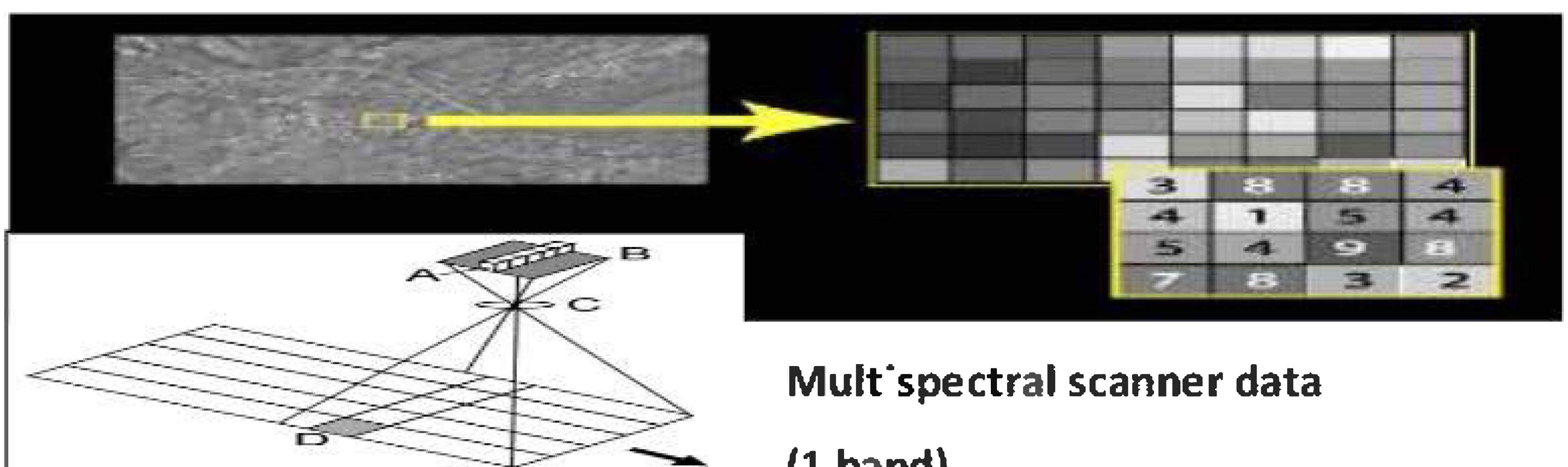
Topic No: 03 Raster Data model, Vector data model, Attribute data model.

Raster Data moda :

A format for storing, processing, and displaying graphic data in which graphic images are stored as values for uniform grid cells or pixels is called raster data model Raster data models incorporate the use of a *grid-cell* data structure where the geographic area is divided into cells identified by row and column. This data structure is commonly called Raster.



Mult˙spectral scanner data

(1 band)

Fig No: 04 To show the Raster data mode

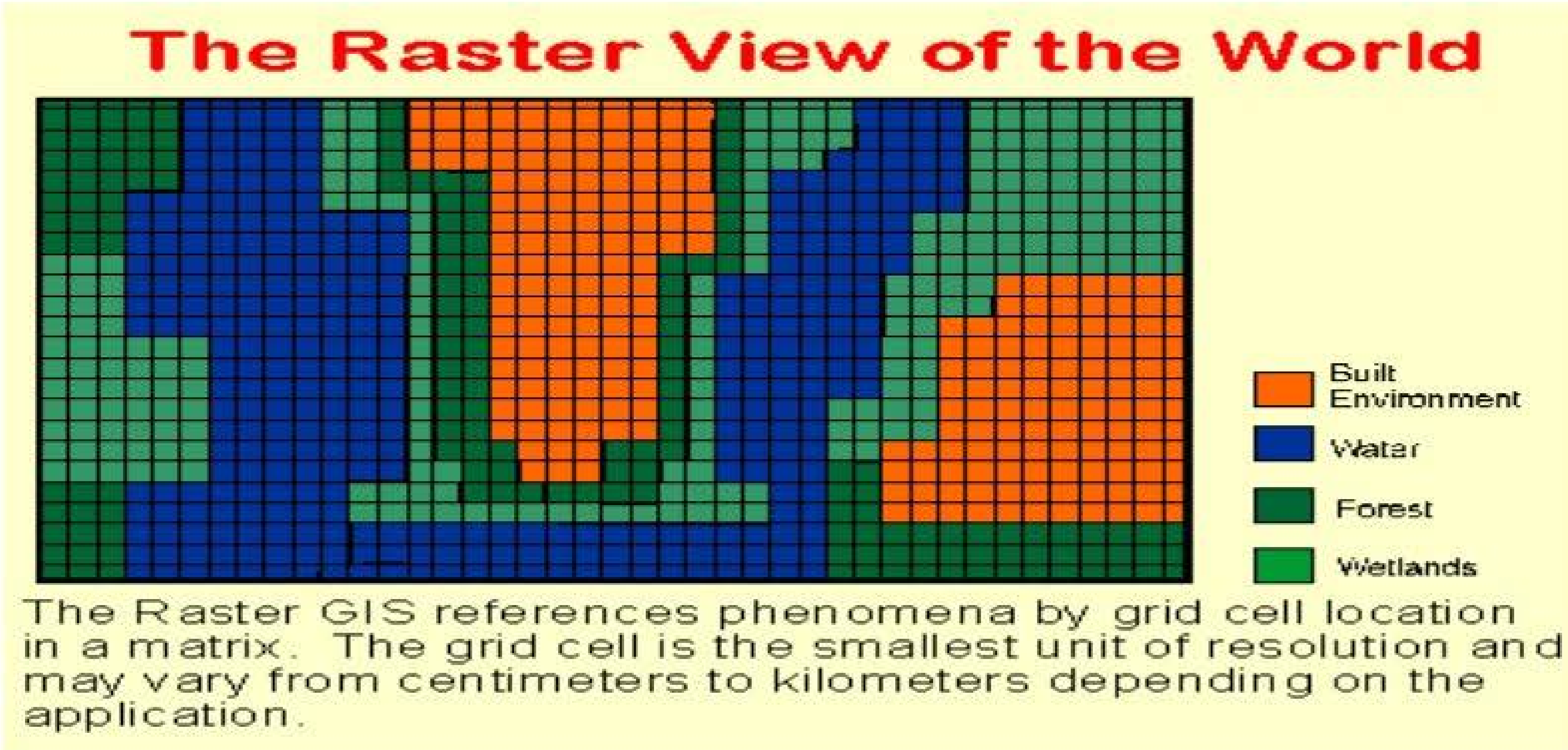


Fig No: 05 the Raster View of the World

Elements of Raster Data modal:

- ✓ A raster data model can be a grid, a raster map, a surface cover, or an image in GIS.
- ✓ A raster represents a continuous surface; however, for data storage and analysis, a raster is divided into rows, columns and cells.
- ✓ Cells are called 'pixels' with images
- ✓ The origin of rows and columns is typically at the upper-left corner of the raster
- ✓ Rows represent *y-coordinates*
- ✓ Columns represent *x-coordinates*

Advantages of raster data model:

- ✓ It has a simple data structure
- ✓ Overlay operations are easily and efficiently implemented.
- ✓ Area and polygon analysis is simple.
- ✓ Overlaying and merging are easily performed.
- ✓ Image processing techniques produce data for integration to GIS in a raster format.
- ✓ The inherent nature of raster maps, e.g. one attribute maps, is ideally suited for mathematical modeling and quantitative analysis.
- ✓ Discrete data, e.g. forestry stands, is accommodated equally well as continuous data, e.g. elevation data, and facilitates the integrating of the two data types.

Disadvantages of raster data model:

- ✓ The cell size determines the resolution at which the data is represented.
  - It is especially difficult to adequately represent linear features depending on the cell resolution. Accordingly, network linkages are difficult to establish.
  - Processing of associated attribute data may be cumbersome if large amounts of data exists. Raster maps inherently reflect only one attribute or characteristic for an area.
  - Since most input data is in vector form, data must undergo vector-to-raster conversion. Besides increased processing requirements this may introduce data integrity concerns due to generalization and choice of inappropriate cell size.
  - Most output maps from grid-cell systems do not conform to high quality cartographic needs. poor at representing points, lines and areas; good at surfaces
- ✓ must often include redundant or missing data
- ✓ network linkages are difficult to establish

Vector data modal:

All spatial data models are approaches for storing the spatial location of geographic features in a database. Vector storage implies the use of vectors (directional lines) to represent a geographic feature.

Advantages of vector data model:

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

Disadvantages of vector data model:

- ✓ The location of each vertex needs to be stored explicitly.
- ✓ For effective analysis, vector data must be converted into a topological structure.

- ✓ Algorithms for manipulative and analysis functions are complex and may be processing intensive. Often, this inherently limits the functionality for large data sets, e.g. a large number of features.
- ✓ Continuous data, such as elevation data, is not effectively represented in vector form. Usually substantial data generalization or interpolation is required for these data layers.
- ✓ Spatial analysis and filtering within polygons is impossible

Characteristics

- ✓ Features are positioned accurately
- ✓ Shape of features can be represented correctly
- ✓ Features are represented discretely (no fuzzy boundaries)
- ✓ Not good for representing spatially continuous phenomena
- ✓ Potentially complex data structure (especially for polygons);
- ✓ can lead to long processing time for analytical operations.

Attribute data model

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

The tabular model is the manner in which most early GIS software packages stored their attribute data. The next three models are those most commonly implemented in database management systems (DBMS). The object oriented is newer but rapidly gaining in popularity for some applications. A brief review of each model is provided.

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 modelling 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

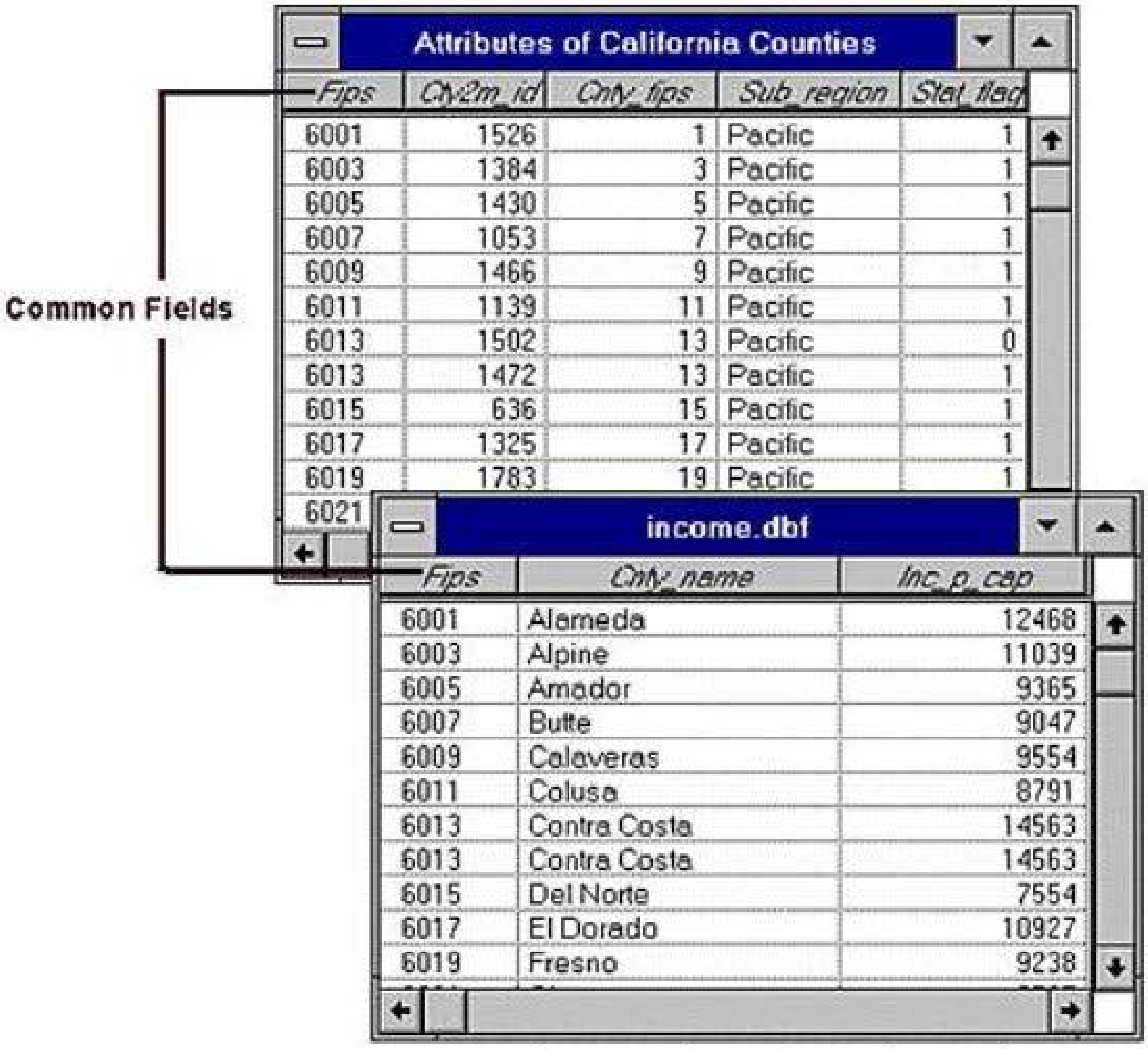
The relational database organizes data in *tables*. Each table, is identified by a unique table name, and is organized by *rows* and *columns*. Each column within a table also has a unique name. Columns store the values for a specific attribute, e.g. cover group, tree height. Rows represent one record 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. The following figure presents a sample table for forest inventory features. This table has 4 rows and 5 columns. The forest stand number would be the *label* for the spatial feature as well as the *primary key*for the database table. This serves as the linkage between the spatial definition of the feature and the attribute data for the feature.

UNIQUE STAND NUMBER	**DOMINANT COVER GROUP**	**AVG. TREE HEIGHT**	**STAND SITE INDEX**	**STAND AGE**
001	DEC	3	G	100
002	DEC-CON	4	M	80
003	DEC-CON	4	M	60
004	CON	4	G	120

Data is often stored in several tables. Tables can be joined or referenced to each other by common columns (relational fields). Usually the common column is an identification number for a selected geographic feature, e.g. a forestry stand polygon number. This identification number acts as the *primary key* for the table. The ability to join tables through use of a common column is the essence of the relational model. Such relational joins are usually ad hoc in nature and form the basis of for querying in a relational GIS product. Unlike the other previously discussed database types, relationships are implicit in the character of the data as opposed to explicit characteristics of the database set up.

The relational database model is the most widely accepted for managing the attributes of geographic

There are many different designs of DBMSs, but in GIS the relational design has been the most useful. In the relational design, data are stored conceptually as a collection of tables. Common fields in different tables are used to link them together. This surprisingly simple design has been so widely used primarily because of its flexibility and very wide deployment in applications both within and without GIS.



Fips	Cty2m_id	Cnty_fips	Sub_region	Stat_flag
6001	1526	1	Pacific	1
6003	1384	3	Pacific	1
6005	1430	5	Pacific	1
6007	1053	7	Pacific	1
6009	1466	9	Pacific	1
6011	1139	11	Pacific	1
6013	1502	13	Pacific	0
6013	1472	13	Pacific	1
6015	636	15	Pacific	1
6017	1325	17	Pacific	1
6019	1783	19	Pacific	1
6021				

Fips	Cnty_name	Inc_p_cap
6001	Alameda	12468
6003	Alpine	11039
6005	Amador	9365
6007	Butte	9047
6009	Calaveras	9554
6011	Colusa	8791
6013	Contra Costa	14563
6013	Contra Costa	14563
6015	Del Norte	7554
6017	El Dorado	10927
6019	Fresno	9238

Fig No:06 Relational Model of Attributes

In the relational design, data are stored conceptually as a collection of tables. Common fields in different tables are used to link them together.

In fact, most GIS software provides an *internal* relational data model, as well as support for *commercial off-the-shelf* (COTS) relational DBMS'. COTS DBMS' are referred to as *external* DBMS'. This approach supports both users with small data sets, where an internal data model is sufficient, and customers with larger data sets who utilize a DBMS for other corporate data storage requirements. With an external DBMS the GIS software can simply *connect* to the database, and the user can make use of the inherent capabilities of the DBMS. External DBMS' tend to have much more extensive querying and data integrity capabilities than the GIS' internal relational model. The emergence and use of the external DBMS is a trend that has resulted in the proliferation of GIS technology into more traditional data processing environments.

The relational DBMS is attractive because of its:

- Simplicity in organization and data modelling.
- Flexibility: data can be manipulated in an ad hoc manner by joining tables.
- Efficiency of storage: by the proper design of data tables redundant data can be minimized and
- The non-procedural nature: queries on a relational database do not need to take into account the internal organization of the data.

The following diagram illustrates the basic linkage between a vector spatial data (topologic model) and attributes maintained in a relational database file.

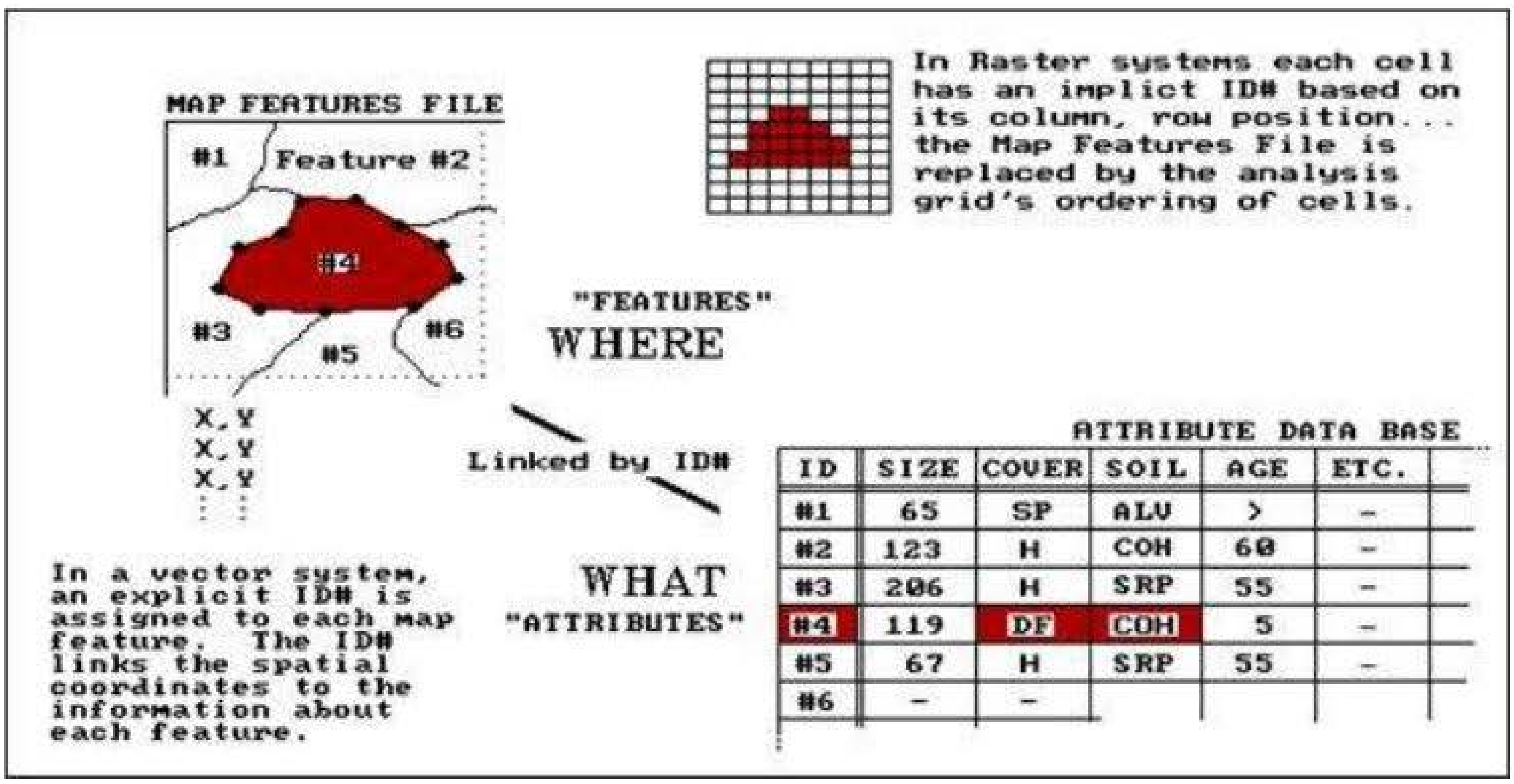


Fig NO: 07 Basic linkages between a vector spatial data (topologic model) and attributes maintained in a relational database file (From Berry).