SALALE UNIVERSITY

COLLEGE OF SOCIAL SCIENCES AND HUMANITIES

DEPARTMENT OF GEOGRAPHY & ENVIRONMENTAL STUDIES

MODULE FOR THE COURSE: Introduction GIS (GeES 3014)

PREPARED BY: Betelhem. M (MSc in GIS & RS)

Academic Year: 2015 E.C

1. Basics of Geographic Information Systems (GIS)

As far as Geographic Information System is concerned different definitions have been given for the past years as they were needed. It is not surprising that GIS can be defined in many different ways. Which definition to choose depends on what you seek. Common to all definition is that spatial data, which is unique because it is liked to a geographical referenced data. Among other definitions, GIS is a powerful set of tools for collecting, storing, retrieving, transforming, and displaying spatial data from the real world for a particular set of purposes. Storage and analysis of spatial data is traditionally done in Geographical Information Systems (GIS).

In general, a GIS provides facilities for data capture, data management, data manipulation and analysis, and the presentation of results in both graphic and report form, with a particular emphasis upon preserving and utilizing inherent characteristics of spatial data. The ability to incorporate spatial data, manage it, analyse it, and answer spatial questions is the distinctive characteristic of geographic information systems. A geographic information system, commonly referred to as a GIS, is an integrated set of hardware and software tools used for the manipulation and management of digital spatial (geographic) and related attribute data.

1.1 Components of GIS

An operational GIS also has a series of components that combine to make the system work. These components are critical to a successful GIS. A working GIS integrates five key components:

Hardware

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

Software

GIS software provides the functions and tools needed to store, analyse, and display geographic information. A review of the key GIS software subsystems is provided above.

Vector-based GIS

- ArcGIS (ESRI)
- ArcView
- MapInfo

Raster-based GIS

- Erdas Imagine (Leica)
- ENVI (RSI)
- ILWIS (ITC)
- IDRISI (Clark Univ.)

Data

Perhaps the most important component of a GIS is the data. Geographic data and related tabular data can be collected in-house, compiled to custom specifications and requirements, or occasionally purchased from a commercial data provider. A GIS can integrate spatial data with other existing data resources, often stored in a corporate DBMS. The integration of spatial data (often proprietary to the GIS software), and tabular data stored in a DBMS is a key functionality afforded by GIS.

People

GIS technology is of limited value without the people who manage the system and develop plans for applying it to real world problems. GIS users range from technical specialists who design and maintain the system to those who use it to help them perform their everyday work. The identification of GIS specialists versus end users is often critical to the proper implementation of GIS technology.

Methods

A successful GIS operates according to a well-designed implementation plan and business rules, which are the models and operating practices unique to each organization. As in all organizations dealing with sophisticated technology, new tools can only be used effectively if they are properly integrated into the entire business strategy and operation. To do this properly requires not only the necessary investments in hardware and software, but also in the retraining and/or hiring of personnel to utilize the new technology in the proper organizational context. Failure to implement your GIS without regard for a proper organizational commitment will result in an unsuccessful system! Many of the issues concerned with organizational commitment are described in Implementation Issues and Strategies. It is simply not sufficient for an organization to purchase a computer with some GIS software, hire some enthusiastic individuals and expect instant success.

1.2 Sub Systems of GIS

A GIS has four main functional subsystems. These are:

- ✓ a data input subsystem;
- ✓ a data storage and retrieval subsystem;
- ✓ a data manipulation and analysis subsystem; and
- ✓ a data output and display subsystem.

Data Input

A data input subsystem allows the user to capture, collect, and transform spatial and thematic data into digital form. The data inputs are usually derived from a combination of hard copy maps, aerial photographs, remotely sensed images, reports, survey documents, etc.

Data Storage and Retrieval

The data storage and retrieval subsystem organize the data, spatial and attribute, in a form which permits it to be quickly retrieved by the user for analysis, and permits rapid and accurate updates to be made to the database. This component usually involves use of a database management system (DBMS) for maintaining attribute data. Spatial data is usually encoded and maintained in a proprietary file format.

Data Manipulation and Analysis

The data manipulation and analysis subsystem allow the user to define and execute spatial and attribute procedures to generate derived information. This subsystem is commonly thought of as the heart of a GIS, and usually distinguishes it from other database information systems and computer-aided drafting (CAD) systems.

Data Output

The data output subsystem allows the user to generate graphic displays, normally maps, and tabular reports representing derived information products. The critical function for a GIS is, by design, the analysis of spatial data. It is important to understand that the GIS is not a new invention. In fact, geographic information processing has a rich history in a variety of disciplines. In particular, natural resource specialists and environmental scientists have been actively processing geographic data and promoting their techniques since the 1960's. Today's generic, geographic information system, is distinguished from the geo-processing of the past by the use of computer automation to integrate geographic data processing tools in a friendly and comprehensive environment. The advent of sophisticated computer techniques has proliferated the multi-disciplinary application of geo-processing methodologies, and provided data integration capabilities that were logistically impossible before.

Data Input **Data Stoage, Retrieval, Manipulation and Analysis** **Data Output**

1.3 Applications of GIS

GIS are needed in part because human population and technology have reached levels such that many resources, including air and land, are placing substantial limits on human action. Thus, GIS help us to identify and to address environmental problems by providing crucial information on where problems occur and who are affected by them. Location and extent of adverse environmental impacts may help us to devise practical plans for monitoring, managing, and mitigating environmental damage.

GIS can be used for different purposes, some applications of it includes:

- Agriculture
- Archaeology
- Geology
- Geography
- Health care
- Business
- Public utilities
- Environmental Science
- Engineering
- Park management
- Military Science
- Natural Resource Management
- Emergency

1.4 Capabilities of GIS

GIS can offer many capabilities. These include: Storage (Geodatabase), Display (Visualization) Analytical, Modelling and Geoprocessing capabilities.

Data Capture: The input of data into a GIS can be achieved through many different methods of gathering. For example, aerial photography, scanning, digitizing, GPS or global positioning system is just a few of the ways a GIS user could obtain data.

Storage (Geodatabase): the data is stored in proprietary format develop by Esri. A file geodatabase can hold more data than a personal geodatabase: up to terabytes. geodatabase store data on central server in a relational database management system. Data Storage Some data is stored such as a map in a drawer, while others, such as digital data, can be as a hardcopy, stored on CD or on your hard drive.

Data Manipulation: The digital geographical data can be edited, this allows for many attributes to be added, edited, or deleted to the specification of the project.

Query and Analysis: GIS was used widely in decision making process for the new commission districts. We use population data to help establish an equal representation of population to area for each district.

Display (Visualization): This represents the ability to display your data, your maps, and information.

2. Parameters of Spatial Reference System

2.1. Coordinate Systems

There are two types of coordinate system: Geographical Coordinate System and Cartesian coordinate system. Geographical coordinate system: Geographers use geographic (earth) coordinate system to locate a point on the surface of the earth. The earth's coordinate system, which is also called spherical coordinate system, is the primary location reference system for the earth. The geographical coordinate system measure’s location from only two values, despite the fact that the locations are described for a three-dimensional surface. The two measures used in the geographic coordinate system are called latitude and longitude. Our planet earth rotates about an imaginary axis called the axis of the earth. If extended one end points to a fixed star, the North Star. The place on earth where this axis of rotation ends referred to as geographic north (North

Pole). The opposite or antipodal point is called geographic south or South Pole. To specify location on earth in the geographical coordinate system we use latitude and longitude.

Each set of numbers corresponds to exactly one position in space, this system of assigned numbers is called a **coordinate system.**

2.1.1. Geographic and Projected Coordinate System

1. The geographical coordinate system: A geographic coordinate system is a reference system for identifying locations on the curved surface of the earth. Locations on the earth's surface are measured in angular units from the center of the earth relative to two planes: the plane defined by the equator and the plane defined by the prime meridian

It is a coordinate system which uses a three-dimensional spherical surface (ellipsoid) to define locations on the earth. A common choice of coordinates is latitude and longitude. For example, Leuven, Belgium is located on 50°52'47" North and 4°42'01" East in the WGS84 coordinate system.

A geographic coordinate system is constituted by a datum (DATUM), a prime meridian (PRIMEM), and unit (UNIT). The datum is constituted by a ellipsoid model (SPHEROID) and a anchor point.

measure's location from only two values, despite the fact that the locations are described for a three-dimensional surface. A geographic coordinate system (GCS) is used to define locations on a model of the surface of the earth. The GCS uses a network of imaginary lines (longitude and latitude) to define locations. This network is called a graticule.

Well, it turns out the earth isn't a perfect sphere. It's a lumpy, bumpy, and uneven rounded surface. There are high mountains and deep ocean trenches. Because the planet spins, the poles are a bit closer to the center of the earth than the equator is. But in order to draw a graticule, you need a model of the earth that is at least a regular spheroid, if not a perfect sphere.

There are many different models of the earth's surface, and therefore many different GCS! *World Geodetic System 1984 (WGS 1984)* is designed as a one-size-fits-all GCS, good for mapping global data. *Australian Geodetic Datum 1984* is designed to fit the earth snugly around Australia, giving you good precision for this continent but poor accuracy anywhere else.

The GCS is what ties your coordinate values to real locations on the earth. The coordinates 134.577°E, 24.006°S only tell you where a location is within a geographic coordinate system. You still need to know which GCS it is in before you know where it is on Earth.

The two measures used in the geographic coordinate system are called latitude and longitude.

- The values for the points can have the following units of measurement: decimal degrees (0), decimal minutes (‘), and decimal seconds (").
- The *prime meridian* is the line of longitude that defines the origin (zero degrees) for longitude coordinates.

One of the most commonly used prime meridian locations is the line that passes through Greenwich, England.

Latitude and longitude

One method for describing the position of a geographic location on the earth's surface is using spherical measures of latitude and longitude. They are measures of the angles (in degrees) from the center of the earth to a point on the earth's surface. This reference system is often referred to as a geographic coordinate system.

Latitude angles are measured in a north-south direction. The ***equator*** *is at an angle of 0*. Often, the northern hemisphere has positive measures of latitude and the southern hemisphere has negative measures of latitude.

These latitude lines call ***Parallels***

Longitude measures angles in an east-west direction. Longitude measures are traditionally *based on the* ***Prime Meridian***, which is an imaginary line running from the *North Pole through Greenwich, England to the South Pole*. This angle is Longitude 0. West of the Prime Meridian is often recorded as negative Longitude and east is recorded as positive

These vertical lines call ***meridians***

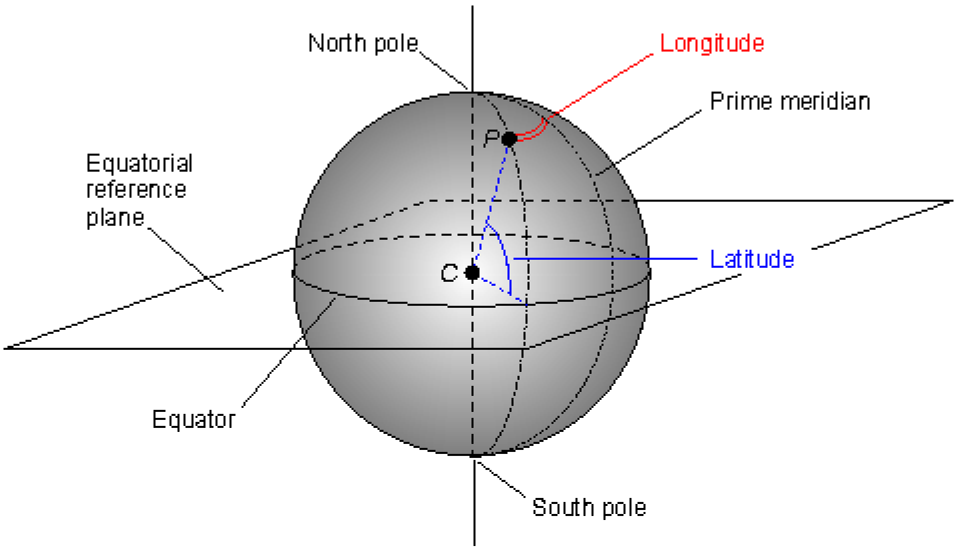
One method for describing the position of a geographic location on the earth's surface is using spherical measures of latitude and longitude. They are measures of the angles (in degrees) from the center of the earth to a point on the earth's surface.

This reference system is often referred to as a **geographic coordinate system.**

Latitude

angles are measured in a north-south direction. The ***equator*** *is at an angle of 0*. Often, the northern hemisphere has positive measures of latitude and the southern hemisphere has negative measures of latitude.

These latitude lines call ***Parallels***

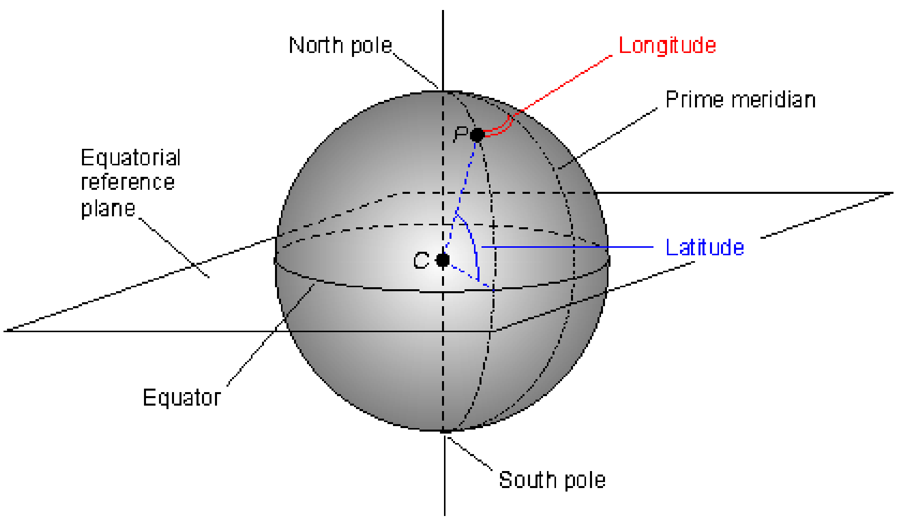


Longitude

measures angles in an east-west direction. Longitude measures are traditionally *based on the **Prime Meridian***, which is an imaginary line running from the *North Pole through Greenwich, England to the South Pole*.

This angle is Longitude 0. West of the Prime Meridian is often recorded as negative Longitude and east is recorded as positive

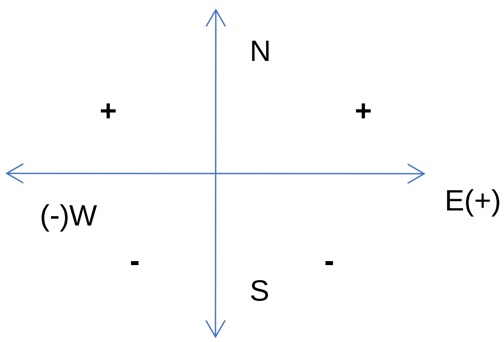
These vertical lines call ***meridians***

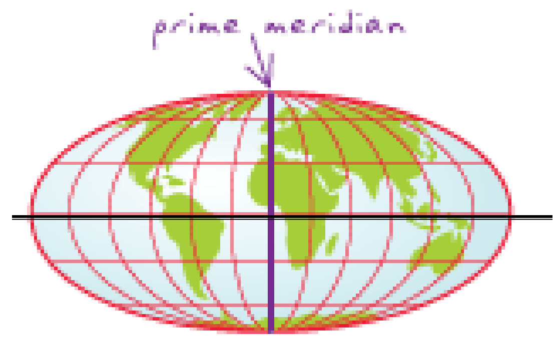


Latitude and longitude values are traditionally measured either in decimal degrees or in degrees, minutes, and seconds (DMS). **Latitude** values are measured relative to the **Equator** and range from **-90° at the South Pole** to **+90° at the North Pole**.

Longitude values are measured relative to the **Prime Meridian**. They range from -180° when traveling **West** to 180° when traveling **East**.

If the prime meridian is at Greenwich, then **Australia,** which is south of the equator and east of Greenwich, has positive longitude values and negative latitude values. (It may be helpful to equate **longitude** values with **X** and **latitude** values with **Y**. Data defined on a geographic coordinate system is displayed as if a degree is a linear unit of measure)





2 Projected Coordinate Systems

The surface of the earth is curved but maps are flat. A projected coordinate system (PCS) is a reference system for identifying locations and measuring features on a flat (map) surface. It consists of lines that intersect at right angles, forming a grid. Projected coordinate systems (which are based on Cartesian coordinates) have an origin, an *x* axis, a *y* axis, and a linear unit of measure. Going from a GCS to a PCS requires mathematical transformations. The myriad of projection types can be aggregated into three groups: **planar**, **cylindrical** and **conical**.

In a **Projected Coordinate System** (**PCS**) you project the geographic coordinate that you have measured to, for example, a cylinder that you can easily roll out on a two-dimensional surface (the map). There exist many different projections and we'll not go in further detail about that here.

Typically, every country, state, or region has its optimal projected coordinate system, which minimizes distortions for particular applications like mapping.

A projected coordinate system is constituted by a geographic coordinate system from which it is projected (GEOGCS) and other projection parameters like the measurement unit (like meter or US Survey Foot), the projection technique, and its projection parameters.

Spatial projection refers to the mathematical calculations performed to flatten the 3D data onto a 2D plane (your computer screen or a paper map). Projecting data from a round surface onto a flat surface, results in visual modifications to the data when plotted on a map. Some areas are stretched and some are compressed. You can see this distortion when you look at a map of the entire globe.

The mathematical calculations used in spatial projections are designed to optimize the relative size and shape of a particular region on the globe.

2.1.2. Universal Transverse Mercator System

Universal Transverse Mercator is a global coordinate system that is defined in meters rather than degrees-minutes-seconds. UTM is a very precise method of defining geographic locations; therefore, it is commonly used in GPS and GIS mapping. When using the UTM coordinate system, a location can be identified within a meter.

The UTM (Universal Transverse Mercator's)

It is a matrix of lat. and long. which give the precise position across the world It is grid system has been widely adopted for topographic maps satellite imagery, natural resources data bases and other applications that require precise positioning. It is a metric system (meter is the basic unit of measurement).

Properties UTM

- □ The most western edge of UTM is zone 1 and the most eastern edge is zone has 6^0.
- □ Each zone has 6-degree longitudinal extent. That means zone 1 extends from 180^0 W to 174o W. The latitudinal interval is 8 latitudes.
- □ The latitudinal extent is form 84^0 N to 80^0 S.

- The rows of quadrilaterals are assigned letters C to X consecutively (with I and O omitted) beginning at 800S latitude.
- □ Row X which extends from 720N to 840N to cover all land areas in the northern hemisphere is having a latitudinal extent of 120.

  Example If Ethiopia existed on 3-15degree North and 33-48 degree East . what is UTM zone of Ethiopia? Answer 37N

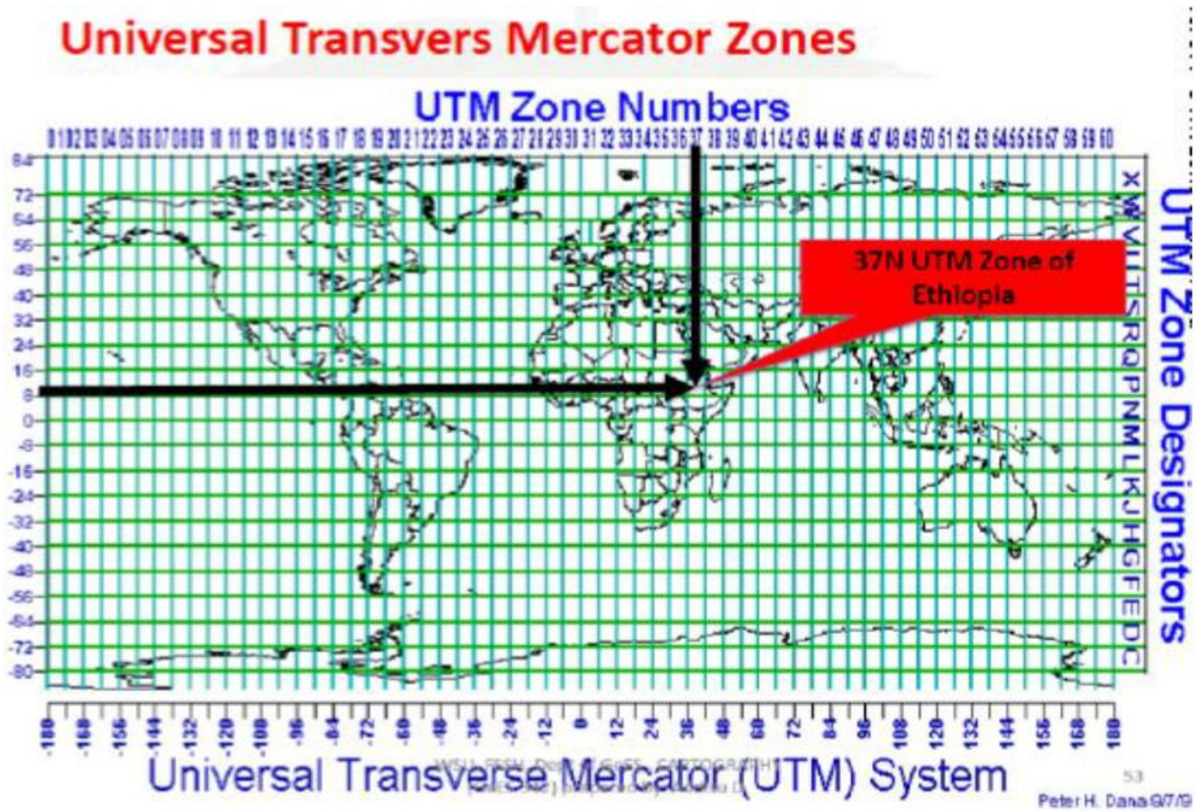


Figure:2.1. UTM Zones of the World

2.2. Datum

A datum provides a frame of reference for measuring locations on the surface of the earth. It defines the origin and orientation of latitude and longitude lines. The solution is to align the geoid with the ellipsoid (or sphere) representation of the earth and to map the earth's surface features onto this ellipsoid/sphere. The alignment can be local where the ellipsoid surface is closely fit to the geoid at a particular location on the earth's surface (such as the state of Kansas) or **geocentric** where the ellipsoid is aligned with the center of the earth. How one chooses to align the ellipsoid to the geoid defines a **datum**. Of the elements that make up a **coordinate reference system** (CRS), none are more fundamental than **datums**. In fact, the term "datum" is often used interchangeably with "**geographic coordinate system"** (or a type of CRS used for specifying the location of a point on the earth).

But a datum is only one component of a geographic coordinate system—albeit an indispensable one. Because understanding the function of datums is crucial to getting the bigger picture of CRSs, we created this quick primer.

While a spheroid approximates the shape of the earth, a datum defines the position of the spheroid relative to the center of the earth. A datum provides a frame of reference for measuring locations on the surface of the earth. It defines the origin and orientation of latitude and longitude lines. Whenever you change the datum, or more correctly, the geographic coordinate system, the coordinate values of your data will change.

Here are the coordinates in DMS of a control point in Redlands, California, on the North American Datum of 1983 (NAD 1983 or NAD83)………..

-1171257.75961

340143.77884

Here's the same point on the North American Datum of 1927 (NAD 1927 or NAD27).

-1171254.61539

340143.72995

The longitude value differs by approximately three seconds, while the latitude value differs by about0.05seconds. NAD 1983 and the World Geodetic System of 1984 (WGS 1984) are identical for Most applications. Here are the coordinates for the same control point based upon WGS 1984.

-1171257.75961

34 01 43.778837

2.2.1. Vertical and Horizontal Datum

ellipsoids and geoids define the size and shape of the earth, to varying degrees of accuracy. But that only gets us halfway to our end goal. To be useful, we also need to know which specific model to use and where to place them in space.

That’s where datums come in. **Horizontal datums** take an ellipsoid and assign its center a point of origin relative to the center of the earth. Once the ellipsoid is “pinned” to the earth, we can use the horizontal datum to start assigning angular units of longitude and latitude to different points on the earth’s surface. When comparing survey datasets, they need to be referenced through the same datums. Otherwise, the same points in the real world will be assigned different points on your map, and your measurements will be off.

As we just established, the Earth’s surface is not uniform. Knowing this, when surveying, you ideally want to use the datum whose ellipsoidal model best fits your area. In other words, you want a **local datum.**

Over the years, more and more stations were benchmarked. The NAD 83 datum expanded on the NAD 27 to over 250,000 stations. It also used terrestrial and Doppler satellite data to correct some of the distortions in distance in the NAD 27. Over the past 37 years, the NAD 83 has been updated several times thanks to advances in GPS and geodesy.

With the advent of GPS, geodesists were able to develop a highly accurate worldwide ellipsoid model and corresponding datum called the WGS 84 that is still used for global mapping purposes today. The WGS 84 is an example of a **geocentric datum.** The center of the WGS 84 ellipsoid is generally considered to be about 2cm off the center of the earth's mass.

All GPS data is referenced through the WGS 84. And while horizontal measurements between the WGS 84 and NAD 83 are on average only one meter, that's a big enough difference in some disciplines—digging a trench on a **construction site** or creating a haul road on a **mining site**, for instance—to make a **datum transformation** from GPS collected coordinates to your local datum necessary.

While horizontal datums pinpoint a location on the earth's surface, a **vertical datum** gets you elevation data. Vertical datums use the surface of a geoid model to establish a zero-point of elevation. When using a vertical datum, you can describe a point as being x-feet above or below the "Mean Sea Level (MSL)."

The MSL actually represents an "equipotential surface," or a surface in which all points experience the same force of gravity. Another way to think about it is that it's how the ocean would cover the earth's topography if tides, wind, and some other factors that restrict its movements didn't exist. The only factor that affects the MSL's shape is the earth's gravitational field. Like horizontal datums, there are localized vertical datums. In the U.S., we use the North American Vertical Datum of 1988 (NAVD 88).

A vertical datum provides a base value from which elevations or heights are determined. On the Earth a vertical datum expresses a surface of zero elevations. There are two types of vertical datums: tidal and geodetic. A tidal datum is created by averaging water levels over a set time period at a given point (e.g., mean sea level MSL). A geodetic datum provides a network of heights known as benchmarks. The current vertical geodetic datum for the United States is the North American Vertical Datum 1988 (NAVD88). A review of vertical datums and their application to mapping follows.

Heights on the Earth's surface are defined by another Earth model called the geoid. Every point on the Earth's surface is being acted upon by a downward force called gravity. That downward force is also called potential energy and is not the same across the Earth's surface. One way to think about the geoid is to imagine what the Earth's surface would look like if the world's oceans were allowed to flow over the land, without any currents or tides, and conform to gravitational pull. As a result, the water's surface (or geoid) is not a regular, but an irregular smoothly changing surface that is influenced by rock density and the Earth's non-spherical shape.

What is height? There are three types of height: orthometric, geoid, and ellipsoidal. These heights are illustrated in Figure 2. Orthometric height (H) is the distance from the geoid to a point on the Earth's surface measured along a plumb line to gravitational pull. The reference surface for orthometric height is a level surface that closely coincides with global MSL (the geoid's surface). Orthometric height is what is commonly associated with elevation of a point on the Earth's surface.

Geoidal height (N) is the distance along a normal line from the reference ellipsoid to the geoid. In the coterminous United States, values of N (between the GRS80 reference ellipsoid and the geoid) are about -8 meters to about -53 meters.

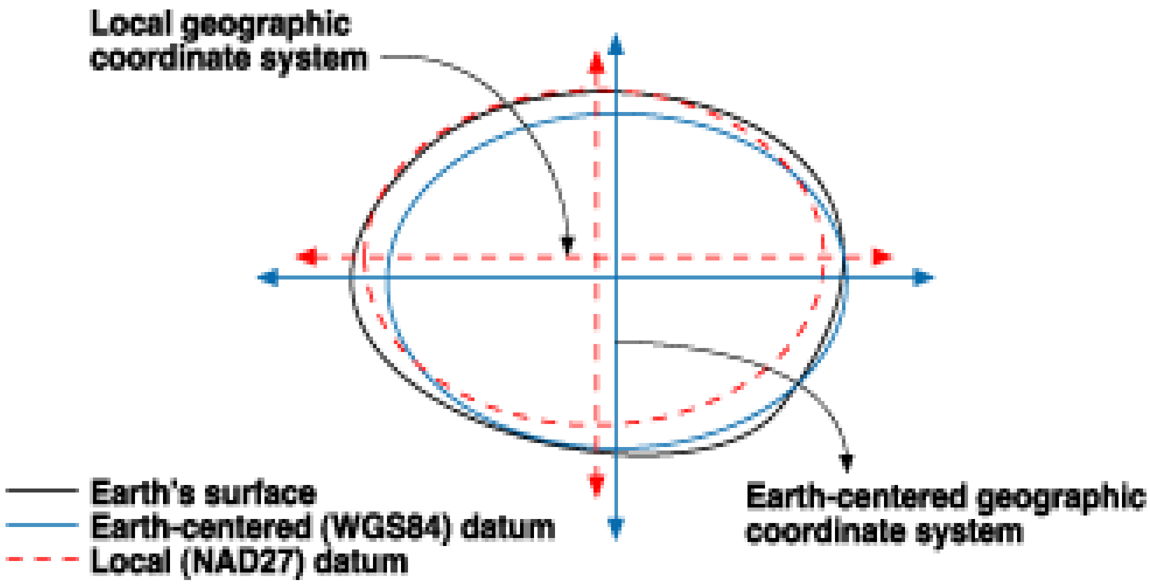
Ellipsoidal height (h) is the distance along a line normal from the reference ellipsoid to a point on the Earth's surface. Calculating H is simply the difference between h and N. H=hN

Height is an important consideration when dealing with data collected via Global Positioning System (GPS). GPS data is considered to be three-dimensional (X, Y, and Z). However, the Z component does not depend upon a direct relation to a gravitational surface. Instead, the X, Y, and Z coordinates are utilized when calculating geodetic latitude, longitude, and h. Then, h can be converted to H so as to be tied into a vertical datum.

2.2.2. Local and Global Datum

A local datum aligns its spheroid to closely fit the earth's surface in a particular area. A point on the surface of the spheroid is matched to a particular position on the surface of the earth. This point is known as the origin point of the datum. The coordinates of the origin point are fixed, and all other points are calculated from it**.**

The coordinate system origin of a local datum is not at the center of the earth. The center of the spheroid of a local datum is offset from the earth's center



North American Datum 1927

NAD 1927 uses the Clarke 1866 spheroid to represent the shape of the earth. The origin of this datum is a point on the earth referred to as Meades Ranch in Kansas. Many NAD 1927 control points were calculated from observations taken in the 1800s. These calculations were done manually and in sections over many years. Therefore, errors varied from station to station.

The North American Datum 1983

The North American Datum of 1983 is based on both earth and satellite observations, using the GRS 1980 spheroid. The origin for this datum is the earth's center of mass. This affects the surface location of all longitude–latitude values enough to cause locations of previous control points in North America to shift, sometimes as much as 500 feet

WGS 1984 (World Geodetic System)

The GRS 1980 spheroid is almost identical to the WGS 1984 spheroid. The WGS 1984 and NAD 1983 coordinate systems are both earth centered. Because both are so close, NAD 1983 is compatible with GPS data. The raw GPS data is actually reported in the WGS 1984 coordinate system.

3. Spatial Data and Geographic Information Systems

3.1 Spatial Data

Spatial data also known as geospatial (coordinate) data or geographic information. It is the data or information that identifies the geographic location of features and boundaries on Earth, such as natural or constructed features, parcels, roads, buildings and more. In other words, it describes the absolute and relative location of geographic or spatial features. Spatial data is usually stored as coordinates and topology, and is data that can be mapped. Spatial data use Cartesian coordinates systems. Two-dimensional Cartesian coordinate systems define x and y axes in a plane. The three-dimensional Cartesian system defines a z axis, orthogonal to both the x and y axes. An origin is defined with zero values at the intersection of the orthogonal axes. Spatial data is often accessed, manipulated or analysed through GIS.

3.1.1 Geographic Entities or Phenomenon

Geographic entities can also be called as geographic phenomena. GIS supports such study because it represents phenomena digitally in a computer. An entity or geographic feature occupies position in space about which data describing the attributes of the entity and its geographic location are recorded. It is a discrete generic class with basic connectedness and interdependence as a single data set, i.e., land use as a class has separate entities of residential, commercial, industrial, agricultural, etc. The class is a set of geographic entities derived from a common set of criteria, thus sharing spatial character and structure, e.g., ownership parcels, intersections, street segments, etc.

In other words, they are features exist in the real world that can be geo-referenced or located on the surface of the earth. True example can be one has to look outside the window. Generally, they are defined as something of interest that:

- Can be named/described – what is it?
- Can be geo-referenced – where is it?
- Can be assigned a time at which it is/was present

For instance: in multipurpose cadastral administration the object of study can be houses, parcels, business areas, schools, roads, buildings, etc. In water management, the object of study can be river basins, agro-ecological zones, measurement of evapo-transpiration, meteorological data, irrigation sites, etc. All these can be named/described, geo-referenced, and provided with time interval at which each exists.

3.1.2 Spatial Data Types

GIS technology utilizes two basic types of data. These are spatial data and non-spatial (attribute) data. Spatial data is the data or information that identifies the geographic location of features and boundaries on Earth, such as natural or constructed features, parcels, roads, buildings and more. Generally, spatial data can be primary and secondary bases on their source. On the other hand, 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.

Attribute Data

Attribute (descriptive) Data (what, how much, when???). Information which describes an entity represented by a graphic feature. Each piece of attribute information is related to a specific graphic element. specifies characteristics at that location, natural or human-created stored in a data base table. GIS systems traditionally maintain spatial and attribute data separately, then "join" them for display or analysis.

Attribute (A spatial) information is the label / name / categorisation / description associated with a spatial object The attributes can be as important as the spatial data Themselves May be more complex than the spatial data May be a simple text label (e.g. town name, river name, county population) Attributes usually stored in some form of database (i.e. DBF, ORACLE table ...) Relational or object oriented DBMS used to spatial objects to complex attribute tables.

Common types of attribute data

Nominal Attribute that distinguishes between locations without any ranking or potential for arithmetic. e.g., Land cover class, Soil types,

Ordinal Attribute that implies a ranking, but arithmetic calculations are nonsensical. E.g., Hierarchy of road type

Interval Attributes where differences make sense. E.g., Celsius temperature, contour interval

Ratio Attributes where it makes sense to divide one measurement by another. E.g., Kelvin temperature, income, distance

3.1.3 Geographic Fields, Geographic objects, and Boundaries

i) Geographic fields

Geographic fields are geographic phenomena at which every point in the study area a value can be determined. They manifest themselves essentially everywhere in the study area. The usual examples of geographic fields are: temp, pressure, elevation, etc. These fields are actually continuous in nature and are characterized by their fuzzy boundary nature. Types of geographic fields: Continuous Fields and Discrete Fields

- In a continuous field, the underlying function is assumed to be 'mathematically smooth', meaning that the field values along any path through the study area do not change abruptly, but only gradually. Continuity means that all changes in field values are gradual

  Example:

  - Air temperature
  - Barometric pressure
  - Soil salinity and
  - Elevation
- Discrete fields cut up the study space in mutually exclusive, bounded parts, with all locations in one part having the same field value.

  Discrete field still assigns a value to every location in the study area

Example:

- land classifications
- soil type
- land use type
- crop type or
- natural vegetation type

ii) Geographic objects

As opposite to the above discussed types of geographic phenomena, many other phenomena do not manifest themselves everywhere in the study area, but only in certain localities. These entities populate the study area and are usually distinguishable one from the other and can be characterized by their discrete boundary nature. The space between them is potentially empty. Examples include: building, road, parcel, river, etc. Their position in space can be determined by a combination of:

- ✓ Location (where is it)
- ✓ Shape (what form is it?)
- ✓ Size (how big is it?)

Orientation (in which direction is it facing?).

We usually do not study geographic objects in isolation, but whole collections of objects viewed as a unit. Examples include: building, road, parcel, river, tree, Earthquake faults etc.

3.2 Spatial Data Models

Vector and raster are two common methods of representing spatial or geographic data in digital environments (GIS).

3.2.1 Vector Data Model

Vector data models represents discrete elements such as points, lines, and polygons of real-world entities e.g., a road, river, city and towns, lakes or wetlands, farm land, etc. In the vector world, the point is the building block from which all spatial entities are constructed. The smallest spatial entity, the point, is represented by a single (x, y) coordinate pair.

Structure of Vector Data

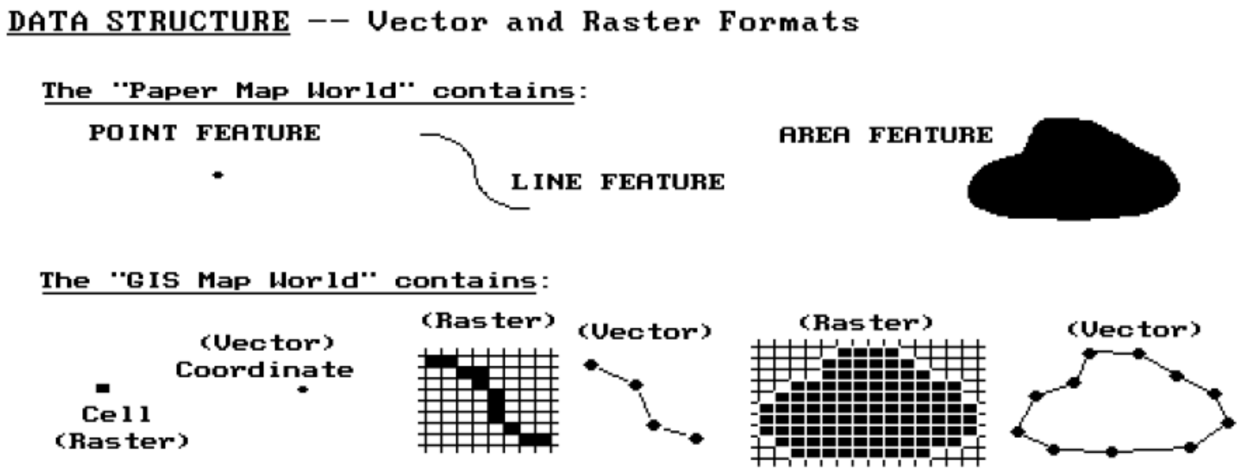
There is a relationship between vector data types. Each data type is often dependent upon one another. When representing geographic data in vector format, data is typically stored as:

- Points: points are zero dimensional objects, and represent geographic features such as wells, sample locations, trees, small buildings, mining holes, electric poles, traffic accident spots and etc. Points are zero dimensional objects which have locations and attribute information but are too small to be represented as areas
- Lines: lines represent one dimensional object, or linear features, such as roads, railways, pipe lines, cracks and stream centerlines. Lines are made up of a series of interconnected points. A line typically starts and end with a special point called a node, and the points that make up the rest of a line are called vertices. Lines are one dimensional object which have length but no area.

- Polygons/Area: polygons represent two dimensional objects such as the boundaries of a region, farm plot or the outline of a building or lake. Polygons are made up of a series of connected lines where the starting point of a polygon is the same as the ending point.

3.2.2 Raster Data Model

Raster data models define the world as a regular set of cells in a grid pattern. Raster data models are the natural means to represent "continuous" spatial features or phenomena. Elevation, precipitation, slope, and pollutant concentration are examples of continuous spatial variables.



3.3 Advantages and Disadvantages of Vector and Raster Data

There are several advantages and disadvantages for using either the vector or raster data model to store spatial data. These are summarized below.

3.3.1 Vector Data

Advantages of Vector Data:

- Data can be represented at its original resolution and form without generalization.
- Graphic output is usually more aesthetically pleasing (traditional cartographic representation)
- Since most data, e.g. hard copy maps, is in vector form no data conversion is required.
- Accurate geographic location of data is maintained.
- Allows for efficient encoding of topology and as a result more efficient operations that require topological information, e.g. proximity, network analysis.

Disadvantages of Vector Data:

- The location of each vertex needs to be stored explicitly.

- For effective analysis, vector data must be converted into a topological structure. This is often processing intensive and usually requires extensive data cleaning. As well, topology is static, and any updating or editing of the vector data requires re-building of the topology.
- Algorithms for manipulative and analysis functions are complex and may be processing intensive. Often, this inherently limits the functionality for large data sets, e.g. a large number of features.
- Continuous data such as elevation is not effectively represented in vector form. Usually substantial data generalization or interpolation is required for these data layers.
- Spatial analysis and filtering within polygons are impossible

3.3.2 Raster Data

Advantages of Raster Data:

- The geographic location of each cell is implied by its position in the cell matrix. Accordingly, other than an origin points no geographic coordinates are stored.
- Due to the nature of the data storage technique data analysis is usually easy to program and quick to perform.
- The inherent nature of raster maps, e.g. one attribute map, is ideally suited for mathematical modeling and quantitative analysis.
- Discrete data e.g. forestry stands is accommodated equally well as continuous data e.g. elevation data and facilitates the integrating of the two data types.
- Grid-cell systems are very compatible with raster-based output devices e.g. electrostatic plotters, graphic terminals.

Disadvantages of Raster Data:

- The cell size determines the resolution at which the data is represented.
- It is especially difficult to adequately represent linear features depending on the cell resolution. Accordingly, network linkages are difficult to establish.
- Processing of associated attribute data may be cumbersome if large amounts of data exists. Raster maps inherently reflect only one attribute or characteristic for an area.

- Since most input data is in vector form data must undergo vector-to-raster conversion. Besides increased processing requirements this may introduce data integrity concerns due to generalization and choice of inappropriate cell size.
- Most output maps from grid-cell systems do not conform to high-quality cartographic needs.

3.4 Attribute Data Models

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. For example, 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. Depending on the specific content of the data, image data may be considered either spatial, e.g. photographs, animation, movies, etc., or attribute, e.g. sound, descriptions, narration's, etc.

Attribute data are used to record the non-spatial characteristics of an entity. Attributes are also called items or variables. Attributes may be envisioned as a list of characteristics that help describe and define the features we wish to represent in a GIS. Color, depth, weight, owner, components vegetation type, or land use are examples of variables that may be used as attributes. Attributes have values, e.g. colour may be blue, black or brown, weight from 0.0 to 500, or land use may be urban, agriculture, or undeveloped.

Attributes are often presented in tables, with attributes arranged in rows and columns. Each row corresponds to an individual spatial object and each column corresponds to an attribute. Attributes of different types may be grouped together to describe the non-spatial properties of each object in the database. These attribute data may take many forms but all attribute data can be categorized as nominal, ordinal, or interval/ratio attributes.

3.5 Spatial Data Base Management

A database is a large collection of interrelated data stored within a computer environment. In such environments, the data is *persistent,* which means that it survives until unexpected software or hardware faces problems (except severe cases of disk crashes).

A DBMS is a collection of software that manages the database structure and controls access to data stored in a database. Small, simple databases that are used by a small number of people can be stored on computer disk in standard files. However, larger, more complex databases with many tens, hundreds, or thousands of users require specialist database management system (DBMS) software to ensure database integrity and longevity. Generally speaking, a DBMS facilitates the process of:

- Defining a database
- Constructing the database
- Manipulating the database
- Querying the database
- Updating the database

3.5.1 Spatial data base components

- Entity
- Object class
- Attributes Table

The lowest level of user interaction with a geographic database is usually the object class (also called a layer or feature class). A table is a two-dimensional array of rows and columns. Each object class is stored as a single database table in a database management system (DBMS). Table rows contain objects (instances of object classes, e.g., data for a single pipe) and the columns contain object properties or attributes.

3.5.2 Characteristics of spatial data base

A spatial database is characterized by the following:

- Full detailed for the intended applications
- Positional accuracy
- compatibility with other information that may be overlain with it
- Internal accuracy, portraying the nature of phenomena without error
- Readily updated on a regular schedule
- Accessible to whoever needs it

3.6 Topology

The topologic model is often confusing to initial users of GIS. A GIS topology is a set of rules and behaviours that model how points, lines, and polygons share geometry. Topology has long been a key GIS requirement for data management and integrity. In general, a topological data model represents spatial objects (point, line, and area features) using an underlying graph of topological primitives. These primitives, together with their relationships to one another and to the features whose boundaries they represent, are defined by representing the feature geometries in a planar graph of topological elements. Such datasets are said to be topologically integrated. In other words, topology describes the spatial relationships between adjacent or connecting features, and uses x, y coordinates to identify the location of a particular point, line, or polygon.

Topology is a mathematical approach that allows us to structure data based on the principles of feature adjacency and feature connectivity. It is in fact the mathematical method used to define spatial relationships. Without a topologic data structure, a vector-based GIS data manipulation and analysis functions would not be practical or feasible. Using such data structures enforces planar relationships, and allows GIS specialists to discover relationships between data layers, to reduce artifacts from digitization, and to reduce the file size required for storing the topological data. Some rule should be considered when applying topological relationships. For instance: Parcels cannot overlap, Valves cover pipes, and contours never cross.

Function of Topology in GIS

1. Basically topology is used to organize spatial relationships between discrete features represented by vector data.

2. Topology can be used to create datasets with better quality control and greater data integrity.

3. Topology rules can be created so that edits made to a dataset can be 'validated' and show errors in that dataset.

In essence, the main functions of topology are to define:

a. feature-to-feature locality or, simply, where a feature is in relation to another feature,(can be used for spatial queries)

b. what is shared between different features, and

c. how features are grouped or connected with each other

3.7 Data Accuracy and Quality

A basic requirement of GIS application is accurate and good quality spatial data. To meet the requirement, we rely on spatial data editing. The editing of spatial data is a time consuming, interactive process that can take a long time. Editing is very important to avoid errors that exist in the data.

Several kinds of errors can occur during data input. They can be classified as:

- Incompleteness of the spatial data. This includes missing points, line segments, and/or polygons.
- Locational placement errors of spatial data. These types of errors usually are the result of careless digitizing or poor quality of the original data source.
- Distortion of the spatial data. This kind of error is usually caused by base maps that are not scale-correct over the whole image, e.g. aerial photographs, or from material stretch, e.g. paper documents.
- Incorrect linkages between spatial and attribute data. This type of error is commonly the result of incorrect unique identifiers (labels) being assigned during manual key in or digitizing. This may involve the assigning of an entirely wrong label to a feature, or more than one label being assigned to a feature.
- Attribute data is wrong or incomplete. Often the attribute data does not match exactly with the spatial data. This is because they are frequently from independent sources and often different time periods. Missing data records or too many data records are the most common problems.
- Topological errors

4. Data Sources and Processes

Data Collection Workflow

In all but the simplest of projects, data collection involves a series of sequential stages, which includes:

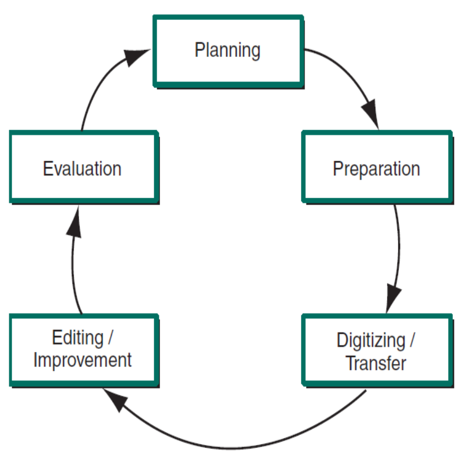
1. Planning

2. Preparation

3. Digitizing and transfer

4. Editing and improvement

5. Evaluation



1. Planning: includes establishing user requirements, gathering resources (staff, hardware, and software) and developing a project plan

2. **Preparation:** includes obtaining data, redrafting poor-quality map sources, editing scanned map images, removing unwanted data and setting up appropriate GIS hardware and software systems to accept data

3. Digitizing and transfer: are the stages where the majority of the effort will be expended.

- It involves inserting the collected data into a computer system.

4. Editing and improvement: This step is designed to: validate data quality, correct errors and improve data quality

5. Evaluation: Is the process of identifying project successes and/or failures.

4.1 Sources of Data

GIS can contain a wide variety of geographic data types originating from many diverse sources. Without geographical data, it is impossible to work with a GIS. The capture and import of data are often the most time consuming, expensive and difficult tasks within the whole chain of work in the GIS environment. Data capture costs can account for up to 85% of the cost of a GIS.

1. Primary Data Sources : It involves GIS data collection by direct observation and measurement of the relevant geographic phenomena.
2. Secondary data source : Are those reused from earlier studies or obtained from other systems.

4.1.1 Primary Data Sources

In GIS primary data sources are:

- **Ground survey with GPS: GPS** uses satellites that orbit Earth to send information to **GPS** receivers that are on the ground. The information helps people determine their location. **GIS** stands for Geographical Information System. **GIS** is a software program that helps people use the information that is collected from the **GPS** satellites.
- **Survey measurements:** Surveying or land surveying is the technique, profession, art and science of determining the terrestrial or three-dimensional positions of points and the

distances and angles between them. It is an important tool for research in many other scientific disciplines

4.1.2 Secondary Data Sources

The secondary data sources for GIS include:

- **Analog map**: An **analog map** is any tangible **map** production that has a continuous appearance and may be viewed directly (generally called a hard copy) Computers are being used in the processing and analysis of remote-sensing data, in the construction of **maps** (known as CAC, computer-assisted cartography), and in the handling of **GIS**.
- **Aerial photography**: is the taking of **photographs** of the ground from an elevated position. The term usually refers to **images** in which the camera is not supported by a ground-based structure. Cameras may be hand held or mounted, and **photographs** may be taken by a **photographer**, triggered remotely or triggered automatically.
- **Satellite data:** Imagery is a type of **data** that is useful for many **GIS** applications and is **defined** as any type of photograph. Normally photographs used for **GIS** projects consist of images gathered from a **satellite** or an aircraft. ... In terms of the specific **GIS data**type, **imagery** is considered raster **data**.
- **Reports and publications:** Reports let you effectively display attribute information about map features in a tabular format that you control. The information displayed in the **report** comes directly from the attribute information stored within the geographic data or stand-alone table in your map.

4.2 Geo Spatial Data Input Techniques

The common geo spatial data input techniques are:

- **Using digitizer tablet:** is a peripheral device that allows users to draw on a computer screen. Tablets are typically used by artists working **with graphics** software such as Adobe Photoshop or Illustrator. Tablets allow for much more precise control than a mouse or trackball does by using a stylus like a pen.
- **Scanning:** Scanning provides a faster means of data entry compared to manual digitizing. In scanning, a digital image of the map is produced by moving an electronic detector across the surface of the map. There are two types of scanner designs: *Flat-bed scanner*: On a flat-bed scanner the map is placed on a flat scanning stage and the detectors move across the map in both the X and the Y directions (similar to copy machine). *Drum scanner*: On a drum scanner, the map is mounted on a cylindrical drum which rotates while the detector moves horizontally across the map. The sensor motion provides movement in the X direction while the drum rotation provides movement in the Y direction. The output from the scanner is a digital image. Usually the image is black and white but scanners can record color by scanning the same document three times using red, green and blue filters.

- **Key board typing:** keyboard is a peripheral device that enables a user to input text into a computer or any other electronic machinery. This device is patterned after its predecessor, the typewriter, from which the **keyboard** inherited its layout, although the keys or letters are arranged to function as electronic switches.
- **GPS:** The Global Positioning System (**GPS**) is a space-based satellite navigation system that provides location and time information in all weather conditions, anywhere on or near the Earth where there is an unobstructed line of sight to four or more **GPS** satellite.

4.3 Data Organization and Storage

Raster data organization

There are four ways to organize your raster data (raster data models): the raster dataset, the mosaic dataset, the raster catalog, and raster that are attributes of a feature.

One of the major design decisions in managing raster data is whether to store all the data in a single dataset or as a cataloged collection of many datasets. In some cases, the application requirements will dictate which approach to use, while in other cases, the choice may not be as obvious. The mosaic dataset provides an excellent solution to this decision, as it is a catalog of raster data that can be viewed as a single raster dataset (mosaicked image).

Raster datasets

A raster dataset is any valid raster format organized into one or more bands. Each band consists of an array of pixels (cells), and each pixel has a value. A raster dataset has at least one band. More than one raster dataset can be spatially appended (mosaicked) together into a larger, single, continuous raster dataset.

Mosaic datasets

A mosaic dataset is a collection of raster datasets (images) stored as a catalog and viewed as a mosaicked image. The raster datasets in a mosaic dataset can remain in their native format on disk or, if required, be loaded into the geodatabase.

Like a raster catalog, each raster within the mosaic dataset can have its own coordinate system, but it will be reprojected to the coordinate system of the mosaic dataset when displayed. Additionally, any footprints or other calculated attributes, such as pixel sizes, will be created using the coordinate system of the mosaic dataset.

The raster data in a mosaic dataset does not have to be adjoining or overlapping but can exist as unconnected, discontinuous datasets.

Raster catalogs

A raster catalog is a collection of raster datasets defined in a table format in which each record represents an individual raster dataset in the catalog. A raster catalog can be large and contain thousands of images. A raster catalog is typically used to display adjacent, fully overlapping, or partially overlapping raster datasets without having to mosaic them into one large raster dataset.

Raster Storage Models

When working with multiple raster datasets, there are four possible storage models: store each raster dataset individually, mosaic them into one large raster dataset, store them as members of a mosaic dataset, or store them as members of a raster catalog.

Storing the raster datasets individually is often the best method when the datasets are not adjacent to each other or are rarely used on the same project.

4.4 Data Editing and Updating

If you have collected data from a variety of sources, chances are that not all layers contain the same coordinate system information. The coordinate system of a data frame in ArcMap can be different from the native coordinate system of the data sources represented by the layers shown in the data frame. In this case, ArcMap projects (on the fly) the features in these layers to the data frame's coordinate system. ArcMap also lets you edit features while they are projected.

Most organizations update or change the geometry and attribute table related to the reference data on a regular basis. This is a basic part of keeping the data current. New subdivisions or streets might be added, alternate names might be used for streets, or a location name might have become commonplace. Additional regions or districts might be added to the database, or previous errors may have been corrected. Each of these changes will alter the reference data. To geocode addresses against the current version of the reference data, the address locator must be rebuilt if you want to update the changes in the locator.

Generally, it is possible to edit geographic data in GIS environment or other software packages. Some tasks related to editing are

- projection or coordinate system editing,
- editing features in topology
- editing attributes
- editing during digitizing

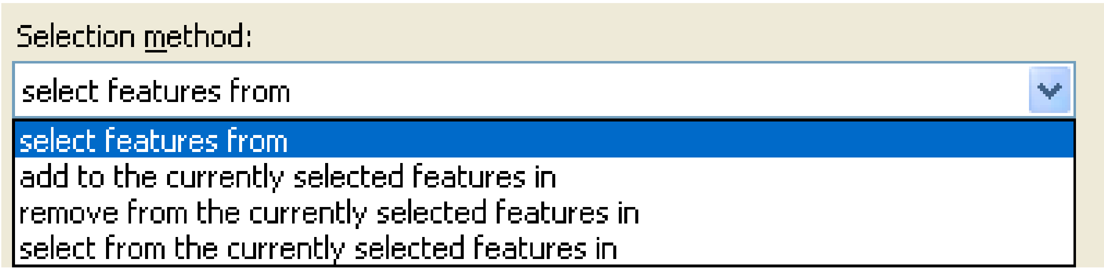
4.5 Data Querying and Retrieval

4.5.1 Querying by Attribute

With GIS, you can identify a subset of map features based on their attributes or their location. You may use queries as an intermediary step, as part of getting to know your data, to create new values, or to answer your research questions.

Select by Attribute

To start a query, from the Selection menu go to “Select by Attributes.” Your first choice is what layer in your map you wish to query (attribute queries are limited to a single layer). In the Method drop down menu, you need to decide whether you are building on a previous query or starting fresh.



From here, the query dialog is similar to the Field Calculator. Build an expression using the field names and functions, connecting statements with “AND” or “OR” as necessary. If you want help working through these steps, click on the Query Wizard button. Click “Verify” to make sure Arc Map likes your expression.

Note that if the values you input are text, rather than numbers, you need to put single quotes around them (for example, ‘BROAD’). When you click “Apply,” the map features that satisfy your query will be highlighted with a bright blue outline.

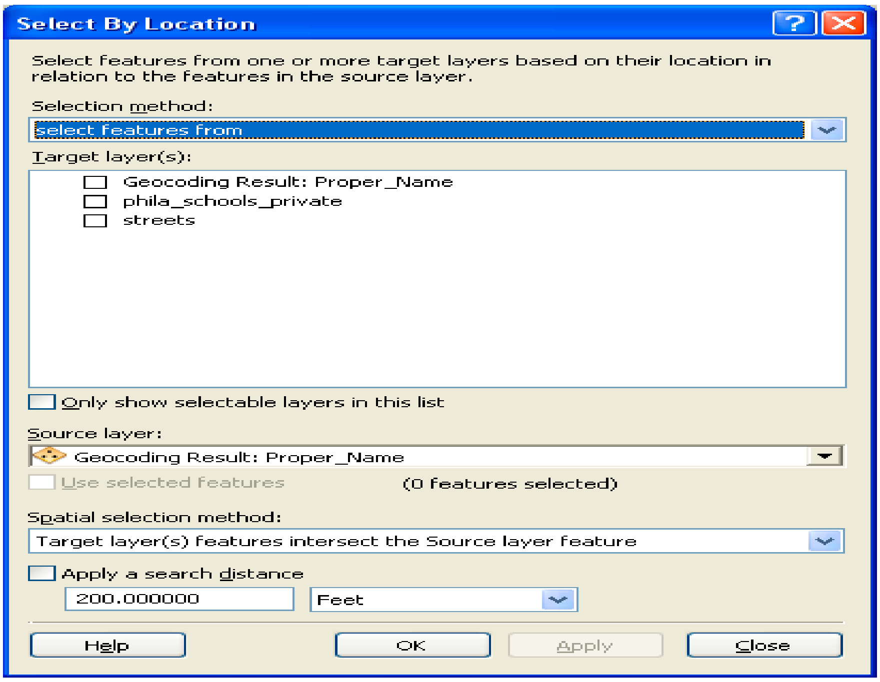
To review the selected records, right click on the map layer and go to “Open Attribute Table.” Click on the Select button at the bottom to view only the selected records. You can calculate values on an existing or new field for only the selected records. You can also look at summary statistics for only the selected records by right clicking on a column name and going to “Statistics.”

4.5.2 Querying by Location

The ability to query based on the location of map features is something unique to GIS, and combining attribute and location queries really takes advantage of GIS functionality.

Select by Location

To select features in a map layer based on their location, from the Selection menu go to "Select by Location." Your first choice is the selection method. Are you starting from scratch or querying a subset of map features already selected? In the next box, put a check mark next to the map layer(s) whose features you wish to select.



The drop-down menu then provides various relationships between the features in that layer and the "source" layer, identified in the next drop-down menu. For distance-based queries, you will be able to select your units as long as the map units have been specified in the Data Frame properties.

5. Data Visualization

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

Visualization: - display/showing capability of GIS, hence it can display vector and raster data in 2D and 3D views. Some visualization techniques are: - Maps /2D views and 3D methods like Arc scene and Arc globe

5.1 GIS and Map

A map is a representation of spatial or geographic information as a series of thematic layers of information for an area of interest. Map is two-dimensional surface that shows the earth features partially/ fully with defined scale. It is simplified representation of feature. It can be analog map/ hard copy map or it can be digital map /soft copy map, most GIS maps are digital map

A printed map also includes additional map elements laid out and organized on a page. The map frame provides the geographic view of information while other elements, for example, a symbol legend, scale bar, north arrow, descriptive text, and a map title, around the map collar help you to understand, read, and interpret the map's contents.

People also work with computer maps—interactive images on computer screens with tools that allow you to interrogate and interact with the map's underlying geographic information.

- Based on their purpose/ information / type map can be grouped in to two

    1) General reference map/ top sheet or top map:- it is show both man made and natural features of the map.
    2) Thematic map/ special purpose map it is a map that focuses on a specific *theme* or subject area such as physical phenomena like temperature variation, rainfall distribution and population density in an area. it is display one feature E.g. Soil map, road map, cadastral map etc

Maps play a special role in GIS:

- They portray logical collections of geographic information as map layers.
- They are at the heart of how GIS is used.
- They provide an effective metaphor for modelling and organizing geographic information as a series of thematic layers.
- Maps identify what is at a location.
- Maps can locate where you are.
- Maps let you identify distributions, relationships, and trends not otherwise visible.
- Maps can integrate data from diverse sources into a common geographic reference
- Maps can find the best path between one place and another
- Maps can model future events

In addition, interactive GIS maps provide the focal point for using geographic information and bringing that information to life. GIS maps are the way that GIS content is shared among professional GIS users and with everyone online.

This topic provides an important context for the role that maps play in delivering GIS to many new users.

Maps are important, and almost everyone understands and appreciates good maps.

And mapping encompasses a lot—from traditional printed maps and imagery to new media maps that are used on computers, across the Web, and on mobile devices.

A new kind of map is a **GIS map**, and each GIS map is more than a static map presentation. It is an interactive window into all geographic information and descriptive data, and into rich spatial analysis models created by GIS professionals.

5.2 Visualization process

Visualization process is considered to be the translation or conversion of spatial data from a database into graphics. These are predominantly map like products. During the visualization process, cartographic methods and techniques are applied. 2D/ hard copy: - length + width/ latitude or longitude 3D/ software x.y.z or latitude, longitude, altitude or elevation.

- Some Examples of Visualization Includes:
  - the creation of a full, traditional topographic map sheet
  - a newspaper maps
  - a sketch maps
  - a map from an electronic atlas
  - an animation showing the growth of a city
  - a three-dimensional view of a building or a mountain
  - or even a real-time map display of traffic conditions

The paper map has long been a powerful and effective means of visualizing and communicating geographic information. In contrast to digital map or data, it is an instance of an *analog* representation, or a physical model in which the real world is scaled.

A paper map is a source of data for geographic databases; an analog product from a GIS; and an effective communication tool. Maps have been so important, particularly prior to the development of digital map, that many of the ideas associated with GIS are actually inherited directly from paper maps.

- The contents of the map can be visualized using six visual Variables: these are:
    - Size
    - Value (lightness)
    - Texture (Grain)
    - Color
    - Orientation and
    - Shape

5.3 Mapping Data

Arc Map provides two ways to view a map: data view and layout view. Each view allows you to view and interact with the map, but in different ways. Data view provides a geographic window for exploring, displaying, and querying the data on your map. You work in real-world coordinates and measurements in data view.

In layout view, you work with the map layout elements, such as titles, north arrows, and scale bars, along with the data frame, all of which are arranged on a page. In layout view, you work primarily in page space (typically, inches or centimetres) except when you are interacting with a data frame in your layout.

Once geographic features and data have been selected, generalized and classified for the map, it is necessary to choose the appropriate graphic representation or symbols for the information. Visual variables include symbol, size, shape, orientation, pattern (texture), hue (colour), and colour value (brightness and lightness). When a data set is large, it is not practical to assign a unique symbol to each data record. Therefore, for mapping it is essential that data is classified or grouped. Before classifying or grouping data, it is necessary to determine whether the data are qualitative or quantitative, and the level of measurement.

These phenomena can be classified into four basic categories: *point* (non-dimensional data), *line* (one-dimensional data), *area* (two-dimensional data), and *volume* (three-dimensional data).

The geographical data must be represented on maps by only three basic symbol types: *point, line,* and *area.*

Before assigning map symbology, it is important to have a good understanding of the data set to be mapped.

5.4 Mapping Qualitative Data

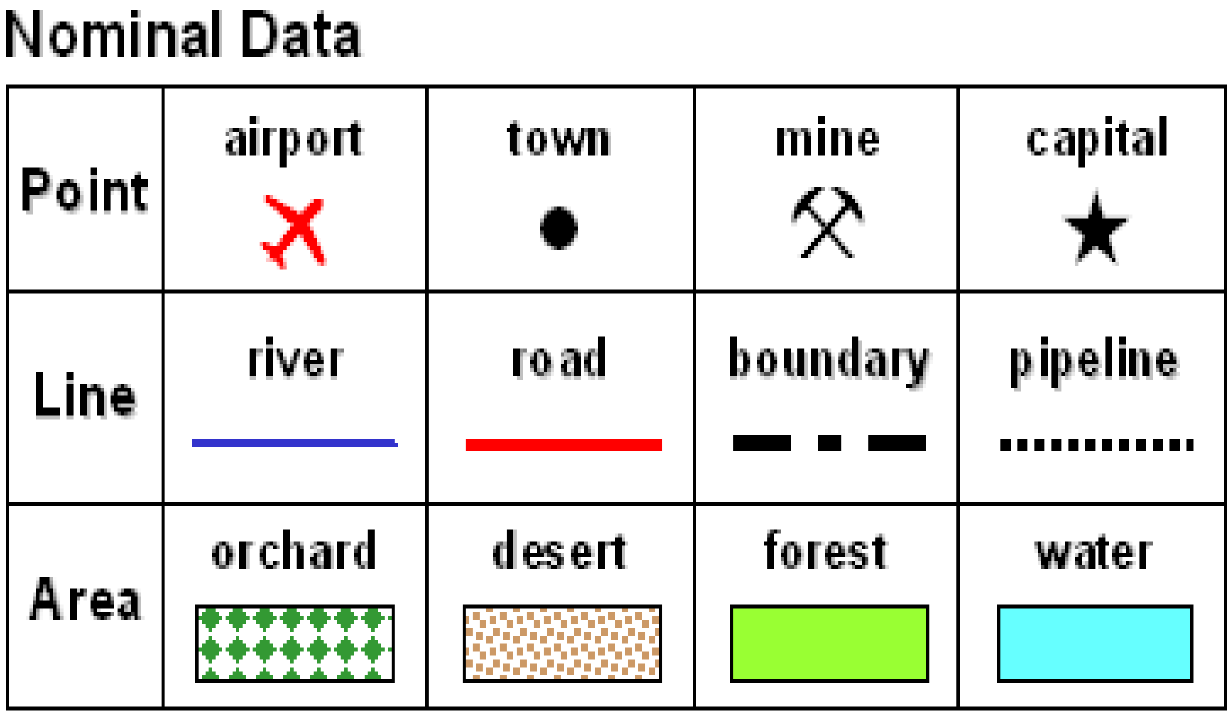
Qualitative data are data that are grouped in classes according to differences in type or quality. It is information about qualities; information that can't actually be measured. Deals with descriptions. Data can be observed but not measured.

- Example: Land Cover classes we can assign the type and its name (farm land, urban area, grass land), Soil Types and Political administration divisions

Qualitative data have no numerical values attached. *Nominal data* comes under this category. Sometimes ordinal data may also be considered qualitative, if no numerical values are involved.

Nominal Data

Nominal data are *discrete* and are classed according to type or quality. For example, a line could represent either a road or river, and a land use polygon could be residential, commercial, or a recreational area. Nominal data are often labeled with numbers or letters, but these labels do not imply ranking. Nominal data can be shown as point, line and area symbology.



5.5 Mapping Quantitative Data

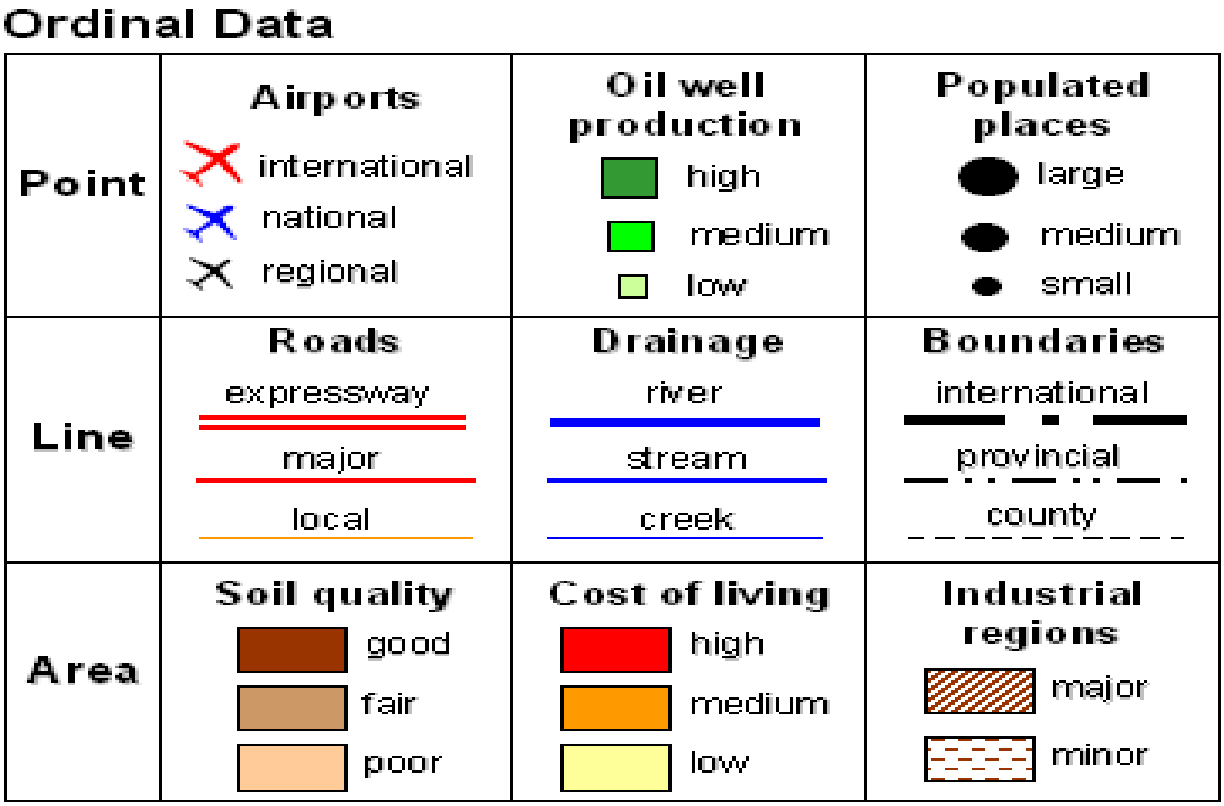
Quantitative data are data that contain attributes indicating differences in amount and can be expressed as numerical values. Quantitative data Included in this category are *ordinal*, *interval*, and *ratio*.

1. Ordinal data

Ordinal data provide information about rank or hierarchy, in other words, relative values. In Ordinal data, therefore, it is possible to describe one item as larger or smaller than another, or as low, medium, or high.

However, it is not possible to measure the differences between ordinal data, because there are **no specific numerical values** attached to them.

Examples of Ordinal Data as point, line & area map



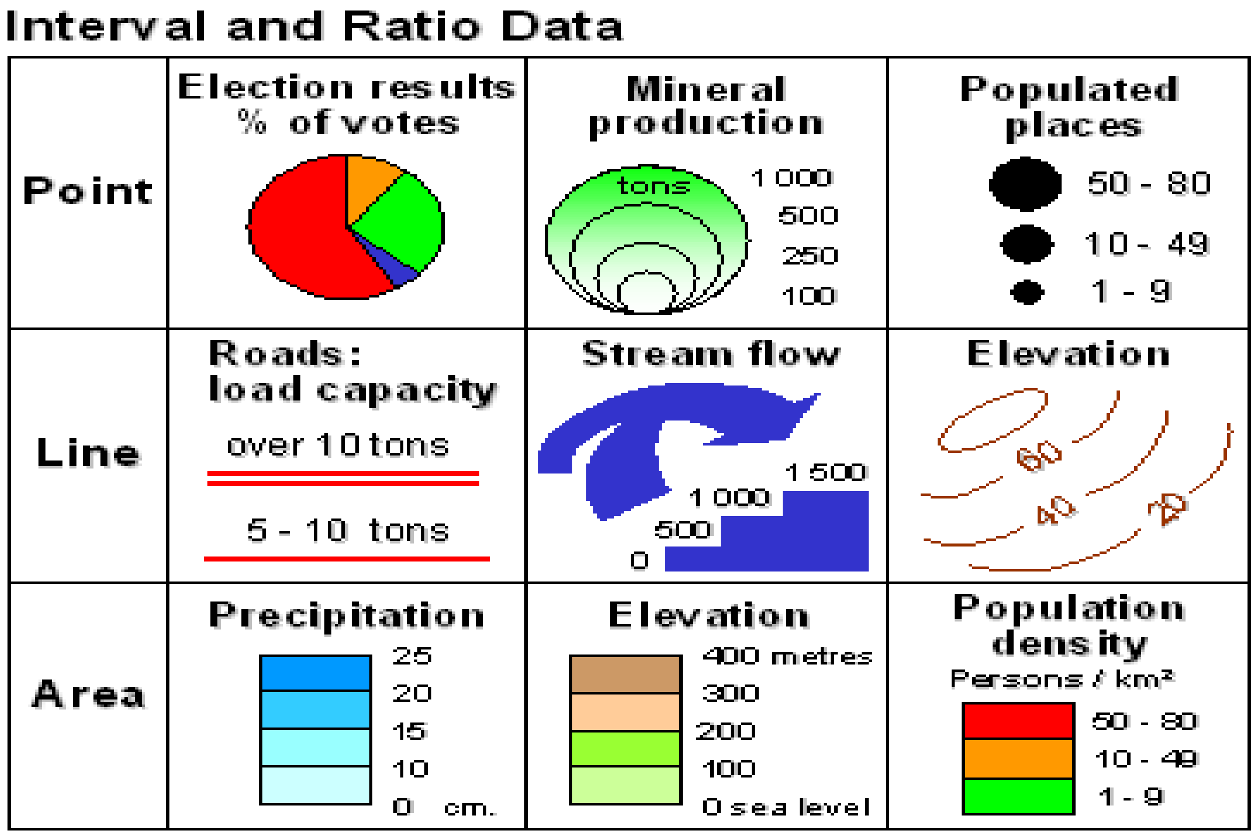
2. Interval data Ratio data

The information can be arranged along a scale using a standard unit. It is possible to calculate the distance or difference between ranks, which must be expressed in terms of a standard unit. For example, a temperature scale uses degrees (°F or °C) as a standard unit of measurement; between 20° and 35° there is a difference of 15°.

As shown by this example of interval data, it cannot be said that 35° is 1.75 times warmer than 20°, because the scale on which temperature is measured is arbitrary. Interval data, as illustrated, have no *natural zero*.

Ratio data are the same as interval data, except there is a *natural zero;* therefore, it is possible to express data as ratios.

Physical measurements of height, weight, and length are examples of ratio variables. With this type of data, it is meaningful to state that a measurement is twice that of another. In some literature and statistical computer programs, no distinction is made between interval and ratio data, calling them both *continuous data*.



There are many statistical methods for the classification or ranging of interval/ratio data.

5.6 Mapping Terrain Elevation

A terrain dataset is a TIN-based dataset that uses geodatabase feature classes as data sources. A triangulated irregular network (TIN) is a data structure used to model surfaces, such as elevation, as a connected network of triangles.

TINs are assembled from a series of data points with x-, y-, and z-values and partition geographic space into contiguous, non-overlapping triangles (called faces).

Difference between DEM and TIN

DEM

- 3D data for raster file
- Pixel is the smallest unit
- Show elevation as a serious of pixel

- Cell/pixel are equal in size
- Important to study slope, contour, hill shed and water shed feature (e.g. flow length, flow direction, basin, catchment are etc.)

TIN

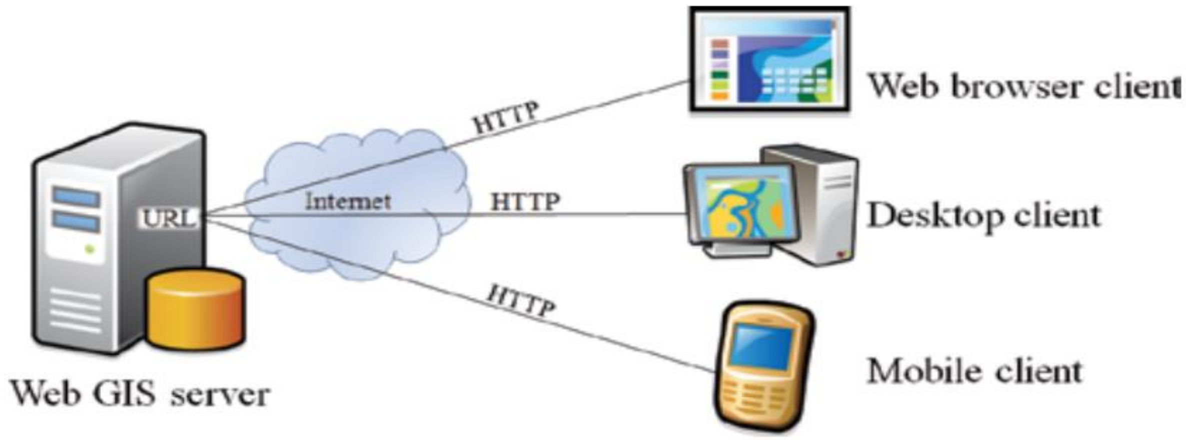
- 3D data model for vector file
- Triangle is the smallest unit
- Show elevation as a network of triangle
- Triangle are a continuous, non-overlapping, and unequal in size
- Used to study land structure like fault valley, crack, depression etc.

6. Web Technology for GIS and Mapping

Definition and concept of Web GIS

Web GIS has considerably changed the way geospatial information is acquired, transmitted, published, shared, and visualized. It is a geographic information system distributed across a networked computer environment to integrate, distribute, and communicate geographic information visually on the World Wide Web.

Web GIS is a type of distributed information system, comprising at least a server and a client, where the server is a GIS server and the client is a web browser, desktop application, or mobile application. Web GIS is a suitable and efficient tool for distributing integrated spatial and non-spatial information related to GIS. Web GIS requires a suitable Internet Map Server.



Web GIS functions

- **Web mapping (Visualization) and query**: Web mapping, as the face of WEB GIS, is the most commonly used functions.

  GIS data and analysis result are usually presented as a map.

- **Collecting/Editing geospatial information:** using the web to create and assemble a trove of geographic information.
- **Dissemination of geospatial information:** Government agencies, the academic sector, and some commercial sectors have long used Web GIS to share geospatial information.
- **Geospatial Analysis:** WEB GIS has gone beyond mere mapping. It also provides analytical functions most notably, those closely related to daily life, such as measuring distance and areas, finding the optimum driving path, finding the location of an address or place, and using proximity analysis to find the businesses nearby**,** Ask for information about features displayed on map, Point based queries on map data ,Allow down-to-top information flow, Crowd sourcing up-to-date information, Wide distribution of information and Provide customized analytical functions.

Advantage and disadvantage of WEB GIS

- **Advantages**
  - GIS Operation online
  - No need of GIS applications
  - No need of GIS knowledge
  - Easy to use
  - Can be done on any device (mobile / pc)
- **Disadvantage**
  - Need powerful server
  - Dependent on server availability
  - Dependent on internet availability
  - Users need to be aware of web service

6. 1 Principles of internet and the web

As Internet GIS if it uses many of services of Internet not only Web service. It uses only Web; we should name it Web GIS. Internet is a huge network connecting computers throughout the world, so that this computer and their contents are easily accessible to each other. This definition makes internet GIS border than Web GIS, in real world Web is the most attractive service of internet.

6.2 Principles of Open standards and web GIS

Like all Web services, GIS services are built to perform a focused set of GIS functions and utilize open Web protocols, such as XML, SOAP, and REST Web services, and HTML. There are a number of types of GIS services:

- **Map services:** It can be added as layers in ArcMap, Arc Scene, or Arc Globe. This is the most common type of GIS service. Map service layers provide the following:
    - *Base map layers* onto which you can overlay and work with operational information.
    - *Mission-critical operational information* for performing specific tasks. For example, you can use a Web map layer containing real-time sensor feeds, weather information, crime locations, traffic conditions, property values, water flows, emergency events, and so on.
    - *Layers* that represent information derived from analytic model results.
    - *Access to the underlying features and raster information for use in queries, graphs, and reports.*
- **Image services:** used as map layers, to generate integrated image mosaics from multiple image sources, and to access the pixel values in imagery and raster datasets.
- **Geo processing services:** used to run analytic models and automate tasks that can be run on a shared server.
- **Geodatabase services:** used to support queries and editing against a centralized, multiuser geodatabase.
- **Locator services:** used to geocode place-names, addresses, and other location information.

GIS Web map services have some important characteristics that influence how they are used. Perhaps the most interesting is that they are georeferenced, which means that you can readily combine GIS services from many servers across the Web. Arc Map, Arc Scene, and Arc Globe support a range of map services.

Most people love the quality of information from, ease of use, and performance of Google Earth, Google Maps, and Microsoft Bing Maps. The primary value of Google Earth, Google Maps, and Bing Maps for GIS is as a base map. These online Web maps provide a common base map framework onto which geographic information can be overlaid and used.

Applications of Web GIS

- **Web GIs as a new business model and a new type of commodity: -** business model is the placement of advertising based on web mapping. It is model used by Google, Microsoft and yahoo.
- **Web GIs as an essential component of daily life: -** Every day you need to answer questions such as where to eat, where to stay, where to shop, and how to get from here to there.
- **Web GIS for Tourism Development: -** Tourists need to find out the relevant distance from the airport to the hotel, the distance between different attractions and the accommodations, the exact position of tourist sites and other facilities easily.
- **WEB GIS TO SUPPORT IRRIGATION MANAGEMENT;** The ability of GIS to analyze and visualize spatial and non-spatial data in the form of maps made it an essential tool for agricultural irrigation management systems.
- **WEB GIS used to improve Environmental monitoring:** The Web-based GIS application is a solution suggested for interactive decision-making allowing better understanding of environmental processes and more effective monitoring and management.