

CHAPTER TWO**2.0****GIS DATA MODELS**

- ❑ GIS does not store a map in any conventional sense. Instead GIS stores the data from which we can draw a desired view to suit a particular purpose.
- ❑ There are two types of data in GIS
- ❑ Spatial Data (Location of a particular feature)
- ❑ Attribute data (information about features. E.g. name of roads, forest type etc.)
- ❑ GIS integrates the two data types and allows the users to derive new data for planning.
- ❑ Spatial Models are important in that way in which information is represented, affects the type of analysis performed and types of graphic display that can be obtained.
- ❑ The two data models common in GIS are Vector data model and Raster data Model.

2.1 Spatial Information:

Spatial characteristics of information can be broadly distinguished between:

- a) Those that describe where things are? Using locations consisting of reference positions, spatial units and spatial relationships.
- b) Those that describe the form of phenomena using qualitative and quantitative description of shape and structure.
- c) Those that describe associations and interaction between different phenomena.

Basic Concepts:**1) All geographic data can be represented by three basic entities:**

- i) Point
- ii) Line
- iii) Area or Polygon plus a label saying what is it

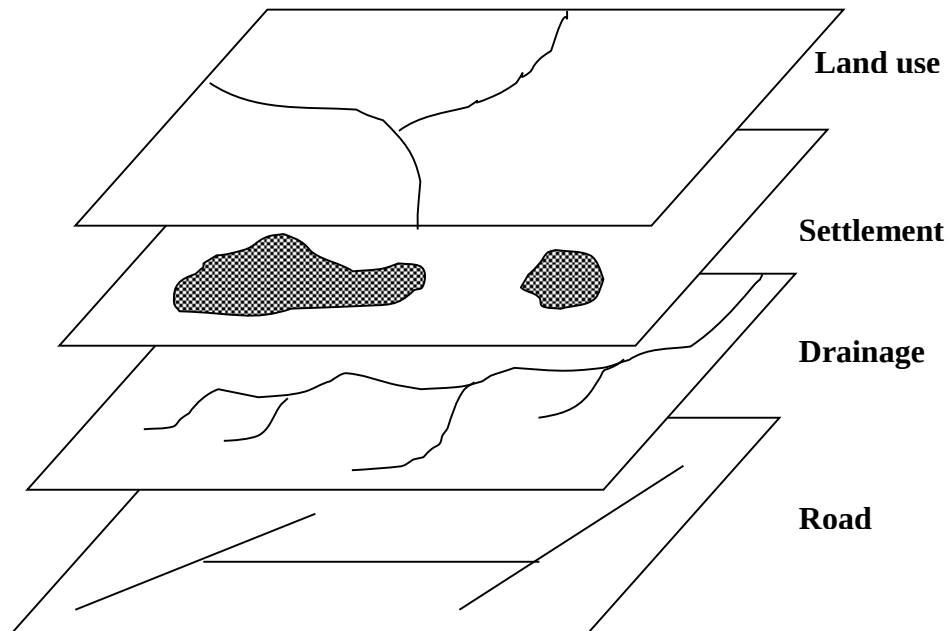
E.g.

- ❖ An oil well could be represented by a single point consisting of X, Y coordinates.
- ❖ Road – represented by a series of X, Y coordinates
- ❖ Forest – represented by a set of X, Y coordinates plus the label forest. The label could be actual name or a special symbol.

2) Layers and Coverages:

- ❑ GIS organize spatial data into layers or coverages

- ❑ Typical layers represent information belonging to particular class. E.g. Roads, Rivers, Vegetation types are different layers.
- ❑ All the layers or coverages pertaining to an area are referenced to a common projection system
- ❑ The layers can be combined with each other in various ways to create new layers that are functions of individual layers.



3) **Data Model:**

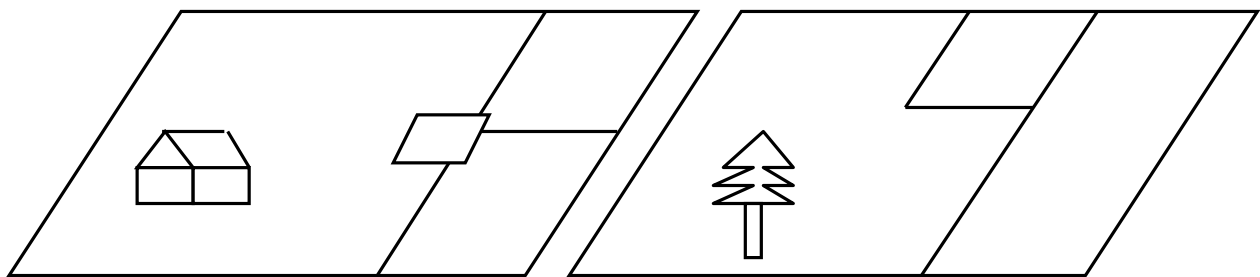
- ❑ In order to represent the spatial information and their attributes, a Data Model – a set of logical definitions or rules for characterizing the geographical data is adopted
- ❑ The Data Model represents the linkages between the real world domain of geographical data and the computer and GIS representation of these features.
- ❑ As a result, Data Model, not only helps in organizing the real world geographical features into a systematic storage/retrieval mechanism, but also helps in capturing user' s perception of these features.

2.2 **Conceptual Models of Spatial Information:**

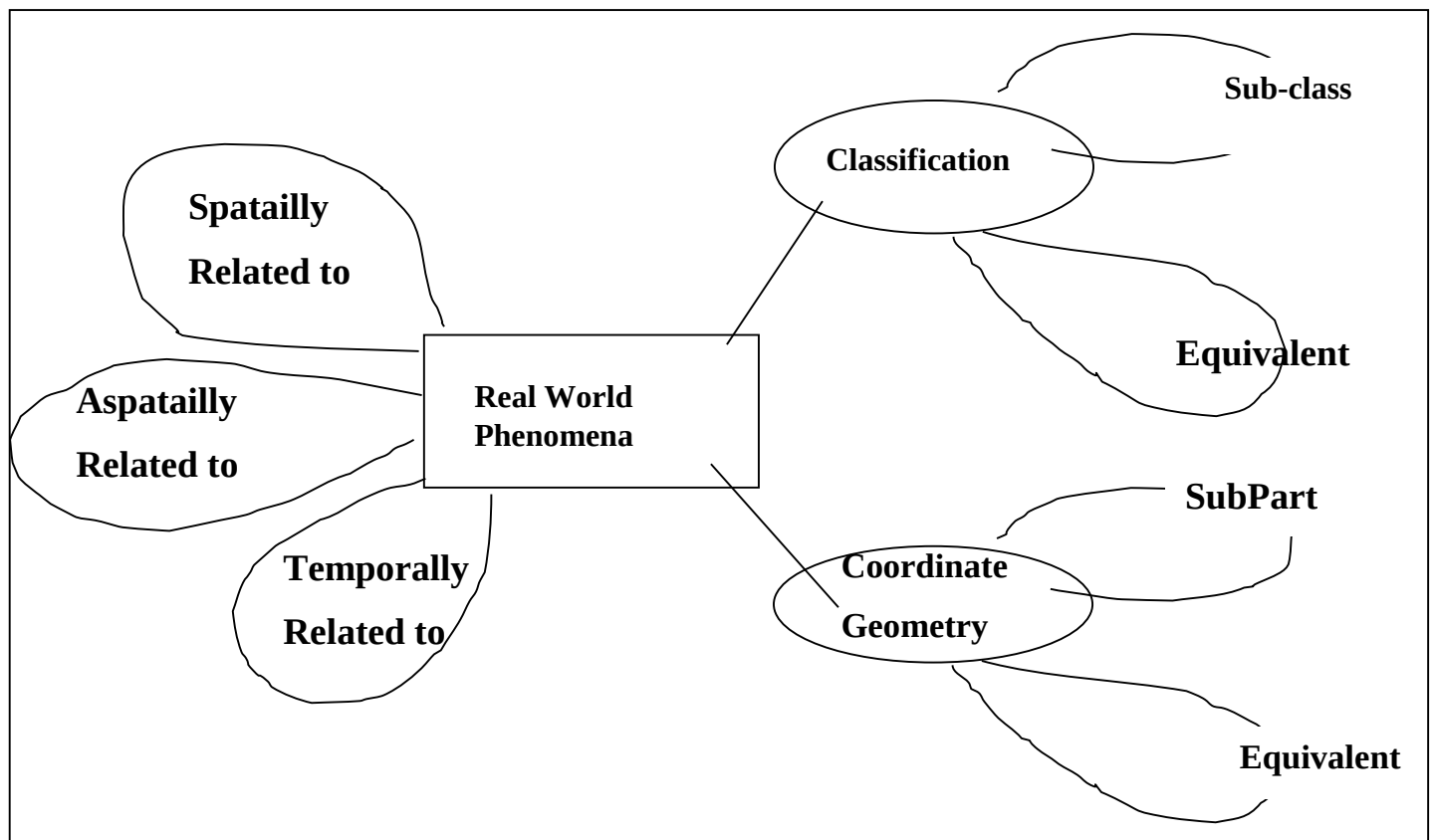
- ❑ There are different models – which have influenced the way in which data are organized and processed within GIS.
- ❑ They are based on Objects, Network and Fields.

2.2.1 Object Based Model:

- ❑ Object based spatial models emphasize individual phenomena that are to be studied in isolation or in terms of their relationship with other phenomena.
- ❑ Any phenomena, however bigger or small, may be designated as an object, provided that it can be separated conceptually from neighboring phenomena.
- ❑ Objects may be composed from other objects and they may have a specific relationship with other separate objects.
- ❑ An object-based view is appropriate to phenomena that have a well-defined boundary.



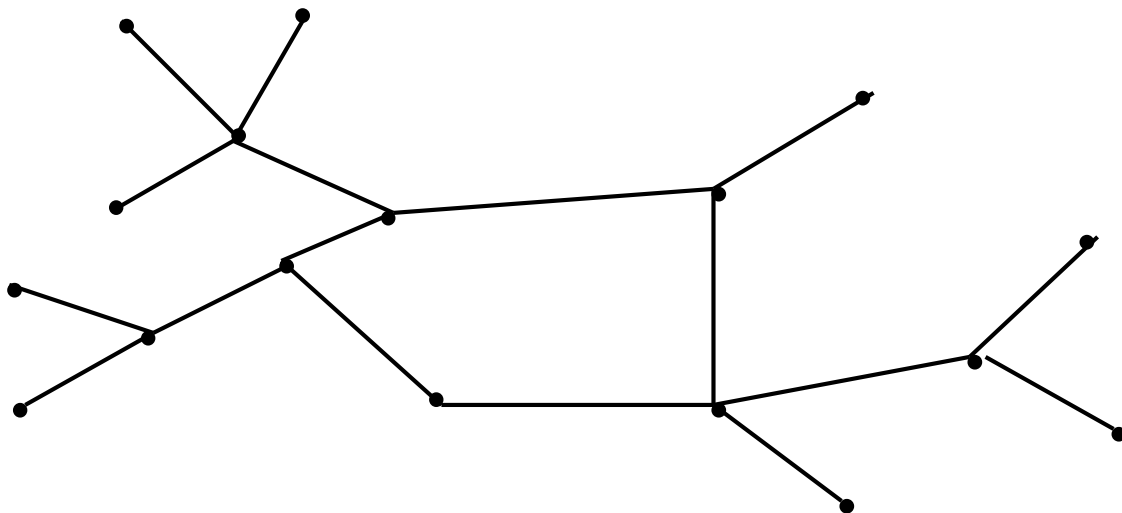
The object based conceptual view



- ❑ Hence it is suited to human made phenomena such as buildings, roads, administrative regions etc.
- ❑ Some natural phenomena such as lakes, rivers, islands and forests are often represented in object-based models because they need to be treated as discrete phenomena.

2.2.2. Network Model:

- ❑ Network based spatial module share some aspects of the object based module in that “ they often deal with discrete phenomenon” .
- ❑ But the essential characteristic is the need to consider interaction between multiple objects, often along discrete path or routes that connect them.
- ❑ The exact shape of phenomena is not of much importance. What is important is some measure of distance and impedance (interaction) between specified phenomena.



- ❑ E.g. Studies of Traffic on road, analysis of flow of water, flow of electricity etc.

2.2.3 Field Model:

- ❑ Field Based model is appropriate for modeling phenomena that are regarded as continuously variable across some region of space.
- ❑ E.g. concentration of pollutants in the air, temperature of ground surface, moisture level of soil, elevation of ground etc.

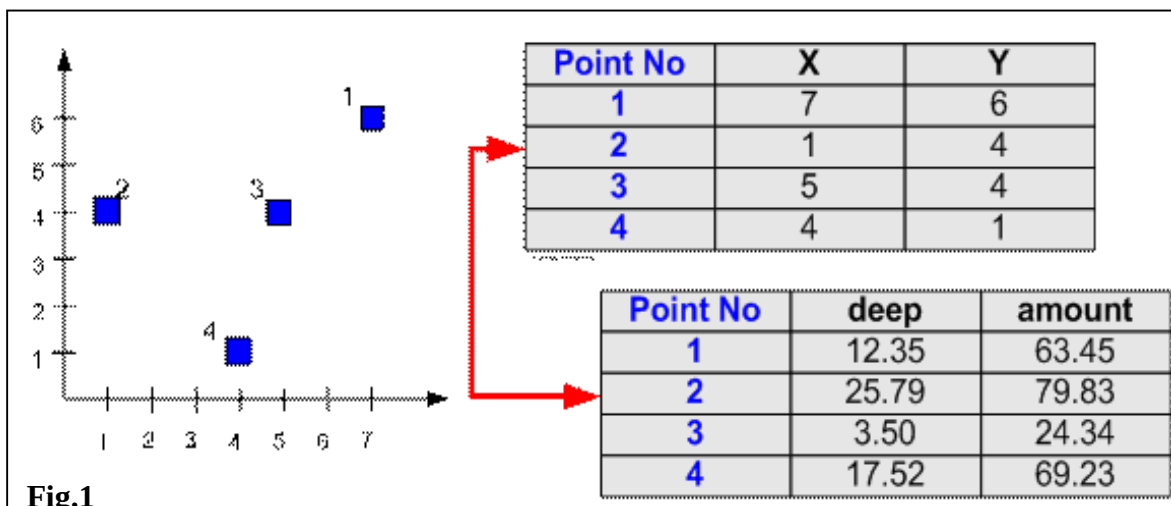
- ❑ Field model may represent either 2 or 3 dimensions depending upon the applications.

2.3 Vector Based Model:

- ❑ A vector based GIS is defined by the vectorial representation of its geographic data. According with the characteristics of this data model, geographic objects are explicitly represented and, within the spatial characteristics, the thematic aspects are associated.
- ❑ The vector representation of an object is an attempt to represent the object as exactly as possible.
- ❑ The geographical phenomena are represented by three basic entities along with their attributes.
Point – City – population, no. of school, no. of houses etc.
Line – Road – Type of road, road name etc.
Area – Land use – class, soil type etc.
- ❑ The coordinate space is assumed to be continuous, allowing all positions, lengths and dimensions to be defined precisely.
- ❑ The vector data structure represents each geographical feature by a set of coordinates.
- ❑ The basic thing is to define a 2D space where coordinates on the two axes represent features.

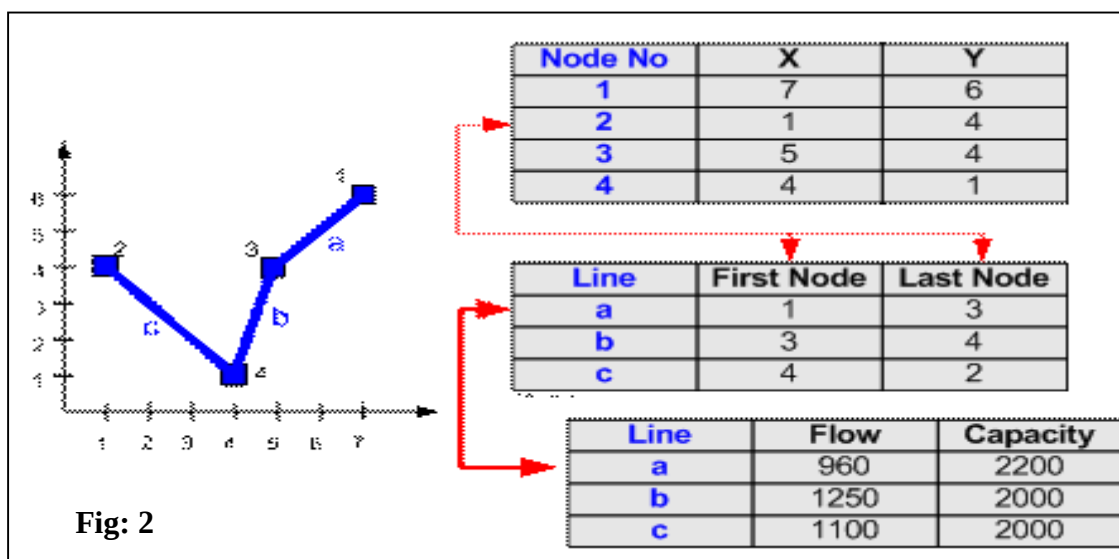
i. **Point Features:**

- ❑ A zero-dimensional abstraction of an object represented by a single X,Y co-ordinate. A point normally represents a geographic feature too small to be displayed as a line or area; for example, the location of a building location on a small-scale map, or the location of a service cover on a medium scale map.
- ❑ Besides the X, Y coordinate, other data must be stored to indicate what kind of point it is and other information associated with it. Fig. 1 shows a typical point data stored in GIS.



ii. Line Features:

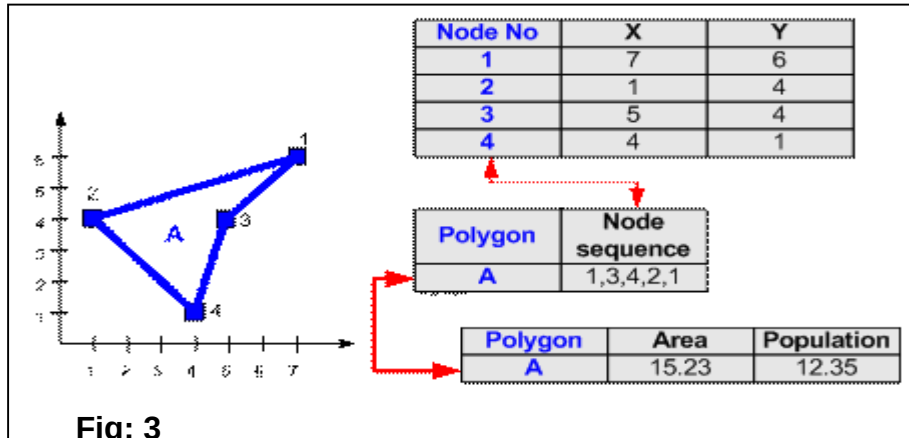
- ❑ A set of ordered co-ordinates that represent the shape of geographic features too narrow to be displayed as an area at the given scale (contours, street centerlines, or streams), or linear features with no area (county boundary lines). A line is synonymous with an arc.
- ❑ Simplest line required the storage of begin point and end point. (Two X, Y coordinates plus a possible record). An arc, a chain or string is a set of n X, Y coordinate pairs describing a continuous complex line.
- ❑ Shorter the line segment and larger the no. of X,Y coordinate pairs, the closer the chain will approximate a complex curve. Fig.2



iii. Area Features (Polygon Features):

- ❑ A feature used to represent areas. A polygon is defined by the lines that make up its boundary and a point inside its boundary for identification. Polygons have attributes that describe the geographic feature they represent.
- ❑ The boundary of area features separates the interior area from the exterior area.

- It may be isolated or connected. Fig:3



Topology:

Topology describes the spatial relationships between the objects and to carryout the measurements. (Length and Area).

e.g. spatial relationship of connectivity and adjacency are examples of topological relationships.

2.4 Raster Based Model:

- Raster based spatial models regard space as a tessellation (resembled mosaic) of cells, each of which is associated with a record of classification or identity of the phenomena that occupies it.
- The raster model represents the 2D location of phenomena as a matrix of grid cell.
- Each cell is known as pixel (Short form of Picture Element).
- Since the cells are of fixed size and location, raster tend to represent natural and human made objects in a blocky fashion.
- The information content in one cell depends upon the size of the cell. If the cells are sufficiently small, the information present in one cell will be more. This is called resolution of the image.
- The raster model or grid cell is relatively simple approach to data representation both conceptually and operationally, and hence has been popular since the earliest days of GIS development.
- The simplest raster data structure consists of an array of grid cells. A row and column number references each grid cell and it contains a number representing the type or value of the attribute being mapped. Fig: 4 a, 4b and 4c explain the raster model, raster representation of location and raster resolution respectively.
- In raster structure a single cell represents a point. A line by a number of neighboring cells string out in a given direction and area by agglomeration (mass) of neighboring cells. Fig: 4d show the raster representation of discrete features, Point, Line and Area.

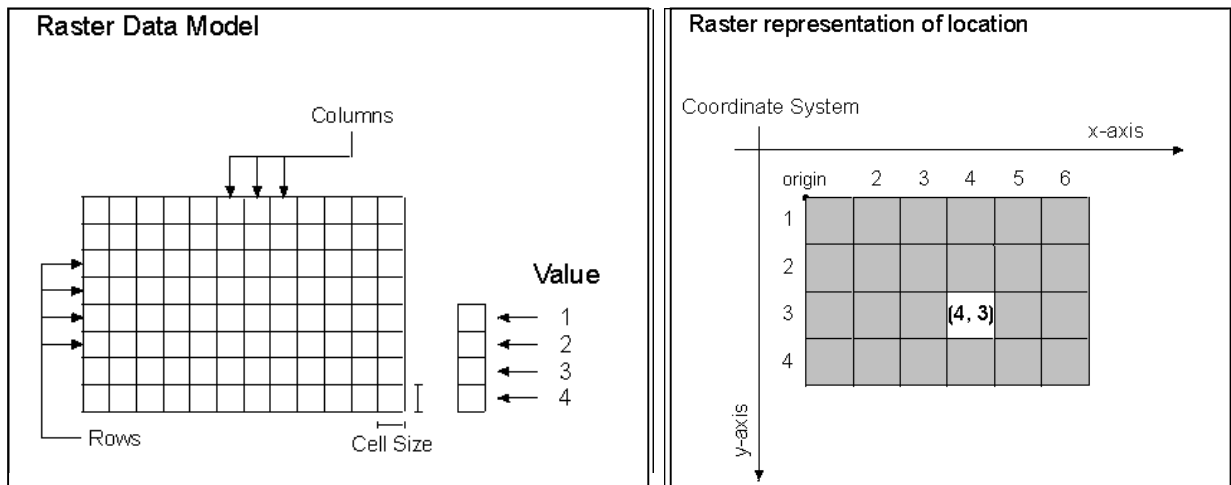


Fig: 4 a

Fig: 4 b

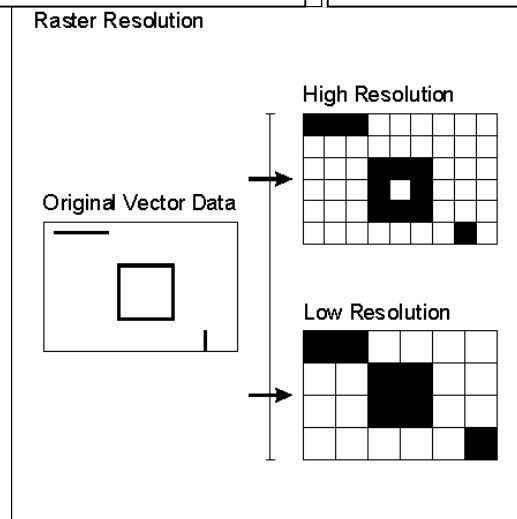


Fig: 4 c

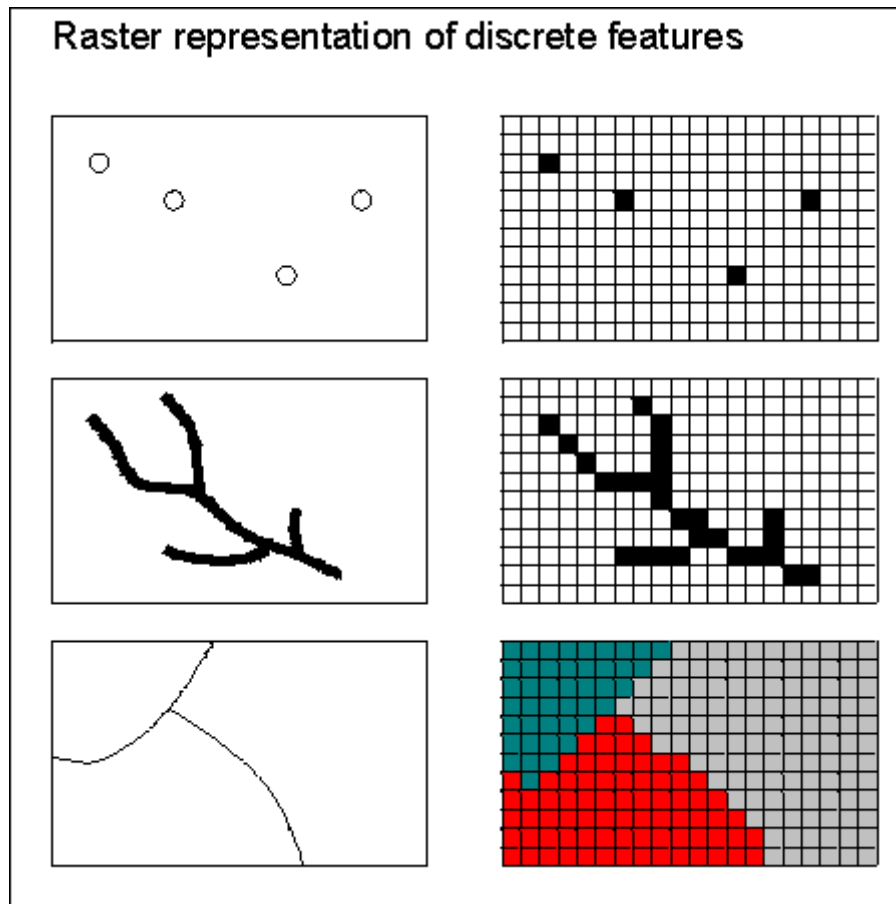


Fig: 4 d

- ❑ Since each cell is associated with a value called cell value or pixel value, it is very easy to carry out overlay operations to compare attributes recorded in different layers.
- ❑ Each attribute associated with a grid cell can be combined logically or arithmetically with attributes in corresponding cells of the other layers to create a new attribute value for the resulting overlay.
- ❑ Transitional areas are poorly represented by raster-based model.

The Choice between Raster and Vector Models

The choice between raster and vector based model depend upon the type of data analysis and other operations to be carried out for a project. However, there is always scope to convert one form to other. i.e., raster to vector or vector to raster

- ❑ Raster method for spatial data structure requires large memory space as compared to vector data.
- ❑ Certain kinds of data manipulation such as polygon intersection, union, clipping, merging etc are complex in raster data model as compared to vector.
- ❑ However, multi-theme overlay operations are more easy in raster data model.
- ❑ Similarly, representation of surfaces is more common in raster-based model.

Vector Data Model:**Advantages:**

- ❑ Good and real representation of geographic data
- ❑ Compact data structure
- ❑ Topology can be completely described
- ❑ Accurate graphic output
- ❑ Less storage space.

Disadvantages:

- ❑ Data structure is complex
- ❑ Combination of several vector polygons create difficulties in handling
- ❑ Simulation is difficult because each unit has a different topological form
- ❑ Display and plotting are expensive.

Raster Data Model:**Advantages:**

- ❑ Simple data structure
- ❑ The overlay of mapped data with remote sensing data is easy
- ❑ Simulation is easy because each spatial unit has same size and shape
- ❑ Good for multiplayer overlay.

Disadvantages:

- ❑ Data is voluminous and require large storage space
- ❑ Use of large cell to reduce data volume loses significant information
- ❑ Crude raster maps have ugly look
- ❑ Network linkages are difficult to establish.

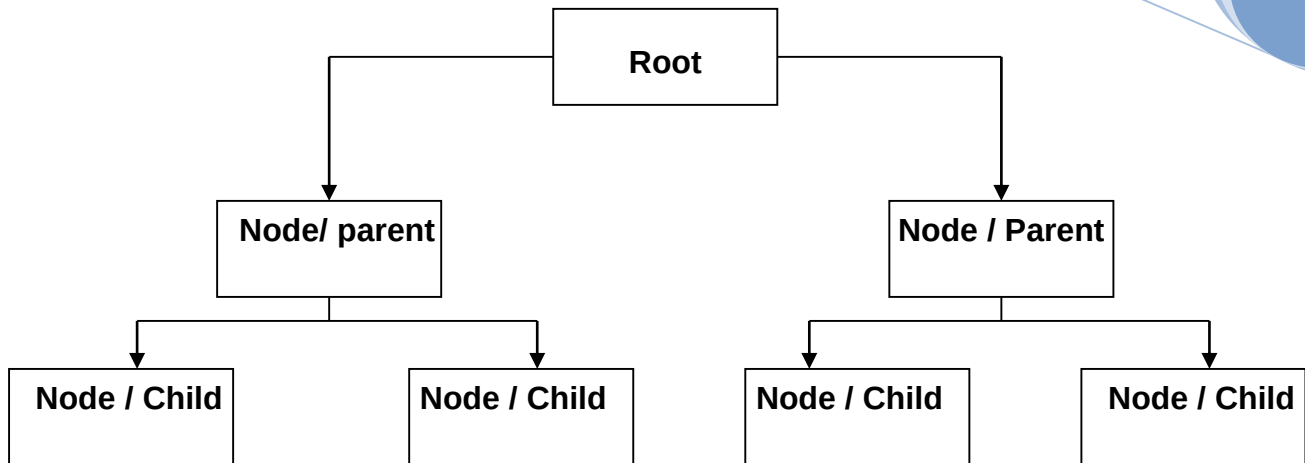
2.5 Conceptual Model for Non-Spatial Information:

- ❑ Non-spatial information, also known as attribute data, is descriptive data that defines spatial data.
- ❑ They are gathered and assembled into records and files

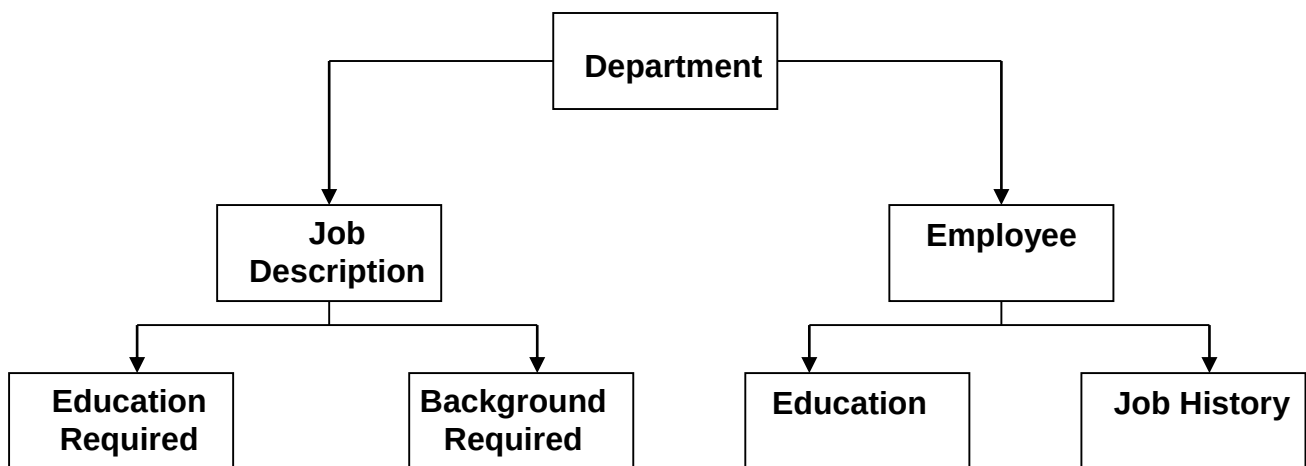
- ❑ A database is a collection of data that can be shared by different users. It is a group of records and files that are organized so that there is little or no redundancy.
- ❑ A database consists of data in many files. In order to access data from one or more files easily it is necessary to have some kind of structures or organization.
- ❑ Data Base Management System (DBMS) is a tool for representing, in computer, real world oriented model of set of data in a predefined structure and organized manner.
- ❑ This high level representation or abstraction is refereed to as Conceptual Model which ensures the data linking, data security, sub-setting, query using logical / arithmetic syntax etc.
- ❑ Most commercial DBMS softwares like Oracle, Dbase, MS Access etc are implemented by three types of data models namely Hierarchical data structure, Network structure and Relational structure.

1. Hierarchical Data Model:

- ❑ It is a tree-based structure. The tree is composed of nodes; the upper most node is called a root.
- ❑ With the exception of this root, every node is related to a node at higher level called its parent. The lower level is called child.



Example:

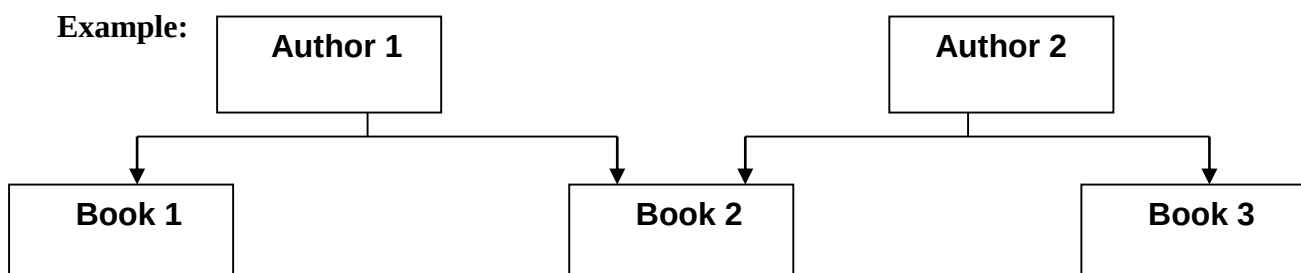


- ❑ This approach is efficient if all desired access paths follow the parent child linkage.
- ❑ However, it requires a relatively inflexible structure and hence linkage with other branch of database is tough or cumbersome. That is why this data base structure is not very common in flexible GIS.

2. Network Structure:

- ❑ Network structure exists when child in a data relationship has more than one parent.
- ❑ An item in such a structure can be linked to any other item.
- ❑ It is good for network-based analysis.

Example:



3. Relational Structure:

- ❑ In this case data are organized in 2D tables consisting of rows and columns. The rows are called records and columns are called items or fields.
- ❑ Such tables are easy to develop and understand.
- ❑ Different sets of tables are created within database and a relationship is established between each table.
- ❑ Because of this, it is easy to create a subset of data for one user or to join two tables for other user to form a large table.
- ❑ The structure can be described mathematically, hence mathematics provide the basis for extracting some columns from the table and for joining various columns.
- ❑ This capability to manipulate relations provides flexibility that is normally not available in hierarchy and network structures.

Relational Operators:

- ❑ Retrieval of data sets from relational model involves creation of new relation, which is a characteristic of permanently stored relations.
- ❑ There are several relational algebra operators that can be used to search and manipulate relations.
- ❑ These operators are implemented by means of Structured Query Language (SQL) using number of commands.
- ❑ E.g SELECT Settlement Name, warda name
 FROM Settlement

Will create a new table from Settlement table that consists of only settlement name and warda name.

```
SELECT
FROM Settlement
WHERE warda name = zero arath
```

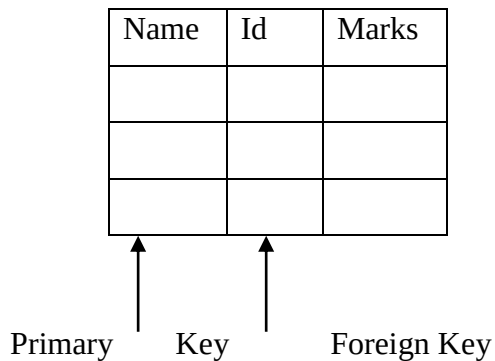
i) Primary and Foreign Keys

ii) Relational Joins

i) Primary and Foreign Keys:

- ❑ Relational approach is used to design database table
- ❑ Since each table or relation represents a set, it cannot have any rows whose entire contents are duplicated.
- ❑ Secondly, as each row must be different to every other, a value in a single column or a combination of values in multiple columns can be used to define a primary key for the table, which allows each row to be uniquely identified.

- ❑ The uniqueness allows the primary key to serve as the sole row level addressing mechanism in the relational database model.
- ❑ A field that stores the key of another table is called foreign key



ii) **Relational join:**

- ❑ The mechanism for linking data in different tables is called relational join.
 - ❑ Values in a column or columns in one table are matched to corresponding values in column in second table.
 - ❑ Matching is frequently based on primary key in one table and foreign key in the second column.
- e.g.**

Table-1

Name	Designation	Employee Id
MARTIN	PROFFESSOR	1107
GEORGE	READER	1206

Table -2

Name	Salary	Experience
MARTIN	5000	12