❖ Introduction :-

- GIS storage and editing subsystems provides a variety of tools for storing and maintaining the digital representation of a study area.
- It also provide tools for examining each theme for mistakes that may have crept into its preparation.
- Before we use these tools, we need to know what these possible mistakes are and how they can be discovered and corrected. If we have been careful in our input, we can encounter relatively few errors.
- There are many aspects for error detection and correction.

❖ Storage of GIS Database :-

- The methods of storing GIS database are themselves also highly dependent on the data model used in your system.
    1. In Raster System:
    2. In Vector System:

- ***In Raster System*** :-

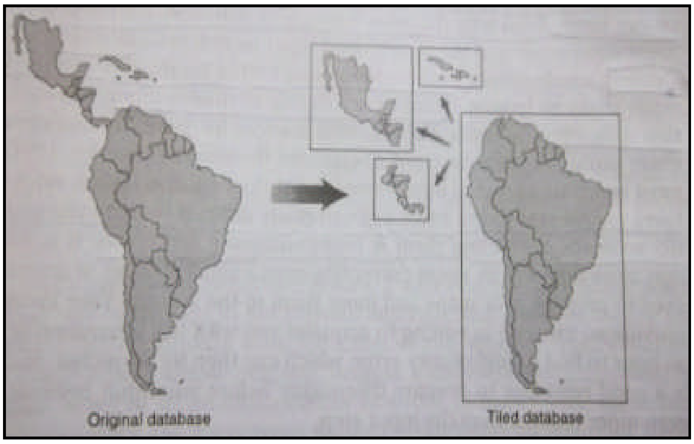
    - ✓ The attribute values for the grid cells are the primary data storage in the computer, usually on the hard drive.
    - ✓ **The locations** of each grid cell are catalogued by their position relatively to the order in which they are placed in columns and rows. In other words, their positions are relative to the location of the other grid cells.
    - ✓ Some raster systems, use **compact methods of storage** such as run-length codes, raster chain codes and quadtrees.
    - ✓ If your raster system allows to **linkage to a DBMS**, the matter becomes somewhat more complicated in that each grid cell has attached to it a number of different attribute codes.

- ***In Vector System*** :-

    - ✓ **The Entities and Attributes** are either store as individual tables with in a single database or as a separate database, linked by a series of pointers.
    - ✓ The separation of entities and attributes requires you to look at the editing procedure applied to entities, attributes and databases.
    - ✓ You can retrieve the graphic entity separately and display them to identify missing object, incomplete links and polygons. Finally, you will be able retrieve part or all of your database and to examine both entity and attribute.

➢ What is tiling ?

- ✓ *Many Vector GIS systems allows you to separately store portion of your large database. This process called tiling, is most often used to reduce the volume of data needed for analysis of extremely large database*.



- ✓ Tiling the portion individually means that you retrieve just a portion of overall database with which you are going to work.
- ✓ Another important purpose of tiling is to allow a system administrator to have final control over an editing and updating process by permitting only certain section of the database.
- ✓ Even when small portions are released for editing and updating, the system maintain an original copy of pre-edited database the system administrator s satisfied.
- ✓ Thus, the **advantages of tiling** are, you can:
  1. Reducing your computation
  2. Reduce overhead and
  3. Increasing the system response.

❖ The Importance of Editing the GIS Database :-

- ➢ Some errors might occur as a result of computational miscalculation and rounding error in a GIS software, most database errors result from improper input.
- ➢ The **causes** include :
  - Simply pushing the wrong button on the digitizer puck,
  - Typing errors during the attribute input, and
  - Even registration difficulties.
- ➢ Such a mistakes are generally very small and extremely difficult to find even with the best GIS software, correcting them is time-consuming and costly.
- ➢ Basically three general types of errors are there:
  1. **Entity Error,**
  2. **Attribute Error and**
  3. **Entity-Attribute Agreement error**

❖ Entity Error :-

- The first type of error you will encounter deal primarily with vector systems it is also called **Positional Error**.
- Entity Errors can take **three different forms** :
  - ✓ Missing Entities,
  - ✓ Incorrectly Placed Entities, and
  - ✓ Disordered Entities.

❖ Attribute Error :-

- Attribute error can occur in both vector and raster system.
- Most often attributes are type in, and volume of typing required for large database often constitute a major source of error.

- In vector system attribute error include using the **wrong code** for an attribute as well as **misspellings**, which make an attribute impossible to retrieve if a query uses a correct spelling.
- In raster the input most often consist of attributes, so the result of typing the wrong code number or placing it in the wrong grid location is a map that displays these incorrectly coded grid cells in the wrong place.

❖ **Entity-Attribute Agreement Error :-**

- Incorrectly placed attribute data comprise this third kind of error also called Logical Consistency, which is also occurs in vector systems when *correctly type codes are attached to wrong entities*.

➢ Of the three basic types of error found in GIS database, **last two are the most difficult to find.**

➢ Mistyped attribute placed in correct location might be found if an active data dictionary is part of the system. These feature generally is helpful if you have violated a rule already establish for the data dictionary.

➢ However misspelled attributes may not be found until you actually perform an analysis. Entity-Attribute Agreement is often even more difficult to find than misspelling or incorrect code.

➢ **How to find the Errors?**

- **In Raster**, the only way you can observe problems of these types is to display the map to identify the misplaced grid cell.
- **In Vector**, you will most often be able to point to entities and display their attributes on monitor.

➢ However, GIS is not likely to be able to tell you that you have the wrong attribute attached to a particular entity. Instead you will need to have a copy of your input map beside you as you display or highlight each entity.

➢ If you create a very complex database, you may have to spend months, evaluating each of the thousands of entities and making comparison to your input document.

➢ It is far better to check your error in small group, as you input the database.

➢ In addition, the input document is already there in front of you.

❖ Detecting and editing te Errors of Different Types :-

➢ As we have seen, a GIS is subjected to error involving:

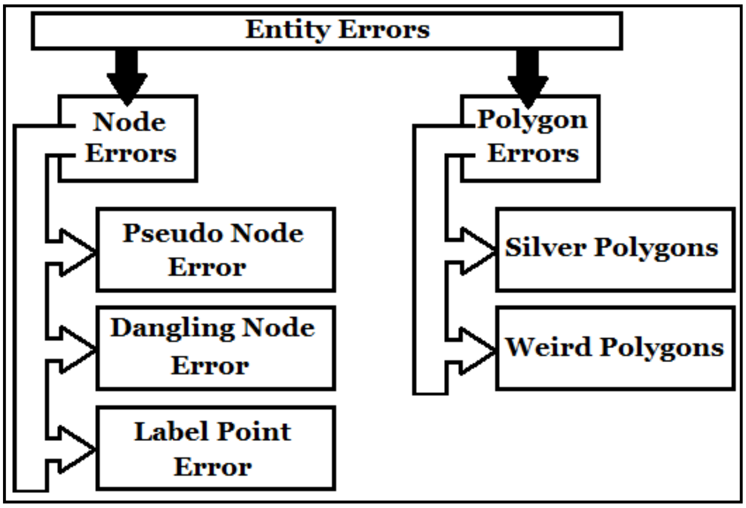
1. **ENTITY Errors : Vector**
2. **Attribute Errors : Raster and Vector Both**
3. **Entity-Attribute Agreement Error**

➢ Most often attribute errors are detected and identified because the entity and attribute failed to agree.

❖ Entity Errors : Vector :-

➢ **How to detect an Entity Errors?**

- Some of these error will be pointed out through text-based error flags telling you that you have a problem.
- Other must be interrupted by looking at database statistics concerning the numbers of entities and the types of entities.



- And some by inspection the graphic displayed in the screen, for error the GIS is not design to detect.

➢ You will be looking for six basic types of errors **represented by the negative case** of the following statements:

1. All entities that should have been entered are present.
2. No extra entities have been designed.
3. The entities are in the right place and are of the correct shape and size.
4. All entities that are supposed to be connected to each other.
5. All polygons have only a single label point to identify them.
6. All entities are within the outside boundary identified with registration mark.

➢ The useful procedure for identifying the errors is to comparing the entity you designed and the original map document.

➢ The latter will allow you the physically overlay the two maps using a standard light table.

➢ *In addition many GIS software provides number of symbols that indicate some errors*.

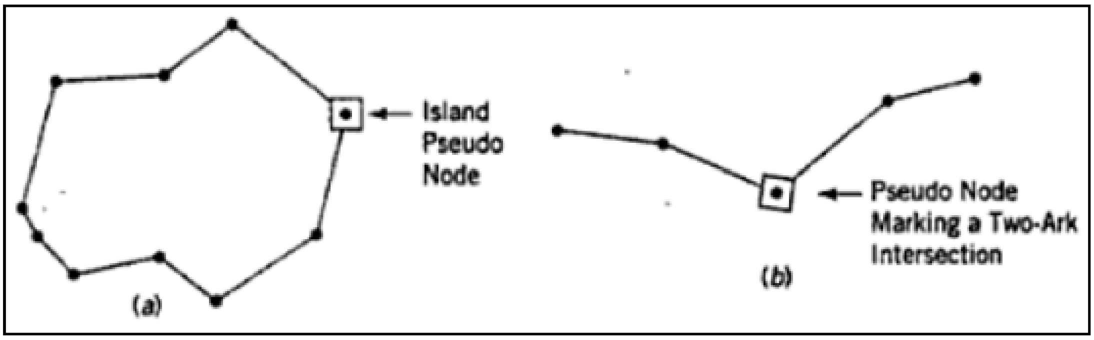
1. Node Errors :

➢ ***What is Node? Nodes*** are special points that indicate a link between lines compose of individual line segment.

➢ ***Use of Node*** **:** Nodes may be used to identify the existence of an **intersection** between two streets or connection between a stream and lake, but they should not occur at every line segment along a line or a polygon.

a) *Pseudo Node Error :*

✓ Thus the ***first type of error*** that can be detected entails false nodes called ***Pseudo Nodes***, which occur where a line connects with each itself or where two lines interact along a parallel path rather than crossing.

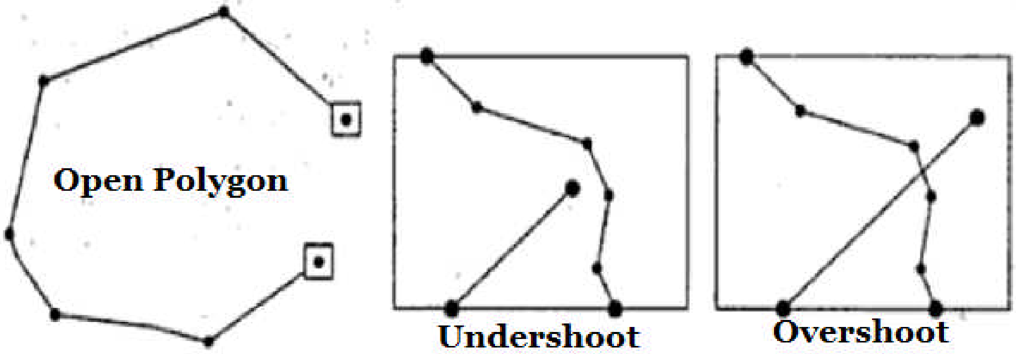


If in figure (a) there is no Island, this indicate an error and must be corrected.

- ✓ A pseudo node connecting a line itself may simply be the beginning and ending of an island polygon (sometimes called a **Spatial Pseudo Node**) in which case its ***flag*** **can be ignore**.
- ✓ **Reason** for generating Spatial Pseudo Node :-
  Either you were trying to create an unclose structure nut place the puck on the wrong place or you were trying to create a polygon that connected to other polygons but push the wrong button.
- ✓ If your software indicate that in your coverage there is one or more Pseudo Nodes, the prepared **map document can be used to help you to correct the mistakes.**

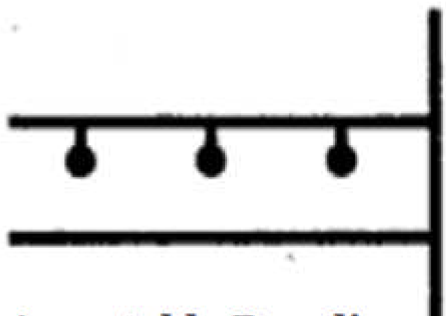
b) ***Dangling Node Error :***

- ✓ Another common node error called ***Dangling Node***, can be defined as a single node connected to a single line entity.
- ✓ In some GIS packages, you often need to have a **From Node** & **To Node** rather than just a single node.
- ✓ Dangling nodes, sometimes called Dangles, can result from **three possible mistakes**
  1. Failure to close a polygon (**Open Polygon**).
  2. Failure to connect the node to the object it was supposed to connected to called an **Undershoot**.
  3. Going beyond the entity you were suppose to connect called **Overshoot**.



- ✓ **Reason** for generating Dangling Node :-
  - o Incorrect Placement of Digitizer Puck.
  - o Fuzzy Tolerance.
- ✓ It is generally easier to find overshoots than undershoots.

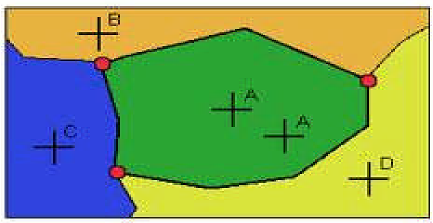
- ✓ If your dangle indicate an open polygon, the GIS will alert you by telling the number of complete polygons in database: if it is differ from the count you had prepared prior to the digitizing, you know you need to look for these dangles as incomplete polygons.
- ✓ **Methods for correcting the Dangling Node Errors :**
    - *In the case of Open Polygon* you merely moved any of the node to connect to the other.
    - *For Undershoots,* the node is identified and it moved or ***snapping*** to the object to which it should have been connected.
    - *Overshoots errors,* are corrected by identifying the intended lines interaction point and ***clipping*** the line so that it connect where it is supposed to.
- ✓ As with pseudo node dangling nodes are intentionally input to the GIS for a particular purpose.
- ✓ Most often these nodes serves as indicator of something important at the end of line or arc.
- ✓ For example, you might use nodes to indicate the location of residential cul-de-sacs.
- ✓ In even more unusual circumstances, a line used to indicate the location of a multistory building may contain numerous nodes, each one indicate the location of separate floor.



Acceptable Dangling Nodes

c) *Label Point Error :*

- ✓ Another node error called ***Label Point Error*** .
- ✓ **What is Label?**
    - "*While digitizing a polygon, you will need to indicate the point inside each polygon that will act as a locator for a label on which you will display text information about the polygon.*"



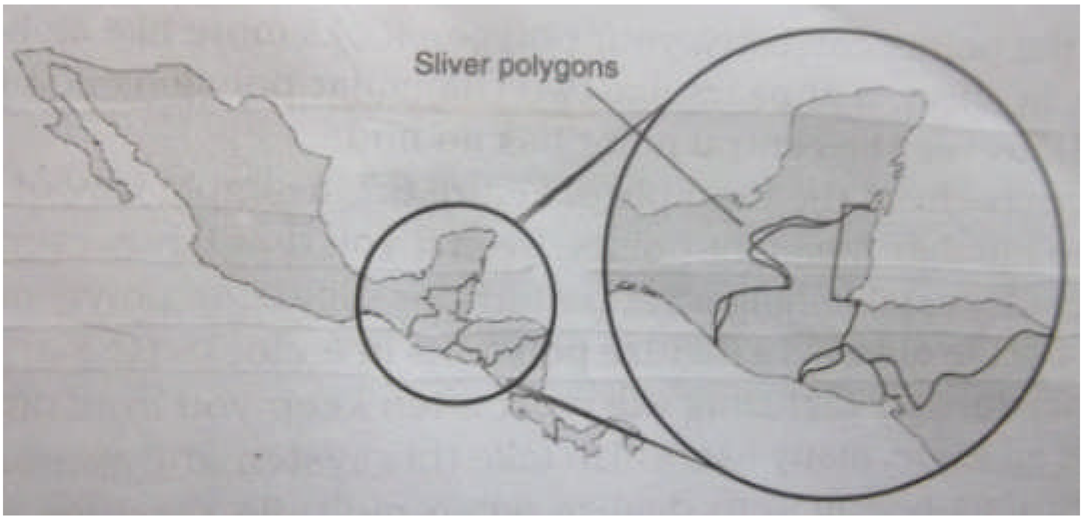
- ➢ ***You need only one Label Node per polygon***.
- ➢ Two types of errors can occur related to the Label Node:
    1. **Missing Label**
    2. **Too Many Labels**
- ➢ Both of these errors are most often caused by failure to keep digitizing process.

2. Polygon Errors :

a) *Silver Polygon :*

- ✓ Another type of digitizing error called **Silver Polygon** most commonly occurs when the software uses a vector data model that treats **each polygon as a separate entity**.

- ✓ In such a cases you are required to digitize the adjacent lines between polygons more than once.
- ✓ **What is Silver Polygon?**
  *"Failure to place the digitizer puck at exactly the correct location for each point along that line will often result in a series of tiny graphic polygons called **Silver Polygon**".*



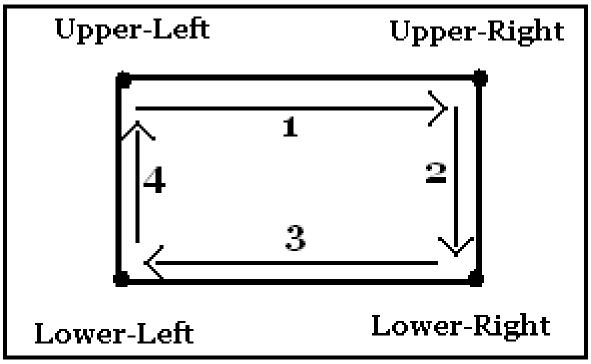
- ✓ **Reason for Silver Polygon**
  Silver polygon can also occur as a result of overlay operation or when each of two adjacent maps is input from a separate projection.
- ✓ **How to avoid Silver Polygon Error?**
  - To avoid Silver Polygon on input is to use the GIS that does not require digitizing the same line twice.
- ✓ **Methods of Finding Silver Polygon.**
  - The method of finding Silver Polygons depends somewhat on whether you actually completed the adjacent lines with nodes that are effectively placed on top of each other.
  - Finding Silver Polygon in the absence of dangling node is more difficult. One way is to compare the number of polygons produced in your digital coverage with that of the original input map.
  - It is very difficult to locate Silver Polygon, however even though you know they are there. Most often you have to move through your image, searching for suspect polygon boundaries, then **Zoom In** to see Silver Polygon.

b) Weird Polygon

- ❖ Another problem related to the polygons is the production of **Weird Polygons**, which are defined as a polygon with missing node that appear to a true polygon but is meassing one or more nodes.
- ❖ Generally, this occur when two or more lines cross over, producing the semblance of a polygon.
- ❖ The most frequent ***cause of this error*** is a point digitize in the wrong place or in the wrong order.
- ❖ **For Example**, you have a rectangular polygon that required four points to define it. You would want to start at the upper-left node, move to the upper-right then lower-

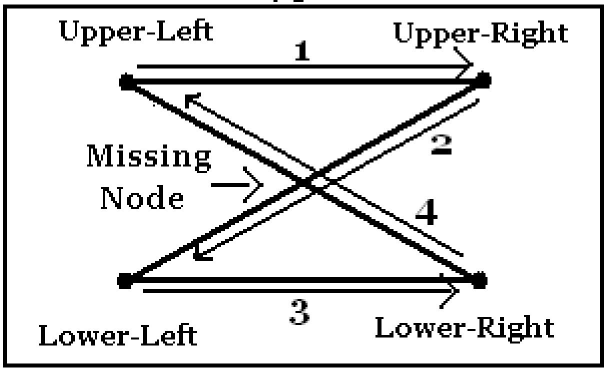
right and lower-left and end at the upper-left from where you begin (**Correct Order for Rectangular Polygon**).

Correct Order for Rectangular Polygon



Correct Polygon

Incorrecr Order for Rectangular Polygon



Weird Polygon

- ❖ You instead go from Upper-Left to Upper-Right, and then by mistake go to the Lower-Left and then Lower-Right and finished at the Upper-Left where you had begun (**Incorrect Order of Rectangular Polygon**).
- ❖ It appear like a two triangle polygons connected at the center point.

❖ **<u>How to avoid this error</u>?**

- ❖ Simply way to avoid the error is to give number the input points.
- ❖ Even if you don't, you can avoid the problem by establishing the set of pattern for digitizing polygons. For Example, you may decide always to digitize a polygon in clockwise fashion.

❖ **<u>How to detect this error</u>?**

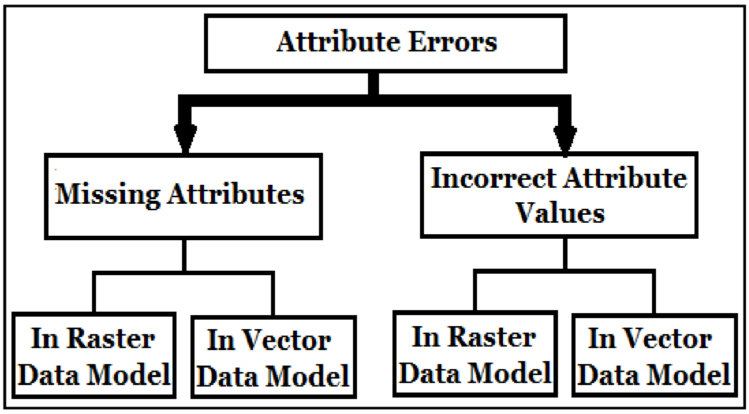
- ❖ Detecting the Weird Polygon is difficult but not impossible.
- ❖ One straightforward method is to highlight the points and display them as a part of the polygon coverage.

❖ **<u>How to edit this error</u>?**

- ❖ Editing the error involves moving the line at the correct location, thereby placing the nodes at the correct sequence.

❖ **Attribute Errors : Raster and Vector :-**

- ➢ Attribute Error, including the entity-attribute agreement error are most difficult to detect because GIS does not know which attribute is correct and which is not.
- ➢ <u>As attribute no equivalent of the topology, there are no rules against which the GIS can check your accuracy</u>.



- **<u>Missing Attribute</u> :**

  Missing attributers are perhaps the only attribute errors that are detectable without direct comparison to the input documents.

1. <u>In Raster</u> :-

- These tends to occur as a map that are missing whole rows or columns or a portion or rows or column.
- These are detectable because familiarity with the original map shape will alert us to the absence of row or column of grid cells.
- ***Reason*** : Missing rows and columns in raster data model are most often cause by missing one's place while typing the grid cell value.
- If you are typing in a grid cell values and missing several of the attribute values, quit often an unusual value will appear in your map, that has little correspondence to the rest of the data.
- ***To Avoid*** this error is to display the text equivalent of the attributes prior to completing the input.
- Alternatively, if you are using the software to input the values by typing them in, the GIS most likely will indicate the row numbers as you enter the data row by row.
- If you missing a complete row of data, or more then one row of data, your map will appear shorter than it should.

2. <u>In Vector</u> :-

- ***Reason*** : Missing Attribute vales in Vector GIS are commonly caused by simply <u>not including anything in the attribute tables for individual points, lines, or polygons</u>.
- ***This can be identify by*** listing the tabular information and identifying missing attribute values in the table, or by outputting the entities and their attributes as a video display.
- If you using the second method, you can edit each entity by selecting it from the rest and inputting the appropriate attribute values.
- And, of course, you need to remember to save your editing work.

- **<u>Incorrect Attribute Values</u> :**

  These are very difficult to detect, both in Raster and in Vector.

1. <u>In Raster</u> :-

- GIS, when they occur as an individual cases, or short row or column segments, they are most ***<u>often causes by</u>*** typing errors when that form of input is used.
- In raster, the incorrectly coded grid cell is most likely to be identifiable as one individual that seems "**<u>Out of Place</u>**" among the sounding grid cell.
- They normally appear as a out-of-place grid cells that **<u>disturb the natural organization of the map</u>**.
- **How to edit?** If they occur as a continuous strips of incorrect grid cell attributes, most software will allow you to use **<u>run-length encoding strategy</u>** to edit these.

Email: ajay.24021985@gmail.com **Mobile:** +91-95588 20298

- **How to detect?**
- You will need to compare the shape of area pattern on the raster map against the original shape of the input map.
- **How to correct?**
- Correcting such a problem usually means reevaluating each grid cell as to its correct attribute.
- In other words, which of the two adjacent area does this grid cell really belong to? Once the answer can be determine each can be selectively edit as before.

2. In Vector :-

- This error is more difficult to identify in vector than in raster because:
- **Reason :** Vector cases generally calls for sources map and its attributes.
- If you are using a form of coding strategy that replace the actual name of the item by, for example a *numeric vale that are sometimes not corresponds to linked tabular information include in other portion of your database*.
- **How to avoid?**
- If you established an appropriate set of rules in you data dictionary especially it is an Active Data Dictionary it should be possible to avoid this error.
- **How to detect?**
- Systematic error would be more easily detected by comparing the tabular data from the database with the data from which the table is produce.

❖ Dealing with Projection Changes :-

- Major function of the storage and editing subsystem is the correction of entity and attribute errors, it is also used to ***convert between the Cartesian Coordinates and real-world coordinates base on the reference glob***.
- The software will require you to identify the projection of the map upon input.
- Conversation to a set of projection is necessary because : (**Reasons for Conversation**)
  - **Any analysis require real-world measurement.**
  - **Not all your input maps will have the same type of projection.**

- ***What to do to make the projection possible?***
  - If your system operate on the real-world coordinate system but has you input the data strictly in a Cartesian coordinate structure, you must define a reference points as latitude/longitude coordinate ***to make the projection possible***.
  - Another important factor is the method by which the GIS stores and manipulate these coordinates.
  - Different formats for storing the Latitude/Longitude Coordinates :
    1. **Degrees, Minutes and Seconds (DMS)**
    2. **Decimal Degree (DD)**
  - ESRI (Environmental Systems Research Institute 1992) has given a formula to convert DMS to DD :

Degree + Minutes/60 + Seconds/3600

  - It gives the number in Degree or Degree of Fractions.

- ***How to convert from DMS to DD?***
  - You will operate first on the reference coordinate by creating a separate map layer with these values only.
  - The values can be read directly from associated tables, then edit by typing their Latitude and Longitude equivalent.
  - The reference points are now in Geographic Coordinate.

- ***Why to convert from DMS to DD?***
  - As you remember, the map from which you input your data will likely have been produce through map projection.
  - You must get your coordinates, which are now in geographic projection, into the same type of projection as the original input document.
  - Usually few commands or even a single command will necessary to set the GIS about the transformation procedure of Scale Change, Rotation and Translation.
  - This transformation require mathematical manipulation of your original geographic coordinates, so errors will always be part of procedure.
  - **No projection procedure is without error.**

❖ Joining Adjacent Maps: Edge Matching :-

- In edge matching, two adjacent maps, usually of the same theme, are physically linked to permit the analysis of the large study area.
- There are ***two sources of difficulty*** when two adjacent maps are input:

1. **First (For Same Projected Maps):**
   - Two maps that are input with the same projection but because they were put separately they are likely to display entity errors that are somewhat different.
     - The maps are registered to the digitizing table separately.
     - Tick marks or reference points were input separately.
     - Entities were input during a separate digitization session.
   - You will need to link all the line and polygon entities that are supposed to be connected.
   - **For Example**, if a road digitized on one map sheet is supposed to be straight line that runs across the two sheets, make sure that when the sheets are connected, both portion of the road will have to be connected so that they exist as a straight line.

2. **Second (For Different Projected Maps):**
   - The second type of difficulty is occur when we are trying to perform edge matching on two adjacent maps that are:
     - ✓ Input using different projections.
     - ✓ Input using same projection but different base line.
     - ✓ Input using different starting points.
   - Edge matching problem can also occur in the raster systems, at least for those that operate on the projected surface, rather than simply in flat Cartesian place.

- A common example in edge matching in a raster system is found in the use of remotely sensed data products.
- Because horizontally adjacent scenes are sensed at different times, the satellite may no be located at exactly the same latitudinal coordinates.
- In addition, **geographic coordinates are most often provided with such data, allowing the user to match the edges** by matching the coordinates along both adjacent images.

❖ Conflation :-

- There are frequently circumstances, in vector, when some ***content of two digital source maps need to be combine***.
- This may be ***require to make a composite map*** that is superior to the original.
- **For Example**, you may have a map of hydrology that is extremely accurate, but you also have map of vegetation where you wish to compare the vegetation along the river to the river. This require you to combine both maps to create a map of riparian vegetation.

  Map of Hydrology + Map of Vegetation = Map of Riparian Venation
- You may also wish to compare multiple maps when you have multiple dates of the same region, but each map has digitize from a different set of aerial photography.
- Another circumstances include the consolidation of multiple thematic datasets within a single organization, or addition of features to an existing map.
- All of these require a complex process called **Conflation** sometimes called **Rubber Sheeting**.
- *"Conflation/Rubber Sheeting is the process of comparing multiple maps, consolidation of multiple maps and addition of features in existing map".*
- "*The process of conflation is involves indentifying features within one reference map that most reasonably correspond to accurate location of real earth objects that are to be combined with one or more target maps*".
- In some cases the process may be repeated to guarantee that the ***best possible outcome is obtain***.

❖ Templating :-

- When you ***look at multiple maps of the same theme for different time frames***, you notice number of graphical discrepancies, including one we have thus far neglected.
- Viewing the same map simultaneously, you also notice that, despite all your to prevent it, the outside dimensions of all study areas seems to ***differ slightly in the shape***.
- When you input such a amps, you choose certain data points as reference point and assign them to true graphic coordinate, yet all maps are not identical.
- Then the difference in location of these reference points from coverage to coverage, combine with the nuances of the projection algorithms and computer rounding error, have produce slightly ***different result for each coverage***.
- If you must latter perform overlay of four maps, there will be numerous area along the margin of the ***some maps that will not have associate area for other maps***.

- "*You must select a coverage you trust most to be representative and use it as a **Template** also called as a **Cookie Cutter***".
- If the boundary of the template is within the boundaries of all other maps, **you simply use this pattern to cut out the study area**.
- However if any of its boundaries beyond those of other maps, you will need to select coordinates somewhat inside the margin of the *template to ensure that all subsequent maps will occupy all area within the template*.

Email: ajay.24021985@gmail.com **Mobile:** +91-95588 20298