UNIT - 1

FUNDAMENTALS OF GIS

1.1. GIS OVERVIEW

GIS refers to three integrated parts.

a) **Geographic**: Of the real world; the spatial realities, the geography.

b) **Information**: Data and information; their meaning and use.

c) **Systems**: The computer technology and support infrastructure.

GIS therefore refers to a set of three aspects of our modern world, and offers new ways to deal with them. The concept of information is indeed the heart of the rapidly growing field of Geographic Information Systems or GIS. As the world moves into the Information Age, meaningful data and information are becoming the major 'currency'. With the continuous advances in computer technology, it is easy to concentrate on dazzling systems and software, but the real value of any product is the data and the information such data provides. At the heart of any GIS is information.

1.1.1. Introduction to GIS

(i) A **geographic information system** (**GIS**) is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data in maps, and present the results of all these operations. GIS (more commonly GIScience) sometimes refers to geographic information science (GIScience), the science underlying geographic concepts, applications, and systems.

(ii) GIS can refer to a number of different technologies, processes, techniques and methods. It is attached to many operations and has many applications related to engineering, planning, management, transport/logistics, insurance, telecommunications, and business. For that reason, GIS and location intelligence applications can be the foundation for many location-enabled services that rely on analysis and visualization.

(iii) GIS can relate unrelated information by using location as the key index variable. Locations or extents in the Earth space–time may be recorded as dates/times of occurrence, and x, y, and z coordinates representing, longitude, latitude, and elevation, respectively. All Earth-based spatial–temporal location and extent references should be relatable to one another and ultimately to a "real" physical location or extent. This key characteristic of GIS has begun to open new avenues of scientific inquiry.

1.2. BASIC SPATIAL CONCEPTS

Spatial Concept seeks to position spatial concepts as the driving force for spatial thinking and for the selection and use of spatial tools. Eight concepts are the focus of spatial reasoning in the use of geographical information. These concepts are demonstrable at all levels of space and time (from sub-atomic to galactic, passed through future, and microseconds to ions). They can be rendered understandable through simple illustrations to young children but they are also sufficiently engaging at advanced levels for thinking about scientific and social problems.

Location - Understanding formal and informal methods of specifying “where”

Distance - The ability to reason from knowledge of relative position

Network - Understanding the importance of connections

Neighborhood and Region - Drawing inferences from spatial context

Scale - Understanding spatial scale and its significance

Spatial Heterogeneity - The implications of spatial variability

Spatial Dependence - Understanding relationships across space

Objects and Fields - Viewing phenomena as continuous in space-time or as discrete

These concepts have been a foundation for researchers for centuries (see Classics in Spatial Thinking). They have been augmented in recent decades with computational and visualization tools and with vast and easily accessible information resources. These concepts and tools must be as central to general education as reading, writing, and arithmetic. In conjunction with the appropriate spatial tools, they provide a basic scaffold for designing research, solving problems, and structuring education programs.

Spatial concepts invites contributions about other concepts for spatial thinking (e.g., in design fields and in the humanities and arts). Examples may include the link between form and function in architecture, the search for pattern in speech and text, the use of spatial notation in music, the use of spatial metaphor in the sciences and humanities, the importance of place in cultural and social studies, and the spatial elements of aesthetics in the visual arts. If you wish to add concepts to the listing.

1.2.1. Location:

Every type of spatial data has two components: a location and some attribute(s). In this way, location can be seen as a fundamental trait that both defines spatial data and separates it from other types of information. Broadly speaking, there are two types of locations: absolute and relative. Absolute location refers to an exact position on the Earth’s surface defined by some coordinate system. Street addresses and latitude/longitude coordinates are good examples of absolute locations. Relative locations, on the other hand, are defined in reference to other objects. For example, one could define UCSB’s relative location as 10 miles west of downtown Santa Barbara.

The method of determining or measuring location is called georeferencing. There are many different ways of defining a georeference, but they all must meet a few requirements. First, the georeference must be unique so that only one location is described. Second, georeferences must have an accepted meaning that is shared so that most users understand its implication. Finally, georeferences must be viable throughout time so that their meaning is not lost. Metric georeferences are those which define location by measurement and are of particular importance in analyzing spatial data. In order to properly define a metric georeference, consideration must be given to the shape of the Earth, map projections and coordinate systems, and positional accuracy.

Representing location is also an important concept in the use of spatial data. Spot locations are represented as zero-dimension points, lines as one-dimensional poly-lines, areas as two-dimensional polygons, and volumes as three-dimensional polyhedra.

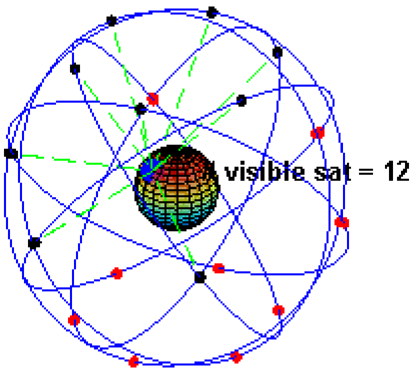


Fig. 1.1. Related Images Location Example (GPS image)

1.2.2. Distance

Distance describes the measurement or separation of two objects or places. At its most basic, measured distance provides a mechanism for describing spatial extent. In spatial reasoning, the classical case is that a positive relationship exists between distance and locational attribute similarity. The relationship, as worded by the Swiss-American cartographer, Waldo Tobler, is, "Everything is related to everything else, but near things are more related to each other." Examining the strength, limitations, and exceptions of the distance-similarity correlation serves as the foundation of much quantitative and qualitative spatial research. Measured distance also provides a necessary dimension for describing position. All real points can be plotted with respect to another by coupling measures of distance with heading and/or time. Common application of distance include spatial analysis and modeling, physics and gravitation, distance decay, buffers, geodesics, route description and optimization, and qualitative comparison of place.

1.2.3. Network

A network is a physical or conceptual system of linkages among entities. Networks offer an infrastructure for representing the anisotropic relationships of various constituents and constituent attributes. Typically, a network connection denotes increased accessibility or relatedness along a link, and linkages may override the default notion that closer features are more accessible and related. For instance, driving from one side of a river to the other often requires a circuitous path many orders of magnitude greater than the direct distance from origin to destination.

Networks commonly are represented by the elements: linkages, nodes, and intersections. Carefully crafted associations among these elements facilitate analyses regarding least-cost path optimization, measures of separation and similarity, and emergent spatial structures. Additionally, attributes can be assigned to various network elements to allow for directionality (e.g., a one-way street), intersection policy (e.g., a no U-turn rule), cost (e.g., a speed limit), and system regulation (e.g., synchronized traffic signals to control flow). Some applications of networks include traffic management, delivery systems, social structures, river hydrology, and communication systems.

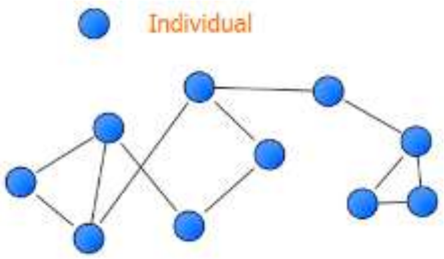


Fig.1.2. Network Example (Social Network Diagram)

1.2.4. Neighborhoods and regions

Neighborhoods and regions define areas surrounding and containing spatial data. They may be formal in nature, such as state and country boundaries, and informal, such as the colloquial use of terms like "downtown." Regions may also be defined in terms of a particular function. For example, the functional region for a pizza restaurant may be the area within a city to which it delivers.

Utilizing neighborhoods and regions can allow one to make inferences about data from its spatial context. Many fields use neighborhoods and regions in just this way. In remote sensing, neighborhood statistics can be calculated using the values of adjacent pixels. In business, service regions are defined to maximize the number and locations of stores or restaurant franchises. In landscape ecology, metrics have been developed to measure the fragmentation of environmental patches within an ecosystem.

There are problems with making inferences between data at different scales, however. The modifiable area unit problem (MAUP), or ecological fallacy, involves two issues that underscore the use of data aggregated to the neighborhood or region. First, there is a problem of scale. In this way, results of analysis at one scale are not comparable to results at higher or lower levels of aggregation. For example, it does not make sense to use population density calculated at the state level to describe density at the county level. Second, there is a problem of aggregation. Changing region boundaries can have immense effects on the neighborhood statistics. The gerrymandering of voting districts is one example of an aggregation problem.

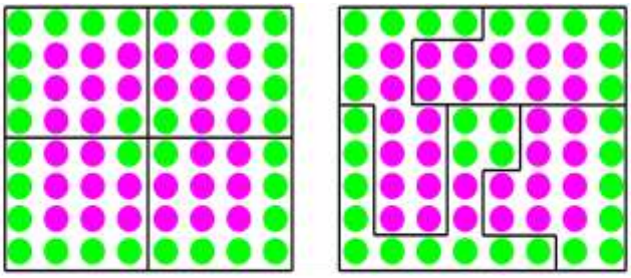


Fig. 1.3. Neighborhood/Region Example (Gerrymandering diagram)

1.2.5. Scale

In spatial reasoning, scale describes the dimensional relationship between a representation and reality. Due to the large variation of all space, scale is used to project reality to more useful and meaningful sizes. For large expanses, scale is reduced (e.g., fitting the entirety of earth's surface on to a paper map), and for miniscule distances, scale is increased (e.g., enlarging and schematizing chemical reactions). Often, scale is denoted as a fraction where a unit of measure in reality is compared to the same unit on the projection. For instance, a paper map of a landscape showing a scale of 1:10,000 would mean that a drawn unit on the map represents 10,000 of the same unit in reality.

Offering spatial data at scales different from their original data collection granularity may imply changes of meaning; thus, data product limitations should be carefully considered by both creators and consumers of spatial information. As a fundamental spatial concept, applications of scale permeate many human activities including data collection, cartography, art, architecture, engineering, and nanotechnology.

1.2.6. Spatial heterogeneity

Spatial heterogeneity refers to the degree of variation in some attribute across places and region. For example, a satellite image of the Pacific Ocean would show little variation and thus would have a low level of spatial heterogeneity; whereas, an image of the patchwork of agricultural fields in the Midwestern U.S. may be considered highly spatially heterogeneous. In the same way that biodiversity defines species variation in biology, spatial heterogeneity defines variation of an attribute in spatial studies.

There are many research implications caused by the fundamental heterogeneity of spatial data. In most cases, spatial data can be said to have a non-constant mean and variance throughout a study area. That is, local statistical parameters change with location and are thus not uniformly distributed. This characteristic of spatial data is termed non-stationarity and directly impacts research in areas such as sampling design. For example, it is very difficult to obtain a representative sample of a region because of the heterogeneous nature of many spaces.

Another characteristic of spatial data that impacts research is the general rule that spatial data tend to become more heterogeneous as the study area gets larger. This means the variance observed in a small region is less than that in a larger region. That is, expected variation around the mean in small regions underestimates and, thus, is not applicable in subsequently larger regions. This uncontrolled variance in spatial data has impacts in the study design of projects involving large areas and times, such as global warming.

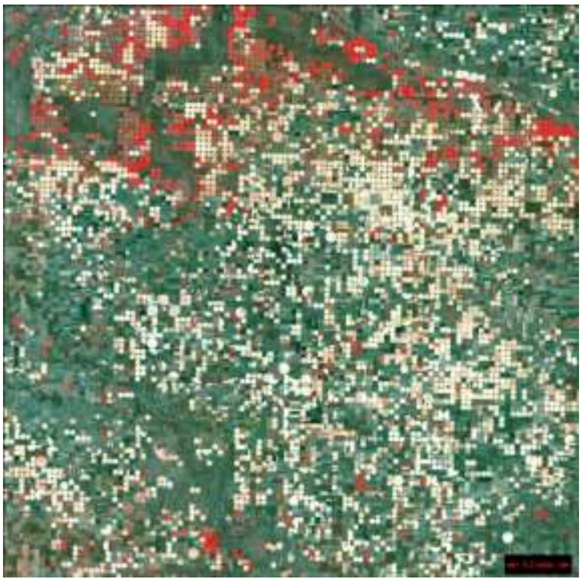


Fig.1.4. Spatial Heterogeneity Example (Landscape Patch image)

1.2.7. Spatial dependence:

Spatial dependence is the manifestation of Tobler's first law of geography, which states, "Everything is related to everything else, but near things are more related to each other." This seemingly simple principle applies to many types of spatial attributes. For example, one would expect the temperature of Santa Barbara to be more similar to the temperature in Los Angeles than to the temperature in Seattle. This relatedness between data based on the distance is also termed spatial autocorrelation. If spatial data were truly random, there would be zero spatial autocorrelation. Applying this idea to the previous example would mean that no inference of similarity could be made between the temperatures in Santa Barbara and Los Angeles. Fortunately, most attributes in the world are not distributed in this manner and thus display a certain degree of spatial dependence.

Researchers can quantify spatial correlation by using indices such as Moran's I and Geary's C. Researchers can also model spatial dependence by using the methods developed in the field of geostatistics. Techniques such as kriging allow researchers to quantify the changes in attribute variance versus distance. This model of spatial dependence is called a semi-variogram, or just variogram. Variograms offer a mechanism to predict attribute values in locations where data are not present. These powerful interpolation tools have many real-world applications in fields such as mining and petroleum discovery, epidemiology, atmospheric science, oceanography, and soil science.

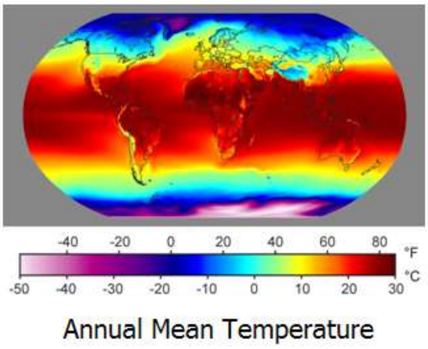


Fig. 1.5. Spatial Dependence Example (Temperature Map)

1.2.8. Objects and fields:

Objects and fields describe two fundamental, dichotomous, conceptualizations of space. Objects are collections of discrete, bounded entities, usually composed of geometric primitives, such as points, lines, curves, and polygons. Conceptually, object-based reality is considered empty space populated by distinct entities. Attributes are associated with bounded features to define and describe the objects. Conversely, field-based representation fills continuous space with attribute measures at all locations. A field offers a conceptual model of spatial variation, and attributes themselves, as opposed to workings of distinct boundaries, define the field.

The object-field dichotomy serves as the underpinning for all methods of spatial representation and analysis, and each perspective offers a characteristic set of abilities and limitations. For instance, objects offer logical concepts like "inside" and "outside," and comparisons such as intersections and buffers. Fields are particularly suited for spatial phenomena considered less bounded, like air, water, temperature, and elevation. Not constrained by distinct edges, fields lend themselves to analyses related to interpolation and global attribute comparison. Analysis and representation of fields often yield areas of peaks, valleys, aspects, and slopes.

Fig. 1.6. Object Example (Cadastral Map)

The representational and analytical potential of both perspectives is mirrored in computer software for modeling reality. Modeling software for analyzing property boundaries, road networks, and individuals, tend to use object-based representation methods. Software for analyzing less-bounded phenomena, like temperature variation, ocean circulation, and hydrologic modeling, typically use field-based models for database storage and analysis.

Ongoing research on the object-field dichotomy has focused on repercussions of accuracy, uncertainty, and usage, the philosophy of spatial reality within the context of such representations, and defining and utilizing object-fields hybrids. Common domains using objects and fields are Geographic Information Systems, spatial ontology, spatial analysis, and database design.

1.3. GEOGRAPHIC COORDINATE SYSTEMS

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 (which crosses Greenwich England). A location is therefore defined by two values: a latitudinal value and a longitudinal value.

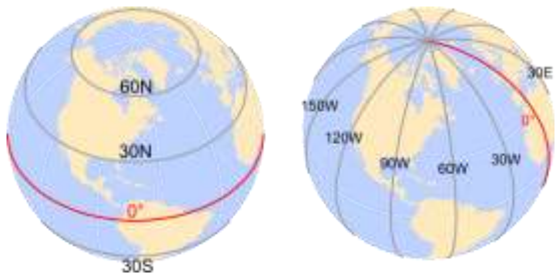


Figure 1.7: Examples of latitudinal lines are shown on the left and examples of longitudinal lines are shown on the right. The 0° degree reference lines for each are shown in red (equator for latitudinal measurements and prime meridian for longitudinal measurements).

A latitude measures the angle from the equatorial plane to the location on the earth's surface. A longitude measures the angle between the prime meridian plane and the north-south plane that intersects the location of interest. For example Colby College is located at around 45.56° North and 69.66° West. In a GIS system, the North-South and East-West directions are encoded as signs. North and East are assigned a positive (+) sign and South and West are assigned a negative (-) sign. Colby College's location is therefore encoded as +45.56° and -69.66°.

Figure 1.8. A slice of earth showing the latitude and longitude measurements.

A GCS is defined by an ellipsoid, geoid and datum. These elements are presented next.

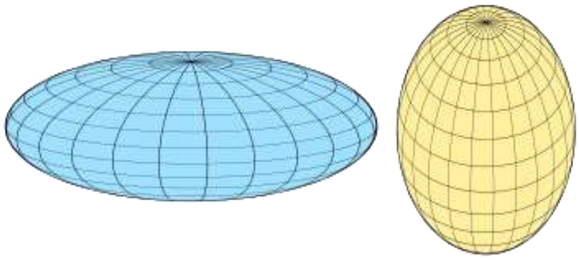
1.3.1. Sphere and Ellipsoid

Assuming that the earth is a perfect sphere greatly simplifies mathematical calculations and works well for small-scale maps (maps that show a large area of the earth). However, when working at larger scales, an ellipsoid representation of earth may be desired if accurate measurements are needed. An ellipsoid is defined by two radii: the semi-major axis (the equatorial radius) and the semi-minor axis (the polar radius).

The reason the earth has a slightly ellipsoidal shape has to do with its rotation which induces a centripetal force along the equator. This results in an equatorial axis that is roughly 21 km longer than the polar axis.

Our estimate of these radii is quite precise thanks to satellite and computational capabilities. The semi-major axis is 6,378,137 meters and the semi-minor axis is 6,356,752 meters.

Differences in distance measurements along the surfaces of an ellipsoid vs. a sphere are small but measurable (the difference can be as high as 20 km) as illustrated in the following lattice plots.



1.9. Fig. Spheroids with vertical rotational axes

1.3.2. Geoid

Representing the earth's true shape, the geoid, as a mathematical model is crucial for a GIS environment. However, the earth's shape is not a perfectly smooth surface. It has undulations resulting from changes in gravitational pull across its surface. These undulations may not be visible with the naked eye, but they are measurable and can influence locational measurements.

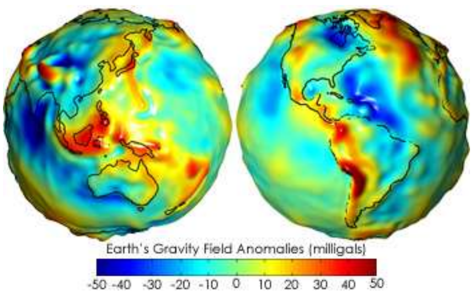


Figure 1.10. Earth's geoid with gravitational field shown in rainbow colors. The ondulations depicted in the graphics are exaggerated for visual effects. (source: NASA)

Note that we are not including mountains and ocean bottoms in our discussion, instead we are focusing solely on the earth's gravitational potential which can be best visualized by imagining the earth's surface completely immersed in water and measuring the sea surface level over the entire earth surface.

The earth's gravitational field is dynamic and is tied to the flow of the earth's hot and fluid core. Hence its geoid is constantly changing, albeit at a large temporal scale.The measurement and representation of the earth's shape is at the heart of geodesy–a branch of applied mathematics.

1.3.3. Datum

So how are we to reconcile our need to work with a (simple) mathematical model of the earth's shape with the on dulating nature of the earth's surface (i.e. its geoid). 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.

Local Datum

There are many local datums to choose from, some are old while others are more recently defined. The choice of datum is largely driven by the location of interest. For example, when working in the US, a popular local datum to choose from is the North American Datum of 1927 (or NAD27 for short). NAD27 works well for the US but it's not well suited for other parts of the world. For example, a far better local datum for Europe is the European Datum of 1950 (ED50 for short). Examples of common local datums are shown in the following table:

Local datum	Acronym	Best for	Comment
North American Datum of 1927	NAD27	Continental US	This is an old datum but still prevalent because of the wide use of older maps.
European Datum of 1950	ED50	Western Europe	Developed after World War II and still quite popular today. Not used in the UK.
World Geodetic System 1972	WGS72	Global	Developed by the Department of Defense.

Geocentric Datum

Many of the modern datums use a geocentric alignment. These include the popular World Geodetic Survey for 1984 (WGS84) and the North American Datums of 1983 (NAD83). Most of the popular geocentric datums use the WGS84 ellipsoid or the GRS80 ellipsoid. These ellipsoids' semi-major and semi-minor axes are nearly identical: 6,378,137 meters and 6,356,752 meters respectively. Examples of popular geocentric datums are shown in the following table:

Geocentric datum	Acronym	Best for	Comment
North American Datum of 1983	NAD83	Continental US	This is one of the most popular modern datums for the contiguous US.

European Terrestrial Reference System 1989	ETRS89	Western Europe	This is the most popular modern datum for much of Europe.
World Geodetic System 1984	WGS84	Global	Developed by the Department of Defense.

1.4. GIS AS AN INFORMATION SYSTEM

As Definition of GIS indicates GIS as a specialized information system stresses "spatially distributed features (points, lines, areas), activities (physical and human-invoked), and events (time).

GIS as an approach to Geographic Information Science

1) research on GIS (algorithms, analytical methods, visualization tools, user interfaces, human-computer-human interaction)
2) research with GIS: GIS as a tool used by many substantive disciplines in their own ways (anthropology, archeology, forestry, geology, engineering, business and management sciences)

1.5. DEFINING GIS

A GIS is a system of hardware, software, data, people, organizations, and institutional arrangements for collecting, storing, analyzing and disseminating information about areas of the earth. It is also defined as an information system designed to work with data referenced by spatial / geographical coordinates. In other words, GIS is both a database system with specific capabilities for spatially referenced data as well as a set of operations for working with the data. A Geographic Information System is a computer based system which is used to digitally reproduce and analyze the feature present on earth surface and the events that take place on it. In the light of the fact that almost 70% of the data has geographical reference as its denominator, it becomes imperative to underline the importance of a system which can represent the given data geographically.

1.6. HISTORY OF GIS

The idea of portraying different layers of data on a series of base maps, and relating things geographically, has been around much older than computers invention. Thousands years ago, the early man used to draw pictures of the animals they hunted on the walls of caves. These animal drawings are track lines and tallies thought to depict migration routes. While simplistic in comparison to modern technologies, these early records mimic the two-element structure of modern geographic information systems, an image associated with attribute information.

Possibly the earliest use of the geographic method, in 1854 John Snow depicted a cholera outbreak in London using points to represent the locations of some individual cases. His study of the distribution of cholera led to the source of the disease, a contaminated water pump within the heart of the cholera outbreak. While the basic elements of topology and theme existed previously in cartography, the John Snow map was unique, using cartographic methods, not only to depict but also to analyze, clusters of geographically dependent phenomena for the first time.

1.12 GIS

The early 20th century saw the development of "photo lithography" where maps were separated into layers. Computer hardware development spurred by nuclear weapon research led to general-purpose computer "mapping" applications by the early 1960s. In the year 1962, the world's first true operational GIS was developed by the federal Department of Forestry and Rural Development in Ottawa, Canada by Dr. Roger Tomlinson. It was called the "Canada Geographic Information System" (CGIS) and was used to store, analyze, and manipulate data collected for the Canada Land Inventory (CLI). It is an initiative to determine the land capability for rural Canada by mapping information about soils, agriculture, recreation, wildlife, forestry, and land use at a scale of 1:50,000.

CGIS was the world's first "system" and was an improvement over "mapping" applications as it provided capabilities for overlay, measurement, and digitizing or scanning. It supported a national coordinate system that spanned the continent, coded lines as "arcs" having a true embedded topology, and it stored the attribute and location specific information in a separate files. Dr. Tomlinson is known as the "father of GIS," for his use of overlays in promoting the spatial analysis of convergent geographic data.

In 1964, Howard T Fisher formed the Laboratory for Computer Graphics and Spatial Analysis at the Harvard Graduate School of Design, where a number of important theoretical concepts in spatial data handling were developed. This lab had major influence on the development of GIS until early 1980s. Many pioneers of newer GIS "grew up" at the Harvard lab and had distributed seminal software code and systems, such as 'SYMAP', 'GRID', and 'ODYSSEY'.

By the early 1980s, M&S Computing (later Intergraph), Environmental Systems Research Institute (ESRI) and CARIS emerged as commercial vendors of GIS software, successfully incorporating many of the CGIS features, combining the first generation approach to separation of spatial and attribute information with a second generation approach to organizing attribute data into database structures. More functions for user interaction were developed mainly in a graphical way by a user friendly interface (Graphical User Interface), which gave to the user the ability to sort, select, extract, reclassify, reproject and display data on the basis of complex geographical, topological and statistical criteria. During the same time, the development of a public domain GIS begun by the U.S. Army Corp of Engineering Research Laboratory (USA-CERL) in Champaign, Illinois, a branch of the U.S. Army Corps of Engineers to meet the need of the United States military for software for land management and environmental planning.

In the years 1980s and 1990s industry growth were spurred on by the growing use of GIS on Unix workstations and the personal computers. By the end of the 20th century, the rapid growth in various systems had been consolidated and standardized on relatively few platforms and users were beginning to export the concept of viewing GIS4 data over the Internet, requiring uniform data format and transfer standards. More recently, there is a growing number of free, open source GIS packages, which run on a range of operating systems and can be customized to perform specific tasks. As computing power increased and hardware prices slashed down, the GIS became a viable technology for state development planning. It has become a real Management Information System (MIS), and thus able to support decision making processes.

1.7. COMPONENTS OF A GIS

A working GIS integrates five key components: -

i) Hardware ii) Software iii) Data iv) People v) Methods

GIS enables the user to input, manage, manipulate, analyze, and display geographically referenced data using a computerized system. To perform various operations with GIS, the components of GIS such as software, hardware, data, people and methods are essential.

1.7.1. Software

GIS software provides the functions and tools needed to store, analyze, and display geographic information. Key software components are (a) a database management system (DBMS) (b) tools for the input and manipulation of geographic information (c) tools that support geographic query, analysis, and visualization (d) a graphical user interface (GUI) for easy access to tools. GIS software are either commercial software or software developed on Open Source domain, which are available for free. However, the commercial software is copyright protected, can be expensive and is available in terms number of licensees.

Currently available commercial GIS software includes Arc/Info, Intergraph, MapInfo, Gram++ etc. Out of these Arc/Info is the most popular software package. And, the open source software are AMS/MARS etc.

1.7.2. Hardware

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 stand-alone or networked configurations. Minimum configuration required to Arc/Info Desktop 9.0 GIS application is as follows:

Product: ArcInfo Desktop 9.0

Platform: PC-Intel

Operating System: Windows XP Professional Edition, Home Edition

Service Packs/Patches: SP 1

SP2 (refer to Limitations)

Shipping/Release Date: May 10, 2004

Hardware Requirements

CPU Speed: 800 MHz minimum, 1.0 GHz recommended or higher

Processor: Pentium or higher

Memory/RAM: 256 MB minumum, 512 MB recommended or higher

Display Properties: Greater than 256 color depth

Swap Space: 300 MB minimum

Disk Space: Typical 605 MB NTFS, Complete 695 MB FAT32 + 50 MB forinstallation

Browser: Internet Explorer 6.0 Requirement:

(Some features of ArcInfo Desktop 9.0 require a minimum installation of Microsoft Internet Explorer Version 6.0.)

1.7.3. Data

The most important component of a GIS is the data. Geographic data or Spatial data and related tabular data can be collected in-house or bought from a commercial data provider. Spatial data can be in the form of a map/remotely-sensed data such as satellite imagery and aerial photography. These data forms must be properly geo-referenced (latitude/longitude). Tabular data can be in the form attribute data that is in some way related to spatial data. Most GIS software comes with inbuilt Database Management Systems (DBMS) to create and maintain a database to help organize and manage data.

1.7.4. Users

GIS technology is of limited value without the users who manage the system and to develop plans for applying it. GIS users range from technical specialists who design and maintain the system to those who use it to help them do their everyday work.

These users are largely interested in the results of the analyses and may have no interest or knowledge of the methods of analysis. The user-friendly interface of the GIS software allows the nontechnical users to have easy access to GIS analytical capabilities without needing to know detailed software commands. A simple User Interface (UI) can consist of menus and pull-down graphic windows so that the user can perform required analysis with a few key presses without needing to learn specific commands in detail.

1.7.5. People

GIS technology has limited value without the people who manage and develop plans for applying it to real world problems. GIS user 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 specialist's vs. end users is often critical to the proper implementation of GIS technology. This is what called 'brain ware' which is equally important as the Hardware and software. Brain ware refers to the purpose and objectives, and provides the reason and justification, for using GIS.

This component of GIS includes all those individuals (such as the programmer, database manager, GIS researcher etc.) who are making the GIS work, and also the individuals who are at the user end using the GIS services, applications and tools.

1.7.6. 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.

1.8. PROPRIETARY AND OPEN SOURCE GIS SOFTWARE

1.8.1. AGISMap

AGIS for Windows is a mapping and simple GIS package specifically designed to be easy to use, and distributed as shareware via the world wide web.

Platforms: *Windows*

1.8.2. Autodesk

Autodesk has a series of software applications designed to meet GIS needs in a variety of areas that interfaces with their CAD software.

Platforms: *Windows*

1.8.3. Bentley Systems, Inc.

Bentley provides software for the "Design, construction, and operation of the world's infrastructure". The company's software serves the geospatial, building, plant, and civil vertical markets in the areas of mapping, architecture, engineering, construction (AEC) and operations. Bentley offers a wide range of products for surveying, GPS, photogrammetry, remote sensing, imaging, conversion, mapping, cartography, and other geospatial applications built on MicroStation Products: Bentley Map – Desktop GIS, Bentley Cadastre – Desktop land management GIS, Bentley Descartes – Desktop image editing, analysis, and processing, Bentley Geo Web Publisher – GIS web publishing and viewing, Bentley PowerMap Field – Field-enabled GIS and redlining.

Platforms: *Windows*

1.8.4. Cartographica

Commercial software package for Mac OS featuring: Support for a huge number of import formats, including popular Raster formats, manual editing and georeferencing, automatic geocoding, integration with online mapping, output to large-format printers. Thirty-day demo available for download.

Platforms: *Macintosh*

1.8.5. DeLorme

DeLorme is the producers of XMap, a GIS application "with 80% of the functionality found in a traditional GIS at 15% of the cost". Performs such functions as geocoding, image rectification, 3D visualization and coordinate transformation.

Platforms: *Windows*

1.8.6. Esri

Environmental Systems Research Institue has been creating GIS software for over 30 years. Recognized as the leader in GIS software, it's been estimated that about seventy percent of

GIS users use Esri products. Esri overhauled their software packages into an interoperable model called ArcGIS (the desktop GIS is referred to as ArcMap). In addition, Esri has developed plug-ins called extensions which add to the functionality of ArcGIS. Demo and light versions of Esri software are available for downloading. You can also find free data to use with Esri products.

Platforms: *Windows*

Further Resources: *ArcGIS, ArcView 3.x (no longer in production)*

1.8.7. Intergraph

Intergraph makes several GIS applications. Most of the GIS packages are designed with an Open GIS in mind and therefore can work with a variety of other GIS software formats. Intergraph has developed products that help merge GIS with information technology (IT) and business process improvement tools. Intergraph offers the Geo-Media family of solutions and Modular GIS Environment MGE Suite of mapping and GIS applications.

Using an open architecture, the Geo-Media product suite integrates geospatial information throughout the enterprise and provides the tools needed to develop business-to-business and custom client applications using industry standard development tools. Geo-Media offers uninhibited access to all geospatial data formats without the need for data translations. Currently in Version 4.0 the Geo-Media family is made up of Geo-Media, Geo-Media Professional, Geo Media Web Map, and Geo Media Web Enterprise.

- GeoMedia is the universal information integrator, serving as a visualization and analysis tool and as an open platform for custom GIS solution development.
- GeoMedia Professional is a product specifically designed to collect and manage spatial data using standard databases.
- GeoMedia WebMap is a Web-based map visualization tool with real-time links to one or more GIS data warehouses.
- GeoMedia WebEnterprise creates dynamic, custom web-mapping applications that can analyze and manipulate geographic data.
- In addition to these products, Intergraph offers MFworks for GeoMedia which provides users of grid-based software the power of visualization, mapping, and analysis. Intergraph also offers SMMS for GeoMedia which is a desktop tool for geographic metadata creation and geographic data management.

The Modular GIS Environment (MGE) product suite provides production-ready capabilities for automating, managing, analyzing, and presenting GIS data, and is completely interoperable with GeoMedia.

1.8.8. Manifold

Manifold System provides comprehensive, professional grade GIS software for $245 that includes a very wide array of features. Manifold imports data from over 80 different GIS formats, including all formats used by Federal government sites for free Internet downloads, and Manifold allows seamless, simultaneous work with vector drawings, raster images, terrain elevations and raster data sets either as 2D displays or 3D terrain visualizations. Manifold includes exceptional DBMS capabilities, full development facilities and includes a built-in Internet Map Server for fast and easy publication of GIS projects to the web without programming. Options include US Street address geocoding and the Enterprise Edition, for centralized geospatial data storage on enterprise servers that can be used by many GIS operators at once. Manifold was the first GIS to attain "designed for XP" status with Microsoft and the Manifold Internet Map Server works perfectly within ASP.NET servers. For info, see http://www.manifold.net/professional.

Platforms: *Windows*

1.8.9. Ortelius

Ortelius is a "map illustration" software package that adds one more selection to the very limited mapping software options out there for Macintosh users. A free trial download is available.

Platforms: *Macintosh*

1.8.10. MapInfo – Pitney Bowes Business Insight (PBBI)

PBBI's flagship software is MapInfo, a suite of GIS software. MapInfo Professional is their leading GIS product containing the most advanced analytical tools. MapInfo also offers plug-ins called add-ons to enhance the functionality of MapInfo Professional. For the development side, MapInfo offers Map-X. Through an Active X component, developers can embed mapping applications into other applications such as Excel. Although it can be used for a variety of analysis, the makers of MapInfo market the software more towards the business sector. Demo versions are available for downloading for some of MapInfo's products.

Platforms: *Windows OS*

Further Resources: *MapInfo, MapBasic, MapInfo Tutorials*

1.8.11 Maptitude

The Maptitude Mapping Software is a full-featured mapping package for Windows. Designed for ease-of-use, data visualization and geographic analysis, Maptitude comes with comprehensive nationwide and worldwide maps, including complete US street maps, and Census tract and ZIP Code boundaries and demographics. Caliper also produces TransCAD for transportation and logistics. TransCAD is used for solving key analytical problems in transportation planning, management, and operations. TransCAD is used extensively for transportation database development and maintenance, demand forecasting, operations management, and vehicle routing and scheduling.

Platforms: *Windows OS*

1.8.12. MyWorld

My World GIS is a full-featured GIS designed for educational use. My World provides a carefully selected subset of the features of a professional GIS environment. These features include multiple geographic projections, table and map views of data, distance-measurement tools, buffering and query operations, customizable map display. They have been selected to provide the greatest value to students without overwhelming them with complexity The features are accessed through a supportive interface designed with the needs of students and teachers in mind. My

World can import data from the industry-standard shapefile format, as well as from tab and comma-delimited text files.

Platforms: *Windows, Macintosh, Linux, Solaris*

1.8.13 Supergeo Technologies

SuperGIS Desktop is a full-featured GIS platform for Windows OS. It allows users to edit, visualize, manage and analyze geospatial data both in vector and raster, including OGC formats and various geodatabases, such as MSSQL, Oracle spatial, and PostGIS. Capabilities can be powered up via extensions such as Network Analyst, 3D Analyst, Spatial Analyst, Biodiversity Analyst, etc., which enables users to conduct complicated analyses and make smarter decisions. By combing Mobile GIS and Server GIS also from Supergeo, you can have a total geospatial solution from data collection in field to data publishing online with reasonable price. The free trial is available on Supergeo's website. Various product resources and friendly technical support are also provided by Supergeo team.

Platforms: *Windows OS*

1.8.14. TatukGIS Editor

Professional, general-purpose desktop GIS mapping and data editing application with built-in scripting environment for customization and feature extensions. Natively supports most GIS/CAD raster/vector/SQL layer data formats, including enterprise spatial geodatabases (such as Oracle Spatial, PostGIS …) State-of-the-art coordinate system support with nearly 5,000 pre-defined coordinate systems and on-the-fly raster/vector layer reprojection, 3D mapping, raster & vector layer rectification, and compatibility with leading database products. Data grid with advanced query and selection tools.

Platforms: *Windows OS*

1.8.15. Terrain Tools

Terrain Tools, produced by Softree, is a software package for surveying and mapping. It is ideal for the forester, geologist, surveyor or resource scientist who is not a GIS specialist, but who needs to quickly produce working maps and site plans.

Platforms: *Windows*

1.8.16. TerrSet Geospatial Monitoring and Modeling System

Developed by Clark Labs, TerrSet is an integrated geospatial software for monitoring and modeling the earth system. The TerrSet software incorporates the IDRISI GISt Analysis and IDRISI Image Processing tools along with a constellation of modeling environments to analyze land change, image time series, ecosystem services, habitat and biodiversity, climate impacts, and REDD. Learn more at www.clarklabs.org.

Platforms: *Windows*

1.8.17. TNT Products

Created by Microimages, The TNT Products is a suite of GIS applications for fully integrated GIS, image processing, CAD, TIN, desktop cartography, and geospatial database management.

Platforms: *Windows, UNIX, Macintosh*

1.9. OPEN SOURCE GIS SOFTWARE

(i) From the UNH Cooperative Extension course "***GIS on Pennies a Day***," this is a list of freely available, open source GIS software. You can download and use the software.

(ii) QGIS and DIVA-GIS are more widely used than the others. Both have MAC versions available for downloading.

1.9.1. Quantum GIS (QGIS)

Quantum GIS (QGIS) is a user friendly Open Source Geographic Information System (GIS) licensed under the GNU General Public License. QGIS is an official project of the Open Source Geospatial Foundation (OSGeo). It runs on Linux, Unix, Mac OSX, and Windows and supports numerous vector, raster, and database formats and functionalities.

1.9.2. DIVA-GIS

DIVA-GIS are a free computer program for mapping and geographic data analysis (a geographic information system (GIS).

1.9.3. Free QGIS Courses from Geo-Academy

Geo-Academy is an independent consortium promoting the use of open source geospatial software. These free courses are based on years of development of their fee-based coursework.

These are other open source GIS software available via the web.

1.9.4. GeoDa

GeoDa is a free GIS software program primarily used to introduce new users into spatial data analysis. It’s main **functionality is data exploration in statistics**.

One of the nicest things about it is how it comes with sample data for you to give a test-drive. From simple box-plots all the way to regression statistics, GeoDa has **complete arsenal of statistics** to do nearly anything spatially.

It’s user base is strong. For example, Harvard, MIT and Cornell universities have embraced this free GIS software to serve as a gentle introduction to spatial analysis for non-GIS users. From economic development to health and real estate, it’s been used as an exciting analytical in labs as well.

From the Center for Spatial Data Science at the University of Chicago, software for geospatial analysis, geovisualization and other techniques.

1.9.5. gvSIG

In 2004, the gvSIG project emerged as a free, open source GIS software option in Spain. We illustrate in this gvSIG guide and review why we like it SO much:

gvSIG really outperforms QGIS 2 for 3D. It really is the best 3D visualization available in open source GIS.

The NavTable is agile in that it allows you to see records one-by-one vertically.

The CAD tools are impressive on gvSIG. Thanks to the OpenCAD Tools, you can trace geometries, edit vertices, snap and split lines and polygons. If you need GIS on your mobile phone, gvSIG Mobile is perfect for field work because of its interface and GPS tools.

1.9.6. MapWindow GIS

An open source GIS software that is being used in History. For more information on its use in History, a manual and exercises.

1.9.7. SAGA (System for Automated Geoscientific Analyses)

SAGA is a free, open source GIS software used for raster-based analyses and the Earth Sciences.

1.9.8. uDig

uDIG is an acronym to help get a better understanding what this Free GIS software is all about.

- u stands for user-friendly interface
- D stands for desktop (Windows, Mac or Linux). You can run uDIG on a Mac.
- I stand for internet oriented consuming standard (WMS, WFS or WPS)
- G stands for GIS-ready for complex analytical capabilities.

When you start digging into uDig, it’s a nice open source GIS software option for basic mapping. uDig’s Mapnik lets you import basemaps with the same tune as ArcGIS

Specifically, its easy-to-use, the catalog, symbology and Mac OS functionality are some of the strong points. But it has limited tools and the bugs bog it down to really utilize it as a truly complete free GIS software package.

1.9.9 Environmental Benefits Mapping & Analysis Program (BenMAP)

BenMAP-CE is a open-source software program that calculates the number and economic value of air pollution-related deaths and illnesses. It incorporates a database that includes many of the concentration-response relationships, population files, and health and economic data needed to quantify these impacts. This is the Community Edition.

1.10. DATA TYPES

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.

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:

The attributes refer to the properties of spatial entities. They are often referred to as non-spatial data since they do not in themselves represent location information. This type of 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.

Spatial Data:

Geographic position refers to the fact that each feature has a location that must be specified in a unique way. To specify the position in an absolute way a coordinate system is used. For small areas, the simplest coordinate system is the regular square grid. For larger areas, certain approved cartographic projections are commonly used. Internationally there are many different coordinate systems in use. This locational information is provided in maps by using Points, Lines and Polygons. These geometric descriptions are the basic data elements of a map. Thus spatial data describes the absolute and relative location of geographic features.

The coordinate location of a forest would be spatial data, while the characteristics of that forest, e.g. cover group, dominant species, crown closure, height, etc., would be attribute data. Other data types, in particular image and multimedia data, have become 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.

1.10.1. Spatial Data

- Describes the absolute and relative location of geographic features.
- Represents spatial data, which has a physical dimension on earth.
- Spatial data consist of digital representations of discrete (spatial) objects. The features are shown on a map, e.g. lakes, buildings and contours can be thought of as discrete objects.
- Thus the contents of a map can be captured in a database by turning map features into database objects (entities).
- Components of spatial data.

Location: The spatial mode of information is generally called location.

Spatial relationship: The connections between spatial objects are described as spatial relationships (e.g. A contains B; A is adjacent to B, A is North of B, etc.).

Attributes: Attributes capture the thematic mode by defining different characteristics of objects. Spatial features in the real world are reduced in the form of point, line, area and surface. GIS will store the data either in tabular form, geographical map, digital map or remotely sensed map.

1.10.2. Non-Spatial Data

The non-spatial data or the attribute data, on the other hand, describes the characteristics of the spatial features. These characteristics can be quantitative or qualitative, also called attribute data.

- Describes the characteristics of the spatial features.
- It holds the characteristics of the spatial features and the descriptive information about the geographic features.
- Represented using colors, textures and symbols.
- Eg: Coordinate location of a sanctuary would be spatial data, while the characteristic's like the cover group, dominant species, nature of vegetation would be attribute data.
- These attributes are given in an organized from in a single table or multiple tables.

1.10.3. Spatial Data Models

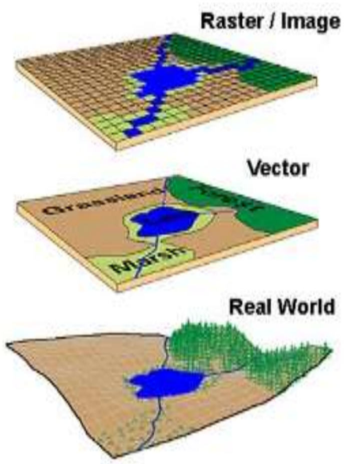
Traditionally spatial data has been stored and presented in the form of a map. Three basic types of spatial data models have evolved for storing geographic data digitally. These are referred to as:

(i) Vector data

(ii) Raster data

(iii) Image data

The following diagram reflects the two primary spatial data encoding techniques. These are vector and raster. Image data utilizes techniques very similar to raster data, however typically lacks the internal formats required for analysis and modeling of the data. Images reflect pictures or photographs of the landscape.



1.10.4. Attribute Data Models

A separate data model is used to store and maintain attribute data for GIS software. These data models may exist internally within the GIS software, or may be reflected in external commercial Database Management Software (DBMS). A variety of different data models exist for the storage and management of attribute data. The most common are:

- Tabular
- Hierarchical
- Network
- Relational
- Object oriented

Tabular Model

The simple tabular model stores attribute data as sequential data files with fixed formats (or comma delimited for ASCII data), for the location of attribute values in a predefined record structure. This type of data model is outdated in the GIS arena. It lacks any method of checking data integrity, as well as being inefficient with respect to data storage, e.g. limited indexing capability for attributes or records, etc.

Hierarchical Model

The hierarchical database organizes data in a *tree* structure. Data is structured downward in a *hierarchy* of tables. Any level in the hierarchy can have unlimited *children*, but any *child* can have only one *parent*. Hierarchical DBMS have not gained any noticeable acceptance for use within GIS.

They are oriented for data sets that are very stable, where primary relationships among the data change infrequently or never at all. Also, the limitation on the number of parents that an element may have is not always conducive to actual geographic phenomenon.

Network Model

The network database organizes data in a network or *plex* structure. Any column in a plex structure can be linked to any other. Like a tree structure, a plex structure can be described in terms of *parents* and *children*.

This model allows for children to have more than one parent. Network DBMS have not found much more acceptance in GIS than the hierarchical DBMS. They have the same flexibility limitations as hierarchical databases; however, the more powerful structure for representing data relationships allows a more realistic modeling of geographic phenomenon. However, network databases tend to become overly complex too easily. In this regard it is easy to lose control and understanding of the relationships between elements.

Relational Model

The relational database organizes data in *tables*. Each table, is identified by a unique table name, and is organized by *rows* and *columns*. Each column within a table also has a unique name.

Columns store the values for a specific attribute, e.g. cover group, tree height. Rows represent one record in the table.

In a GIS each row is usually linked to a separate spatial feature, e.g. a forestry stand. Accordingly, each row would be comprised of several columns, each column containing a specific value for that geographic feature.

Object-Oriented Model

The object-oriented database model manages data through *objects*. An object is a collection of data elements and operations that together are considered a single entity. The object-oriented database is a relatively new model.

This approach has the attraction that querying is very natural, as features can be bundled together with attributes at the database administrator's discretion. To date, only a few GIS packages are promoting the use of this attribute data model.

However, initial impressions indicate that this approach may hold many operational benefits with respect to geographic data processing. Fulfillment of this promise with a commercial GIS product remains to be seen.

1.11. TYPES OF ATTRIBUTE DATA

Attribute data can be store as one of five different field types in a table or database: character, integer, floating, date, and BLOB.

1.11.1. Character Data

The character property (or string) is for text based values such as the name of a street or descriptive values such as the condition of a street. Character attribute data is stored as a series of alphanumeric symbols.

Aside from descriptors, character fields can contain other attribute values such as categories and ranks. For example, a character field may contain the categories for a street: avenue, boulevard, lane, or highway. A character field could also contain the rank, which is a relative ordering of features. For example, a ranking of the traffic load of the street with “1” being the street with the highest traffic.

Character data can be sorted in ascending (A to Z) and descending (Z to A) order. Since numbers are considered text in this field, those numbers will be sorted alphabetically which means that a number sequence of 1, 2, 9, 11, 13, 22 would be sorted in ascending order as 1, 11, 13, 2, 22, 9.

Because character data is not numeric, calculations (sum, average, median, etc.) can’t be performed on this type of field, even if the value stored in the field are numbers (to do that, the field type would need to be converted to a numeric field). Character fields can be summarized to produced counts (e.g. the number of features that have been categorized as “avenue”).

1.11.2. Numeric Data

Integer and floating are numerical values (see: the difference between floating and integer values). Within the integer type, the is a further division between short and long integer values. As would be expected, short integers store numeric values without fractional values for a shorter range than long integers. Floating point attribute values store numeric values with fractional values. Therefore, floating point values are for numeric values with decimal points (i.e numbers to the right of the decimal point as opposed to whole values).

Numeric values will be sorted in sequentially either in ascending (1 to 10) or descending (10 to 1) order. Numerical value fields can have operations performed such as calculating the sum or average value. Numerical field values can be a count (e.g. the total number of students at a school) or be a ratio (e.g. the percentage of students that are girls at a school).

1.11.3. Date/Time Data

Date fields contains date and time values.

1.11.4. BLOB Data

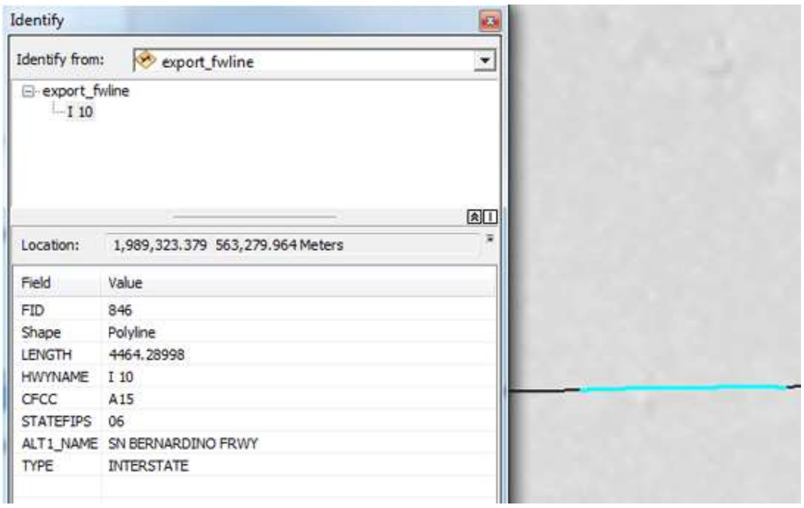


Fig.1.11. Attribute data for a road in gis.

BLOB stands for binary large object and this attribute type is used for storing information such images, multimedia, or bits of code in a field. This field stores object linking and embedding (OLE) which are objects created in other applications such as images and multimedia and linked from the BLOB field Attribute data for a road in gis.

1.12. SCALE OF MEASUREMENT/LEVEL OF MEASUREMENTS

Types of Attribute Data:

Types used in a GIS and in computer programming include character strings, integers, floating points or real numbers, dates and time intervals. Each field in an attribute table is defined with a data type, which applies to the domain of the field.

Another method is to define attribute data by measurement scale. The measurement scale concept groups attribute data into nominal, ordinal, interval and ratio data.

Nominal Data:

Nominal data describe different kinds of different categories of data such as land use types or soil types.

Ordinal Data:

Ordinal data differentiate data by ranking relationship. For example-cities may be grouped into large, medium and small cities by population size.

Interval Data:

Interval data have known intervals between values such as temperature reading. For example- a temperature reading of 700 F is warmer than 600 F by 100 F.

Ratio Data:

Ratio data are the same as interval data except that ratio data are based on a meaningful or absolute zero value. Population densities are an example of ratio data, because a density of 0 is an absolute zero.

Measurement scale (or level) of attribute can be broadly divided into two categories. Some attribute are measured in a numerical scale (such as job accessibility) whereas others are not (such as world language)

Distribution of Indian tribes & language:

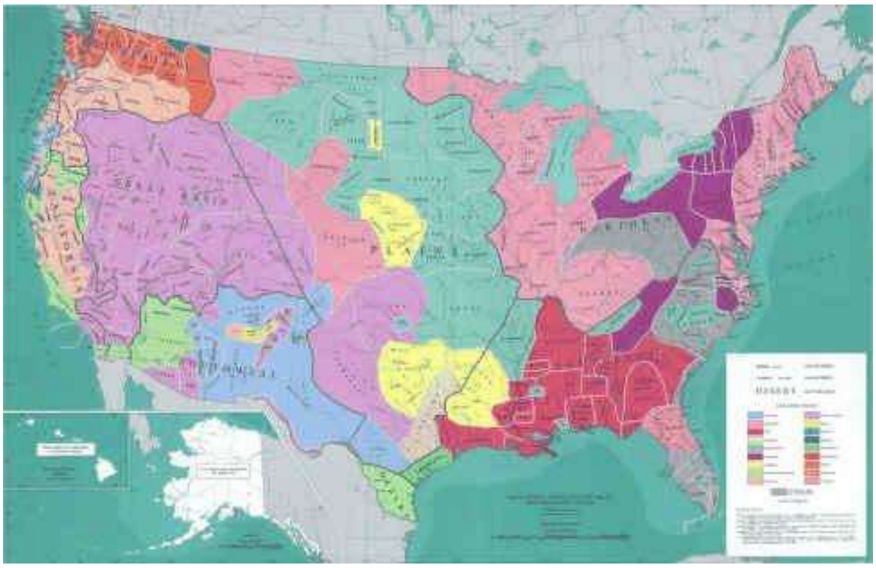


Fig.1.12. National Atlas of United States

This map can be fallen into qualitative thematic map because the measurement level of attribute portrayed (tribe/language) is nominal (not measured in number), and displays one or more particular themes.

Distribution of urban and rural population

Fig.1.13. National Atlas of United States

This map can be fallen into quantitative thematic map because the measurement level of attributed portrayed (population) is countable (measure in number), and displays one particular theme.

More specifically, measurement of attributes is organized into four levels: nominal, ordinal, interval, and ratio, listed in increasing order of sophistication of measurement

1.12.1. Nominal scaling

Only has a value either 0 or 1 (false or true)

Suppose we have two values say region A and region B; we can't determine if A > B or

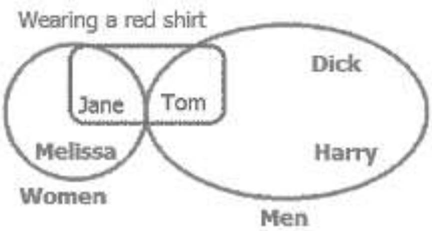
A < B, but we can determine if A=\=B or A = B.

e.g. agricultural region (corn regions, wheat regions, soy-bean regions)

Political party affiliation (Democrat, Republican, Independent)

Sex (male, female)

Response (yes, no)



1.12.2. Ordinal scaling

Value is arranged in a hierarchy of rank

Can determine if A > B or A < B, but can't determine how much they are different

e.g. social power (more, less)

agreement (strongly agree, strongly disagree)

Order of arrival of contestants in footrace

	Women's race	Men's race
First	Jane	Tom
Second	Melissa	Dick
Third	Leila	Harry

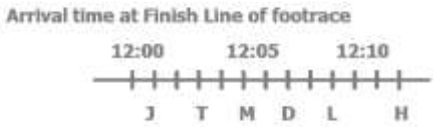
1.12.3. Interval scaling

Ranked

Know the distance between ranks

But it is not measured in an absolute scale; they are relative (has no natural origin)

e.g. Fahrenheit



1.12.4. Ratio scaling

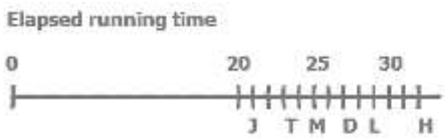
Ranked

Know the distance between ranks

It is measured in an absolute scale (has a natural origin)

e.g. weight, elevation

convey more information and permit more analytical treatment



Level of measurement prescribe the information required for an attribute reference system

Level of measurement	**Information required**
Nominal	Definitions of categories
Ordinal	Definitions of categories plus ordering
Interval	Unit of measure plus zero point
Ratio	Unit of measure