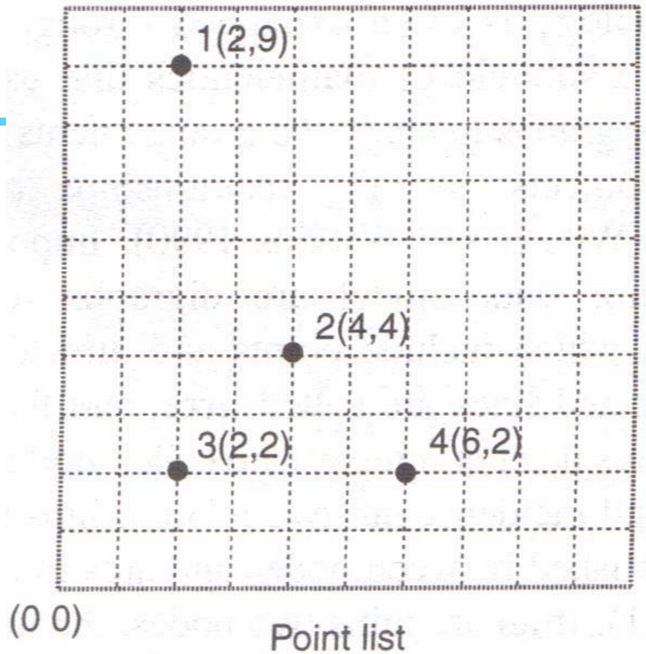
Data formats by ESRI based on topology and data model

ESRI's Coverage Model

- **Coverage** is a topology-based vector data format.
- The coverage model supports three basic topological relationships:
  - **Connectivity:** Arcs connect to each other at nodes.
  - **Area definition (Containment):** An area is defined by a series of connected arcs.
  - **Contiguity (Adjacency):** Arcs have directions and left and right polygons.

Coverage Data Structure

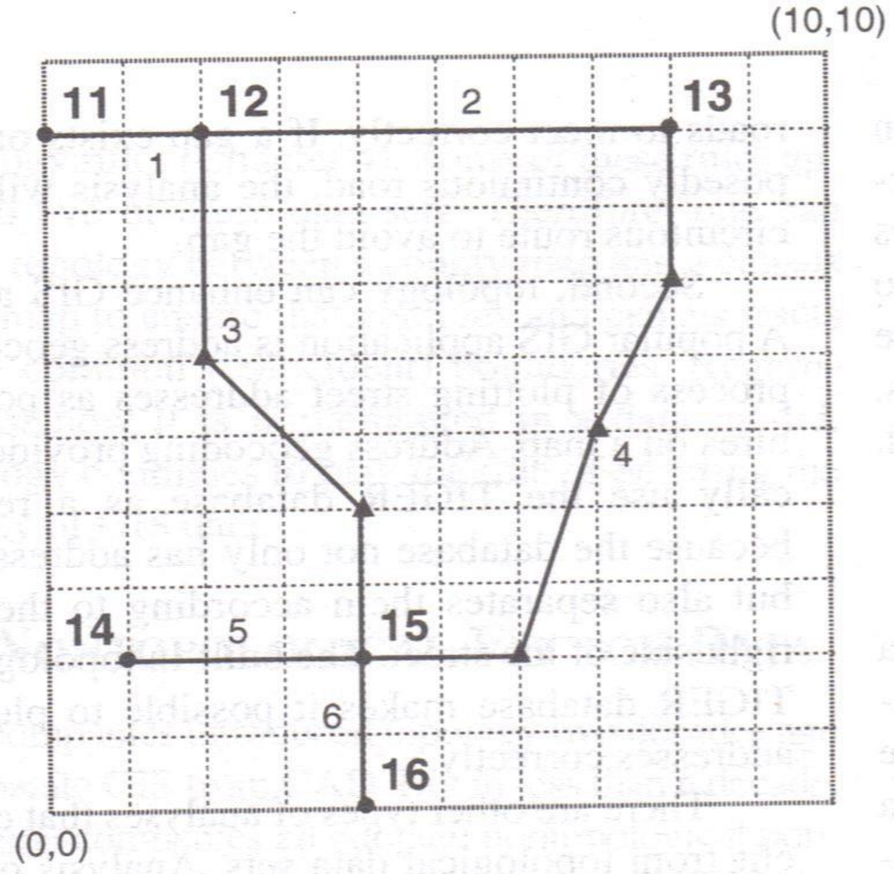
- The coverage model incorporates the topological relationships into the structure of feature data.
- The data structure of a point coverage contains feature identification numbers (IDs) and pairs of *x*- and y-coordinates.



Point list

ID	*x,y*
1	2,9
2	4,4
3	2,2
4	6,2

The data structure of a line coverage



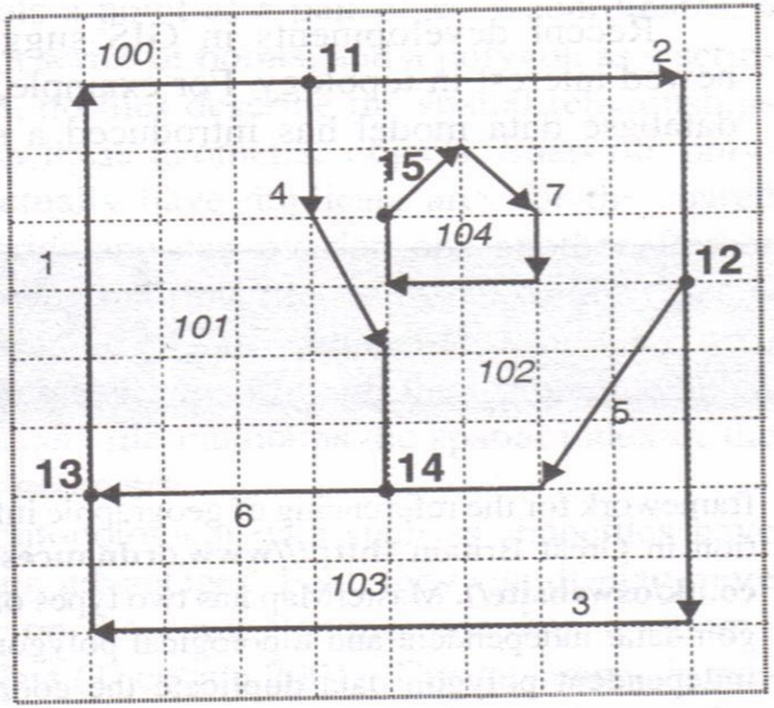
Arc-node list

Arc#	F-node	T-node
1	11	12
2	12	13
3	12	15
4	13	15
5	15	14
6	15	16

Arc-coordinate list

Arc#	*x,y* Coordinates
1	(0,9) (2,9)
2	(2,9) (8,9)
3	(2,9) (2,6) (4,4) (4,2)
4	(8,9) (8,7) (7,5) (6,2) (4,2)
5	(4,2) (1,2)
6	(4,2) (4,0)

The data structure of polygon coverage



Left/right list

Arc#	L-poly	R-poly
1	100	101
2	100	102
3	100	103
4	102	101
5	103	102
6	103	101
7	102	104

Polygon/arc list

Polygon #	Arc#
101	1,4,6
102	4,2,5,0,7
103	6,5,3
104	7

Arc-coordinate list

Arc#	x,y Coordinates
1	(1,3) (1,9) (4,9)
2	(4,9) (9,9) (9,6)
3	(9,6) (9,1) (1,1) (1,3)
4	(4,9) (4,7) (5,5) (5,3)
5	(9,6) (7,3) (5,3)
6	(5,3) (1,3)
7	(5,7) (6,8) (7,7) (7,6) (5,6) (5,7)

Raster data model uses **grid and grid cells** to represent spatial variation of a feature.

A raster represents a continuous surface, but for data storage and analysis, a raster is divided into rows, columns, and cells.

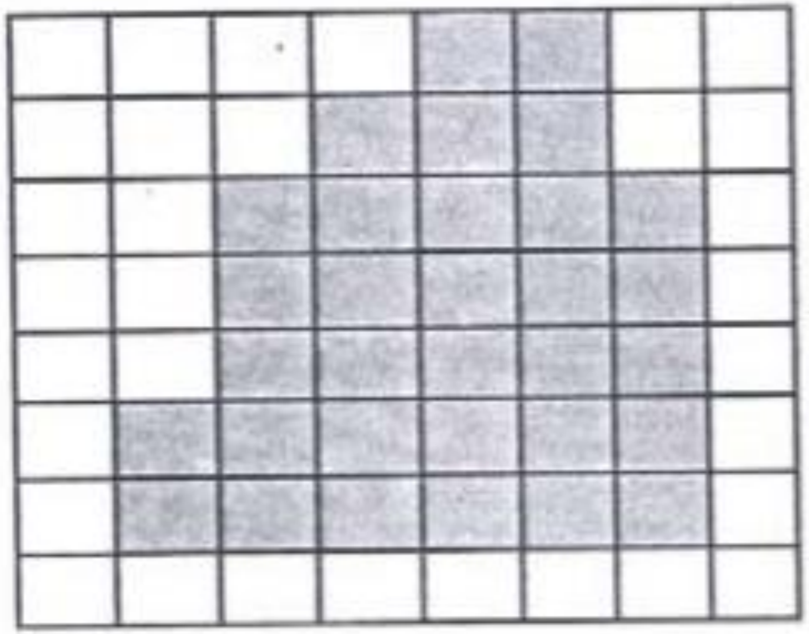
Raster data represent points by single cells, lines by sequences of neighboring cells, and areas by collections of contiguous cells.

Raster Data Structure

Refers to storage of raster data so that it can be processed by the computer.

Cell-by Cell Encoding

A raster model is stored as a matrix. Its cell values are written into a file by row and column. Ideal to store the cell values that change continuously, e.g.,DEM.



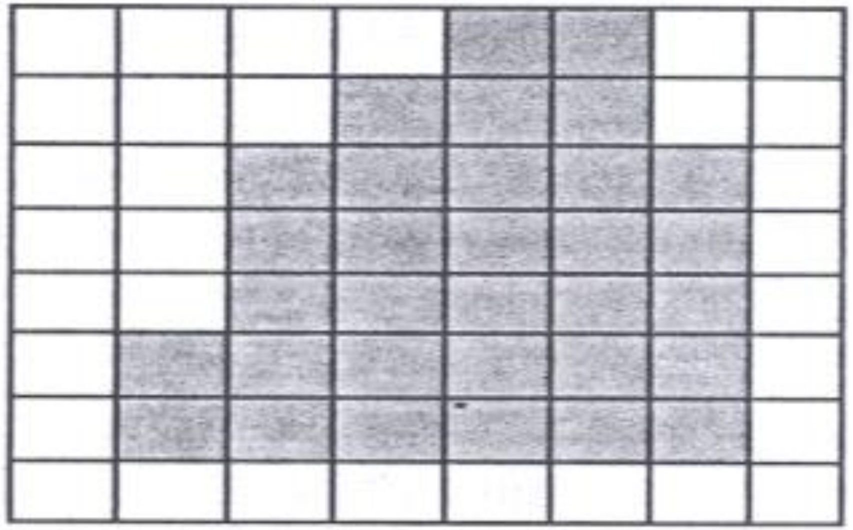
Row 1: 0 0 0 0 1 1 0 0
Row 2: 0 0 0 1 1 1 0 0
Row 3: 0 0 1 1 1 1 1 0
Row 4: 0 0 1 1 1 1 1 0
Row 5: 0 0 1 1 1 1 1 0
Row 6: 0 1 1 1 1 1 1 0
Row 7: 0 1 1 1 1 1 1 0
Row 8: 0 0 0 0 0 0 0 0

Figure 7.2
The cell-by-cell data structure records each cell value by row and column.

Run-length Encoding

Records the cells by row and by group. Each group includes a cell value and the number of cells with that value. If all cells in a row contain the same value, only one group is recorded, hence save computer memory.

The run length encoding method records the cell values in runs. Row 1, for example, has two adjacent cells in columns 5 and 6 that are gray or have the value of 1. Row 1 is therefore encoded with one run, beginning in column 5 and ending in column 6. The same method is used to record other rows.

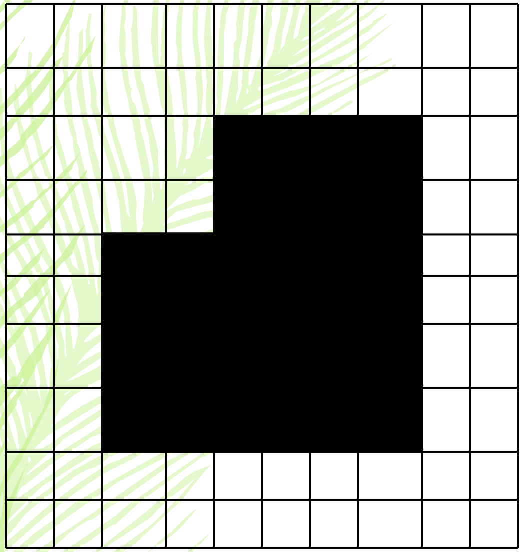


Row 1: 5 6
Row 2: 4 6
Row 3: 3 7
Row 4: 3 7
Row 5: 3 7
Row 6: 2 7
Row 7: 2 7

Figure 7.3
The run length encoding method records the cell values in runs. Row 1, for example, has two adjacent cells in columns 5 and 6 that are gray or have the value of 1. Row 1 is therefore encoded with one run, beginning in column 5 and ending in column 6. The same method is used to record other rows.

Raster Data Storage – Run Length Encoding

This approach takes advantage of **patterns** in the data, taking advantage of the **repetition** of values in a row:



0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	0	0
0	0	0	0	1	1	1	1	0	0
0	0	1	1	1	1	1	1	0	0
0	0	1	1	1	1	1	1	0	0
0	0	1	1	1	1	1	1	0	0
0	0	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

header → 10,10,1

row by row

0, 10
0, 10
0, 4, 1, 4, 0,2
0, 4, 1, 4, 0,2
0, 2, 1, 6, 0,2
0, 2, 1, 6, 0,2
0, 2, 1, 6, 0,2
0, 2, 1, 6, 0,2
0, 10
0, 10

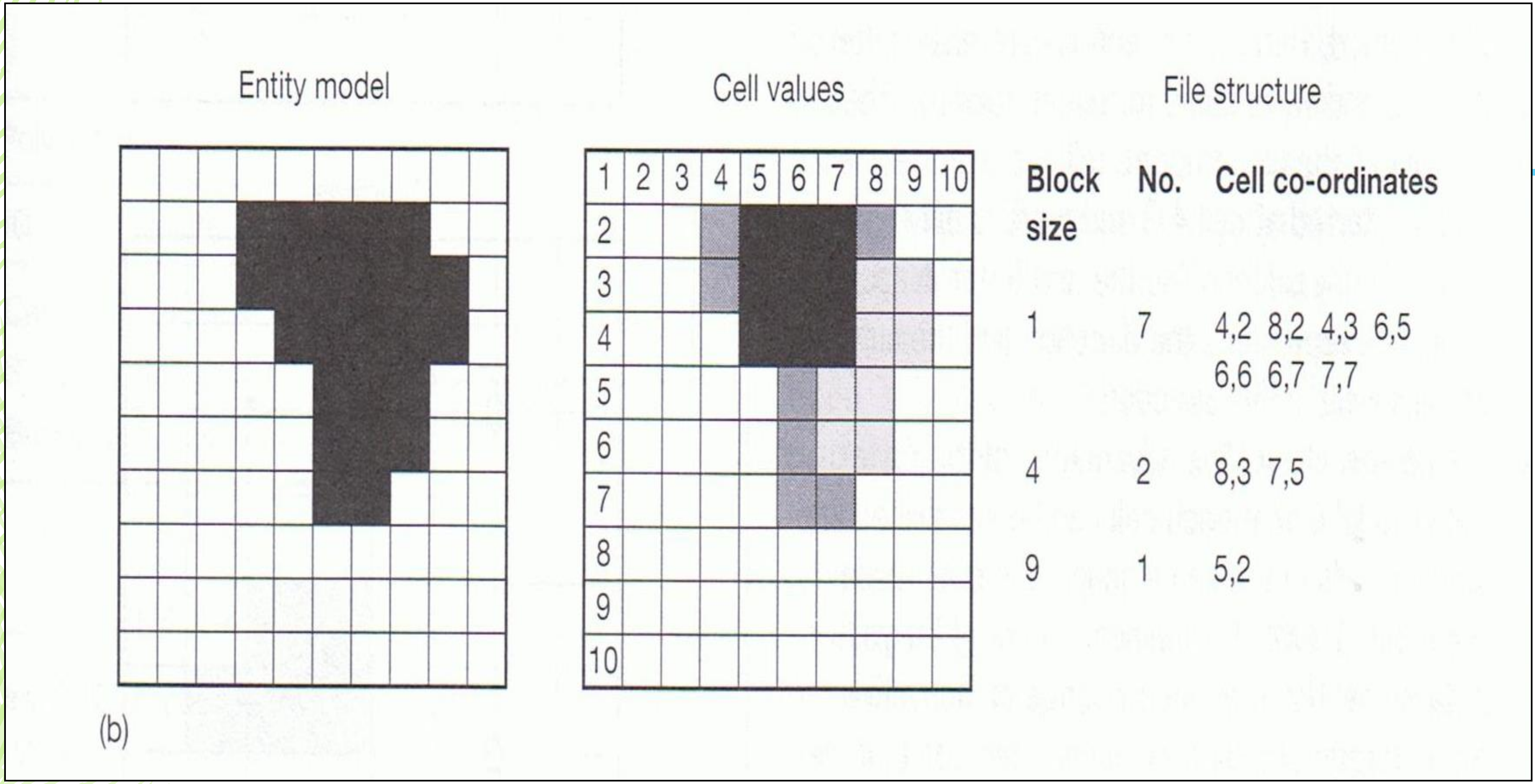
45 values

There is a tendency towards **spatial autocorrelation**; for **nearby cells** to have **similar values** - values often occur in runs across several cells

Block Code Method

Uses square blocks to represent the region. Each square block is coded only with the location of a cell (lower left of the block), and the side length of the block.

Raster Data Compression Models: Block Encoding



Run-length encoding in 2-D: Uses a series of square blocks to encode data

From *An Introduction to Geographic Information Systems*, Heywood et al. (2002)

Chain Code Method

Represent the boundary of a region by using a series of cardinal directions and cells. – Ex: N1 means moving north by 1 cell, S4 means moving south by 4 cells.

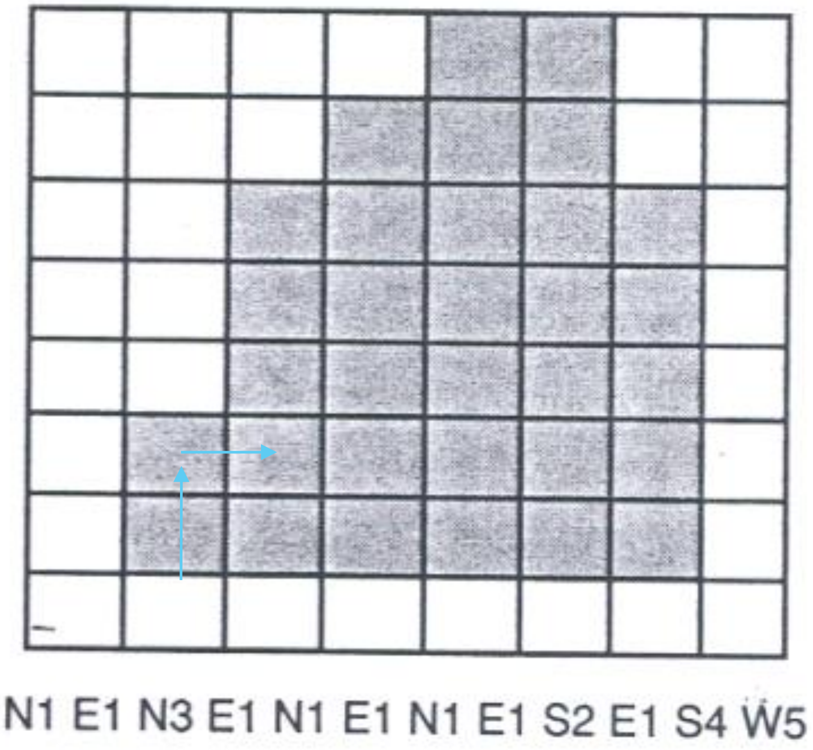
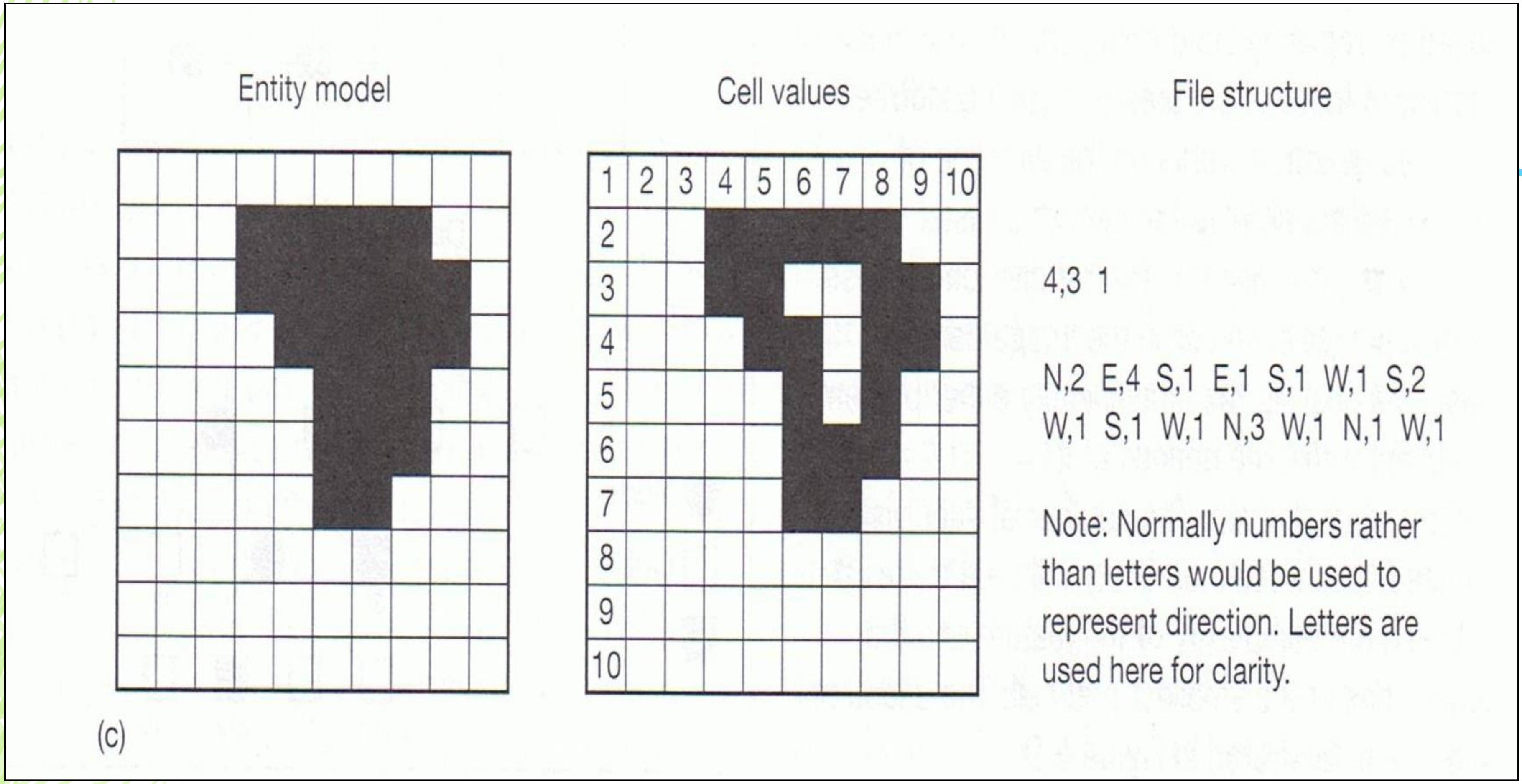


Figure 7.4
Starting at the lower left cell of the region, the chain codes method records the region's boundary by using the principal direction and the number of cells. In this example, the recording follows a clockwise direction.

Raster Data Compression Models:
Raster Chain Codes



Reduces data by defining the boundary of entity

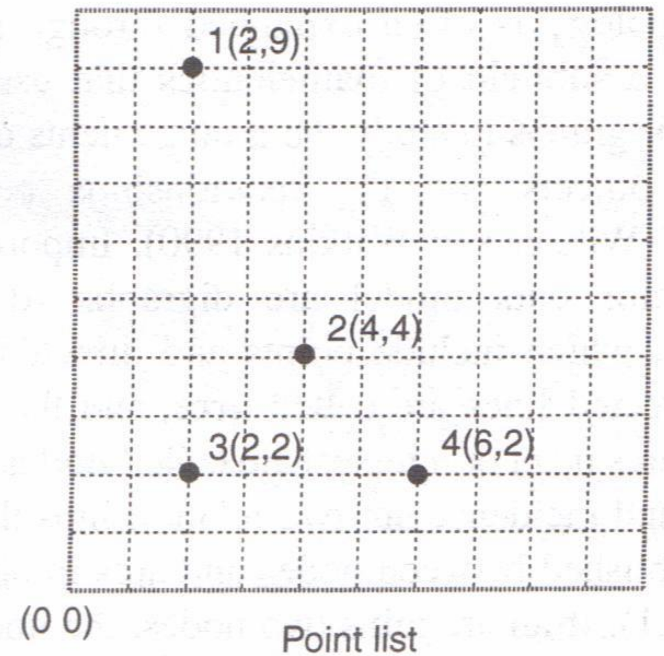
From *An Introduction to Geographic Information Systems*, Heywood et al. (2002)

ESRI's Coverage Model

- **Coverage** is a topology-based vector data format.
- The coverage model supports three basic topological relationships:
  - **Connectivity:** Arcs connect to each other at nodes.
  - **Area definition (Containment):** An area is defined by a series of connected arcs.
  - **Contiguity (Adjacency):** Arcs have directions and left and right polygons.

Coverage Data Structure

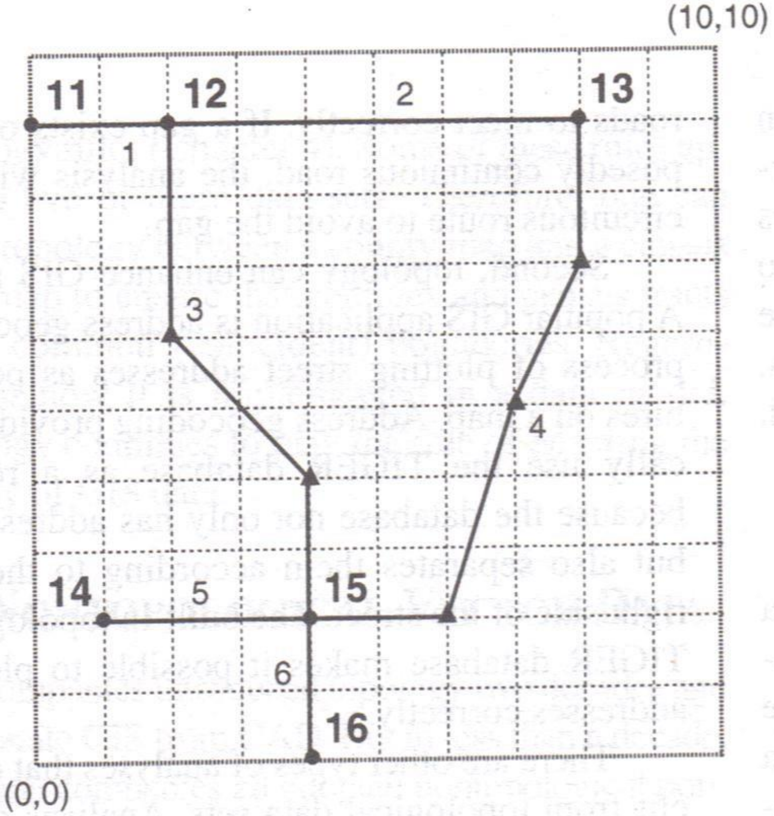
- The coverage model incorporates the topological relationships into the structure of feature data.
- The data structure of a point coverage contains feature identification numbers (IDs) and pairs of *x*- and y-coordinates.



Point list

ID	*x,y*
1	2,9
2	4,4
3	2,2
4	6,2

The data structure of a line coverage



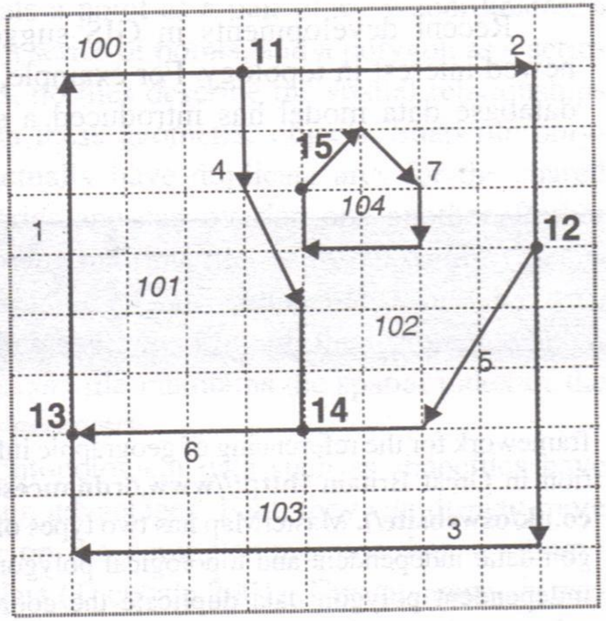
Arc-node list

Arc#	F-node	T-node
1	11	12
2	12	13
3	12	15
4	13	15
5	15	14
6	15	16

Arc-coordinate list

Arc#	*x,y* Coordinates
1	(0,9) (2,9)
2	(2,9) (8,9)
3	(2,9) (2,6) (4,4) (4,2)
4	(8,9) (8,7) (7,5) (6,2) (4,2)
5	(4,2) (1,2)
6	(4,2) (4,0)

The data structure of polygon coverage



Left/right list

Arc#	L-poly	R-poly
1	100	101
2	100	102
3	100	103
4	102	101
5	103	102
6	103	101
7	102	104

Polygon/arc list

Polygon #	Arc#
101	1,4,6
102	4,2,5,0,7
103	6,5,3
104	7

Arc-coordinate list

Arc#	x,y Coordinates
1	(1,3) (1,9) (4,9)
2	(4,9) (9,9) (9,6)
3	(9,6) (9,1) (1,1) (1,3)
4	(4,9) (4,7) (5,5) (5,3)
5	(9,6) (7,3) (5,3)
6	(5,3) (1,3)
7	(5,7) (6,8) (7,7) (7,6) (5,6) (5,7)

Raster Model	Vector Model
Advantages: • Simple data structure • Easy and efficient overlaying of grids • Compatibility with satellite imagery • High spatial variability which helps in efficient representation • Simple structure for own programming.	**Advantages:** • Compact data structure that is efficient for network analysis and projection transformation • Accurate map output.
Disadvantages: • Inefficient use of computer storage • Errors in perimeter and shape • Difficult network analysis • Inefficient projection transformations • Loss of information when using large cells which tend to give out less accurate maps.	**Disadvantages:** • Complex data structure • Difficult overlay operations • High spatial variability is inefficiently represented • Not compatible with satellite imagery.

Software Modules

GIS Software Producers and their main products

AUTODESK Inc: Autodesk Map
Baylor University: GRASS (Geographic Resources Analysis System),Open Source
Bentley Systems: Microstation
Caliper Cooperation: TransCad, Maptitude
Clark Labs: IDRISI
ESRI (Environmental System Research Institute): ArcGIS, Arc View 3x
Intergraph Corperation: MGE, GeoMedia
International Institute for Aerospace Survey and Earth Sciences: ILWIS
Keigan Systems: Mfworks, Keigan Grid
Manifold.net: Manifold Systems
Map Info Corporation: MapInfo
PCI Geomatics: Geomatica
QuantumGIS (Open Source)