Database design & development in support of Disaster Risk Management

Workshop 2: Technical Issues Towards Effective Applications of Geospatial Technologies and Data in DRM

Date : Monday 6-8 September 2021

Presented by Dr. Jacob Opadeyi,
Disaster Risk Management Consultant

Geospatial Database Design Concepts

Databases

- The “I” in GIS
- GIS are often split into two components
  - Coordinate information (describes object geometry or spatial information)
  - Attribute information (describes other non-spatial properties associate with it)
    - Often referred as tabular data as they are presented in tabular form

Databases - cont.

Name	FIPS	Pop90	Area	PopDn
Whatcom	53073	128	2170	59
Skagit	53057	80	1765	45
Clallam	53009	56	1779	32
Snohomish	53061	466	2102	222
Island	53029	60	231	261
Jefferson	53031	20	1773	11
Kitsap	53035	190	391	485
King	53033	1507	2164	696
Mason	53045	38	904	42
Gray Harbor	53027	64	1917	33
Pierce	53053	586	1651	355
Thurston	53067	161	698	231
Pacific	53049	19	945	20
Lewis	53041	59	2479	24

Attribute Information Presentation

- In GIS, attribute information are typically entered, analyzed, and presented using a database management system (DBMS)

- DBMS is a specialized computer program for data organization and manipulation

DBMS Functions

- DBMS incorporates a special set of software tools to manage the GIS non-spatial tabular data
  - Efficient data storage
  - Data retrieval
  - Data indexing
  - Data reporting
- DBMS often provides data independence
  - Valuable while working with large data sets
- Multiple user views
- Centralized control and maintenance of important data

Database software...

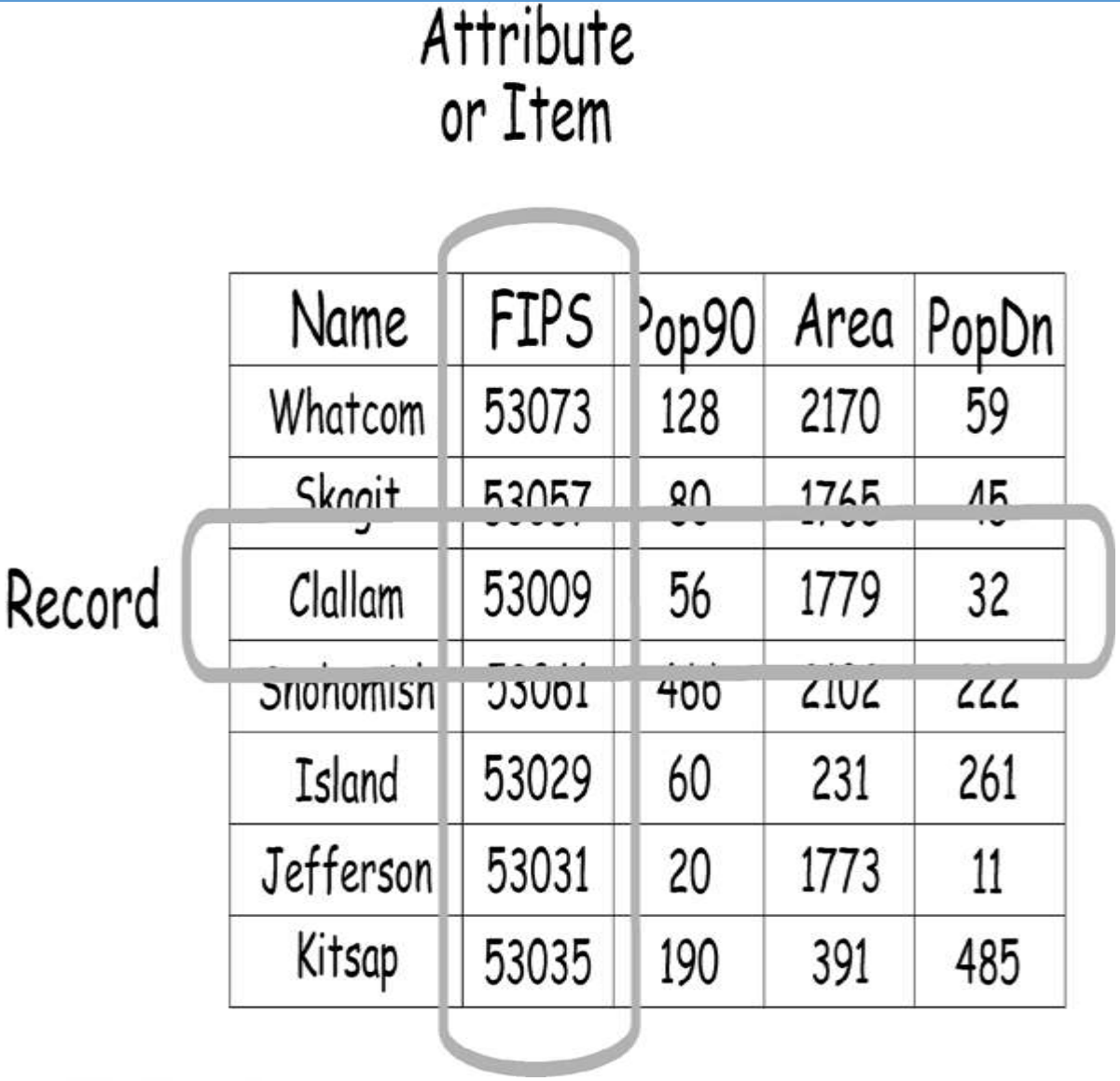
- Light Duty

- Medium Duty

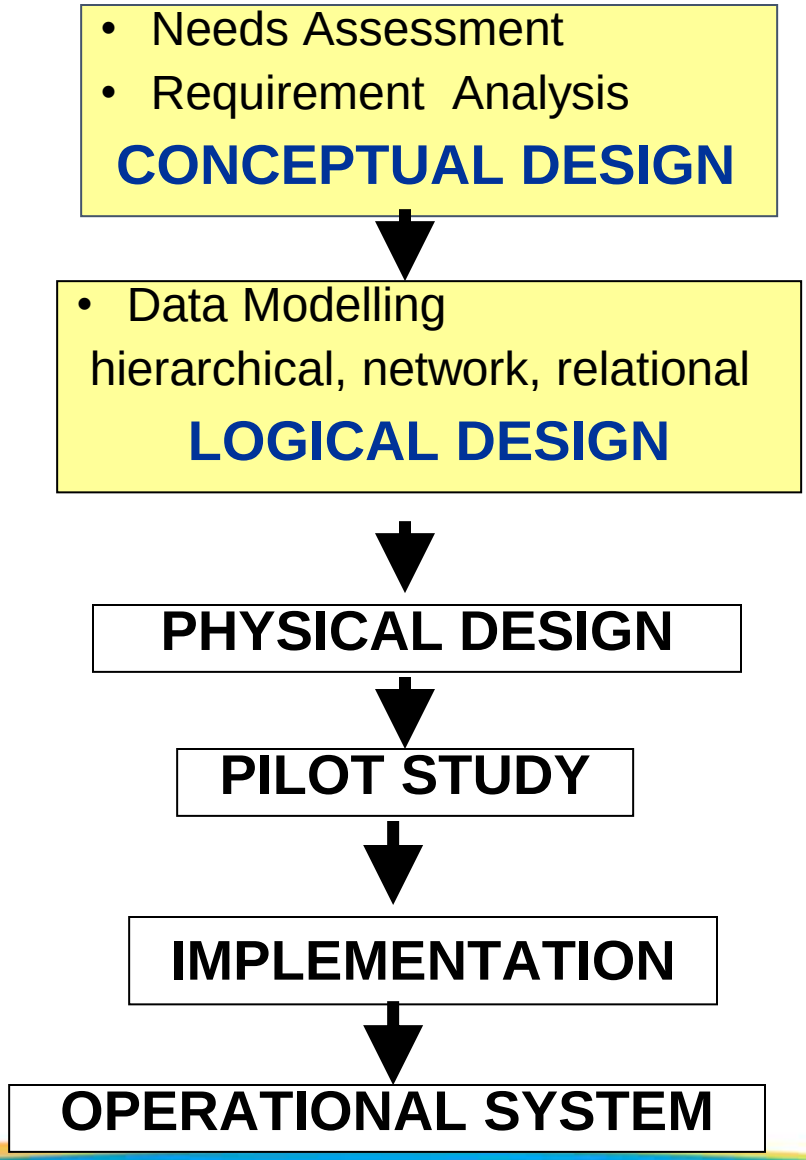
- Heavy Duty

Database Components

- Attribute/item
- Entity
- Instance
- Record
  - Row/n-tuple



Database Design Procedure



Conceptual Design

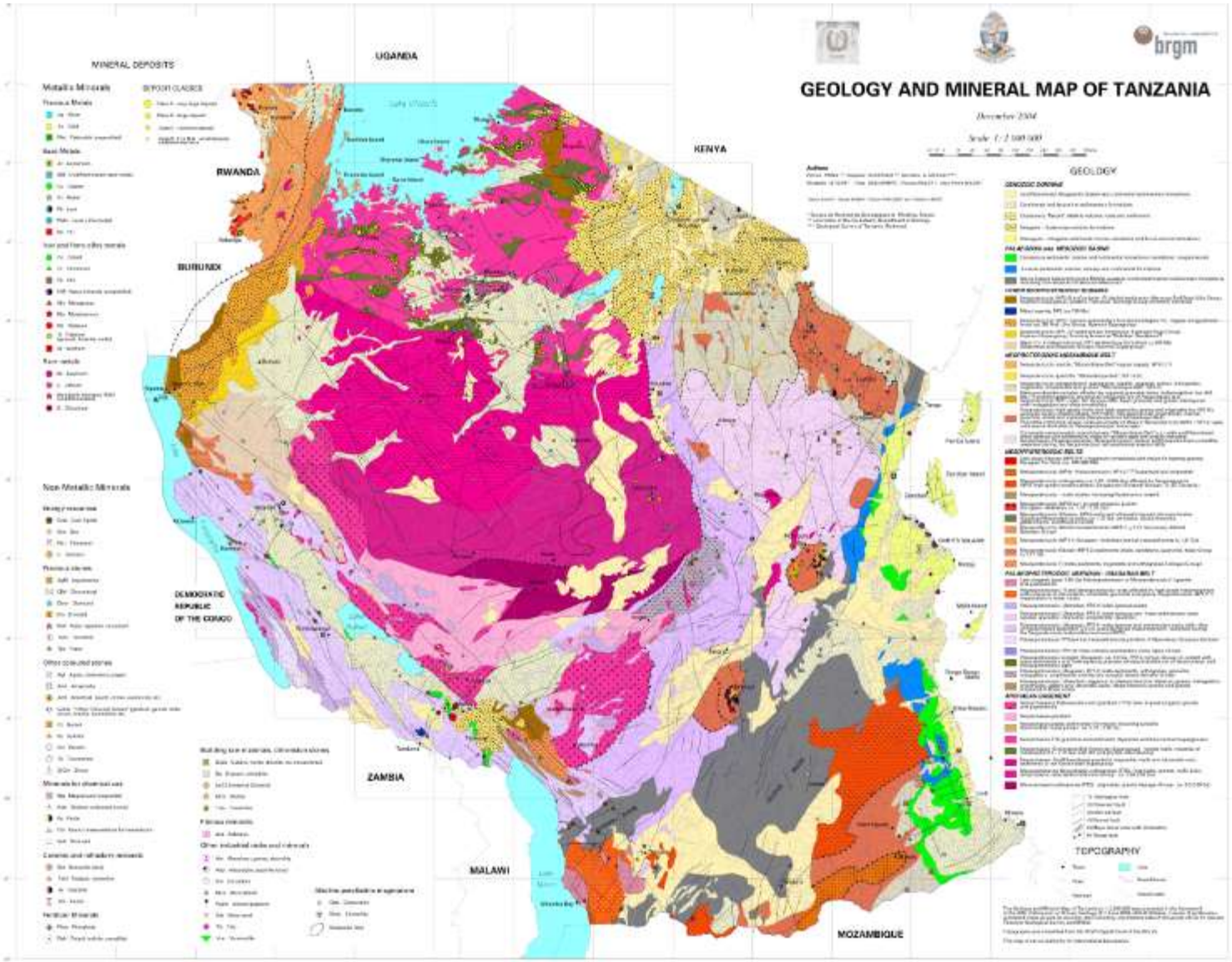
- Identifies and evaluates alternatives
- Builds framework for database
- Represents a generalized ‘user’ view of data
- Software and hardware independent
- Identifies how entities will be represented in database
  - e.g. point, lines, areas; cells
- Identifies how relationships will be represented

Stages in Conceptual Design

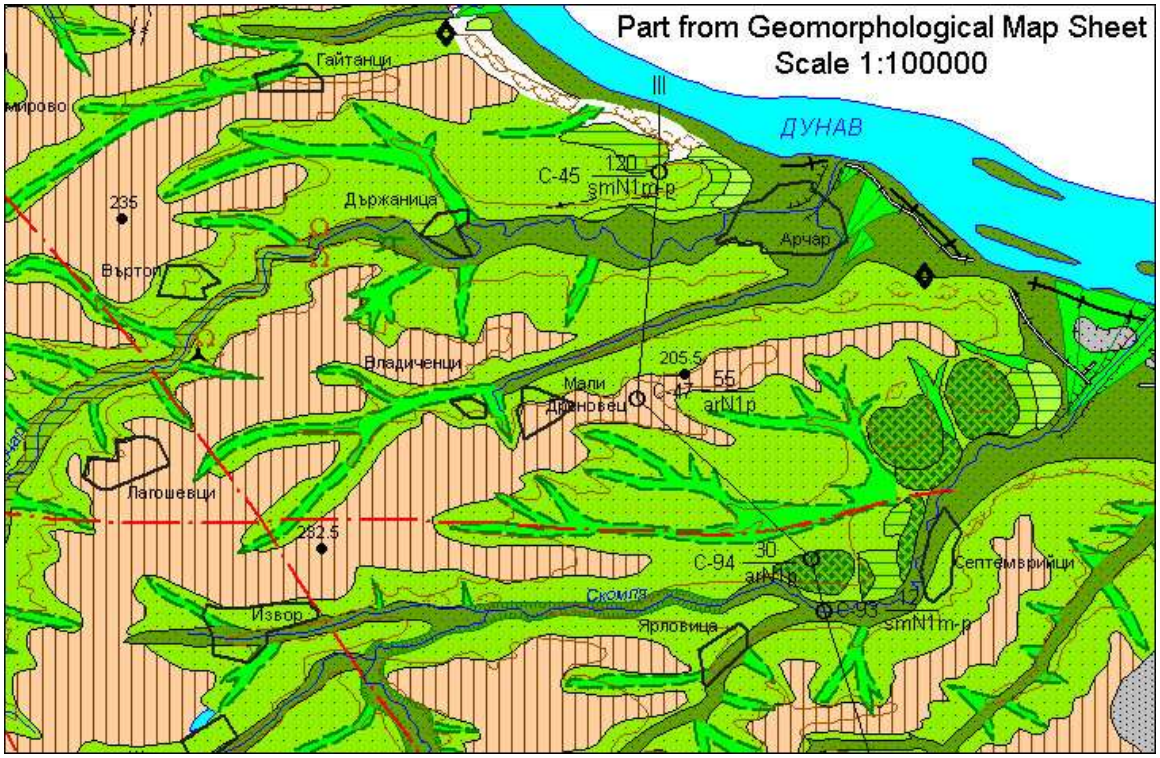
- Identify database issues
- Evaluate data sources
  - examine format
  - categorize : standards, coding
  - eliminate redundancy
  - prioritize
- Select data model
- Conceptualize table design
- Develop implementation plan
- Prepare documentation
- Present, review and approve

Geospatial Data

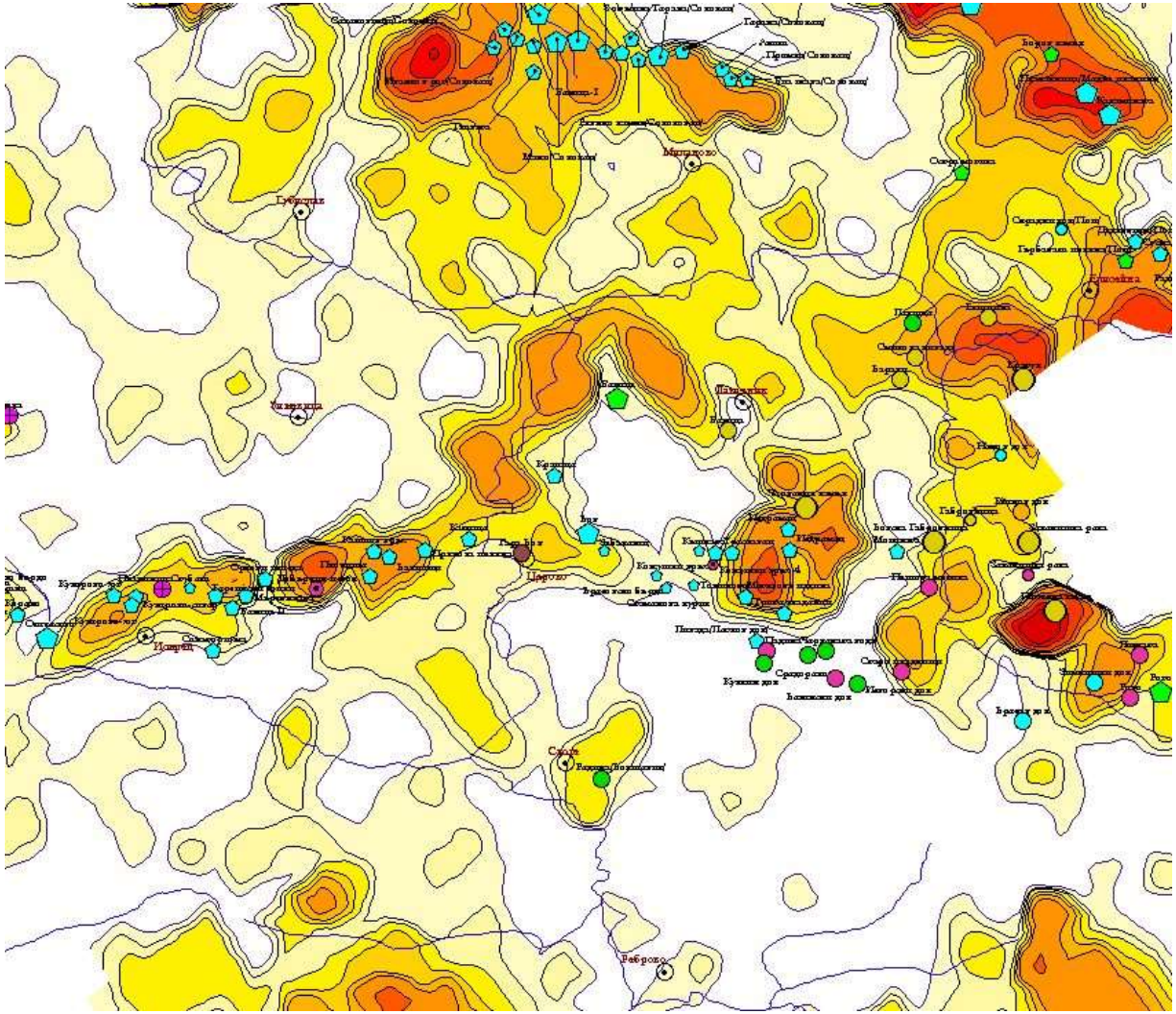
Geological Maps



Geomorphologic Map

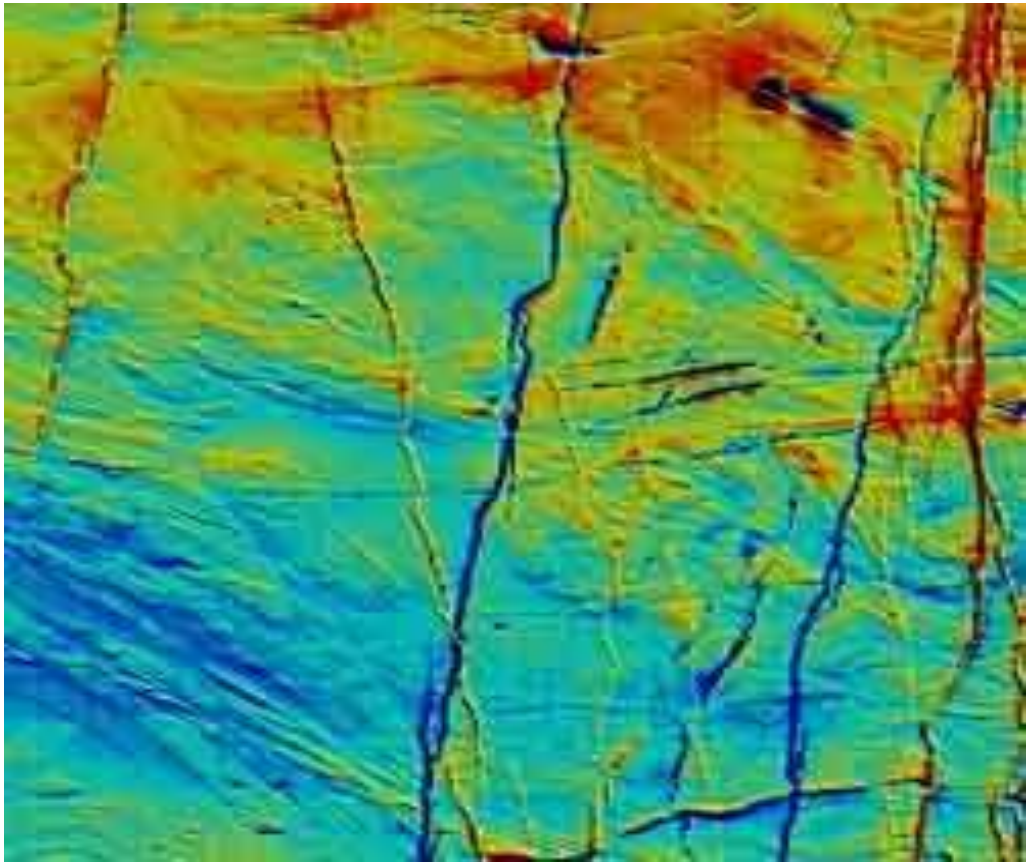


Geochemical Maps



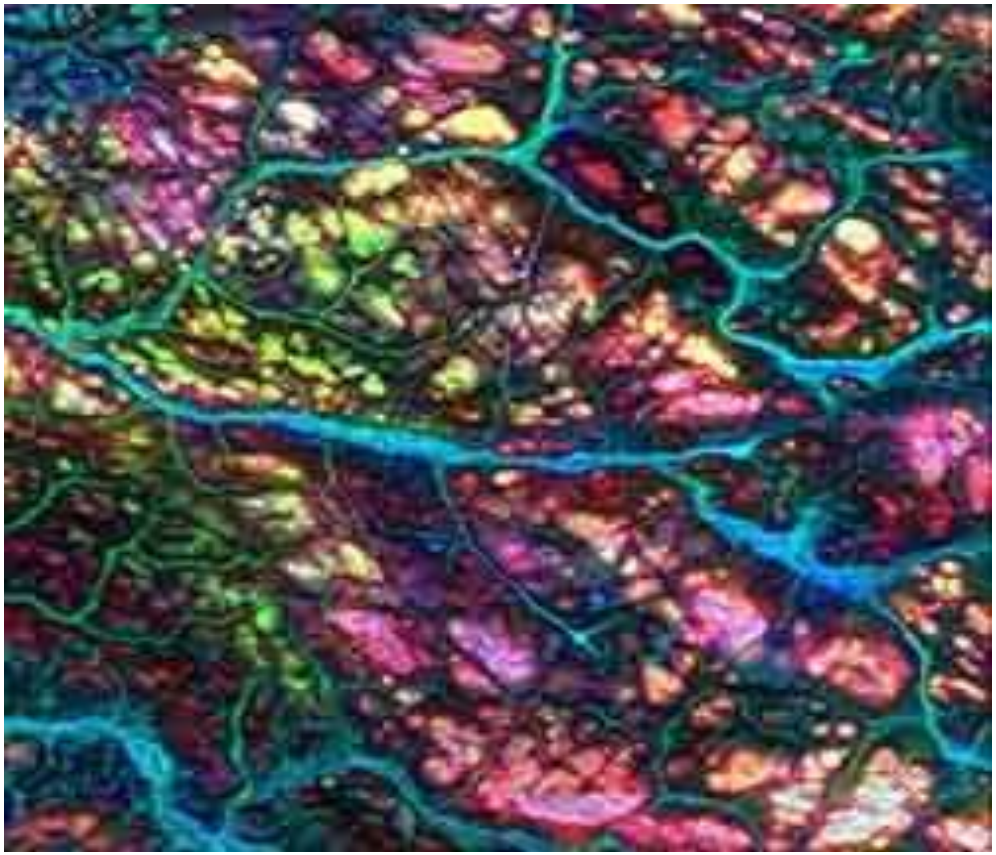
Geophysical Maps

High resolution magnetic map, Scale 1:100000



Geophysical Maps

Total count radiometric map, 1:100000.



Attribute Data

What is Attribute Data ?

- Spatial databases containing geographic features
- Attribute data is descriptive information about features
- It is stored as data items in tables
- Features may be classified
- Attribute data facilitates interpretation and analysis of the data set

RDBMS

- Almost all of the tabular data used in a GIS are stored in relational database tables
- RDBMS models are more flexible than the other similar database models
- Table structure is simple to understand and does not restrict any kind of query
- Relational database supports eight primary operations important for GIS
  - Restrict, project, union, intersection, difference, product,, join, and divide (Bolstad, 2005)

Attribute Characteristics

Attribute data defines the different characteristics of an object such as:

- Name
- Date
- Time
- Magnitude
- Value
- Nature
- Condition
- Number
- Ownership

Domains

- A set of values an attribute may have
- Includes type, length and size of data

Column	Domain	
	Data Type	Size
Customer_Id	Integer	20
Name	Character	30
Street	Character	30
City	Character	25
State	Character	2
Zip	Character	5

Column	Domain	
	Data Type	Size
Customer_Id	Integer	20
Account_Number	Integer	15
Account_Type	Character	2
Date_Opened	Date	
Balance	Real	12,2

Business Rules

Business rules specify constraints on the data that can appear in tables and the operations that can be performed on data in tables. For example:

- An account balance can never be negative.
- A Customer can not be deleted if they have an existing (open) account.
- Facilities must have at least one building.
- Building height must be > 0 and < 22.
- Age must be < 120

- How do we enforce business rules ?
  - Setting Properties in the database
  - Applications

RDBMS

- Outside of a GIS, tabular data are commonly held and manipulated in relational databases:
    - dBASE, rBase, ACCESS, Oracle, SQLServer, INFORMIX, DB2, Rational Rose
- Object-oriented Relational Database Management Systems...

Primary and Foreign Keys

- Primary key is the unique identifier for each record
  - It may consist of one or more attributes (columns)

- Each table should have a primary key to facilitate *indexing*

- A Foreign key is an attribute (item) which is included in a table and is a primary key in another table
  - The process is referred to as the *posting* of the key into the table
  - they facilitate the relating (linking) of tables
  - Basis upon which referential integrity is enforced
- Preferred Key Characteristics
  - short length
  - numeric
  - non-volatile

Database Tables

Forests

Forest Name	Forest-ID	Location	Size
Nantahala	1	N. Carolina	184,447
Cherokee	2	N. Carolina	92,271

Trails

Trail Name	Forest-ID
Bryson's Knob	1
Slickrock Falls	2
North Fork	1
Cade's Cove	1
Cade's Cove	2
Appalachian	1
Appalachian	2

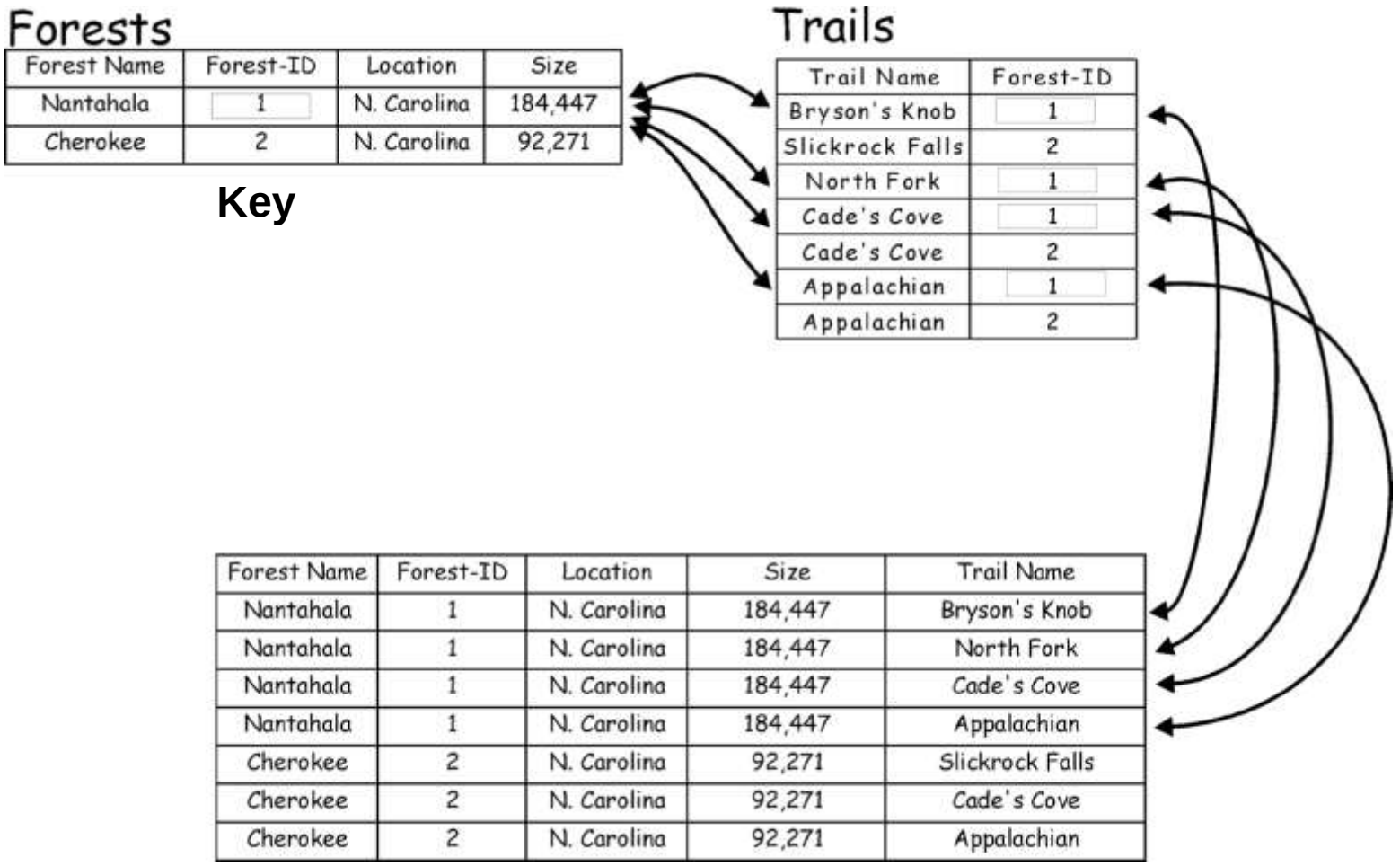
Characteristics

Trail Name	Feature	Difficulty
Bryson's Knob	Vista	E,M
Bryson's Knob	Ogrth	E,M
Slickrock Falls	Ogrth	M
Slickrock Falls	Wfall	M
North Fork	-	M
Cade's Cove	Ogrth	E
Cade's Cove	Wlife	E
Appalachian	Wfall	M,D
Appalachian	Ogrth	M,D
Appalachian	Vista	M,D
Appalachian	Wlife	M,D
Appalachian	Cmp	M,D

Recreational features

Feature	Description	Activity1	Activity2
Wfall	Waterfall	Photography	Swimming
Ogrth	Old-Growth Forest	Photography	Hiking
Vista	Scenic Overlook	Photography	Viewing
Wlife	Wildlife Viewing	Photography	Birding
Cmp	Camping	Camping	-

Relationships



Bolstad, 2005

RDBMS Supporting Functions in GIS

a) restrict

ID	type	color	size	age
1	a	blue	big	old
2	c	green	big	young
3	a	red	small	mid
4	d	black	big	older
5	x	mauve	tiny	oldest
6	g	dun	huge	young
7	c	ecru	small	mid

restrict →

ID	type	color	size	age
1	a	blue	big	old
4	d	black	big	older
6	g	dun	huge	young
2	c	green	big	young

b) project

ID	type	color	size	age
1	a	blue	big	old
2	c	green	big	young
3	a	red	small	mid
4	d	black	big	older
5	x	mauve	tiny	oldest
6	g	dun	huge	young
7	c	ecru	small	mid

project →

ID	color	size
1	blue	big
2	green	big
3	red	small
4	black	big
5	mauve	tiny
6	dun	huge
7	ecru	small

c) product

No.	Dir.
1	N
2	S

product

App.
Yes
Yes
No

→

No.	Dir.	App.
1	N	Yes
2	S	Yes
1	N	No
2	S	No

d) divide

type
m
n
r

divide by

size
1
2

per

type	size
m	1
m	2
m	3
m	4
n	2
r	1
r	3

→

type
m

a) union

ID	type	color	size	age
1	a	blue	big	old
6	g	dun	huge	young

ID	type	color	size	age
2	c	green	big	young
4	d	black	big	older

union →

ID	type	color	size	age
1	a	blue	big	old
4	d	black	big	older
6	g	dun	huge	young
2	c	green	big	young

b) intersect

ID	color	size
1	blue	big
2	green	big
3	red	small
4	black	big
5	mauve	tiny
6	dun	huge
7	ecru	small

ID	color	size
1	blue	big
5	mauve	tiny
9	ivory	big

intersect →

ID	color	size
1	blue	big
5	mauve	tiny

c) difference

ID	color	size
1	blue	big
2	green	big
3	red	small
4	black	big
5	mauve	tiny
6	dun	huge
7	ecru	small

ID	color	size
1	blue	big
5	mauve	tiny
9	ivory	big

difference →

ID	color	size
2	green	big
3	red	small
4	black	big
6	dun	huge
7	ecru	small

d) join

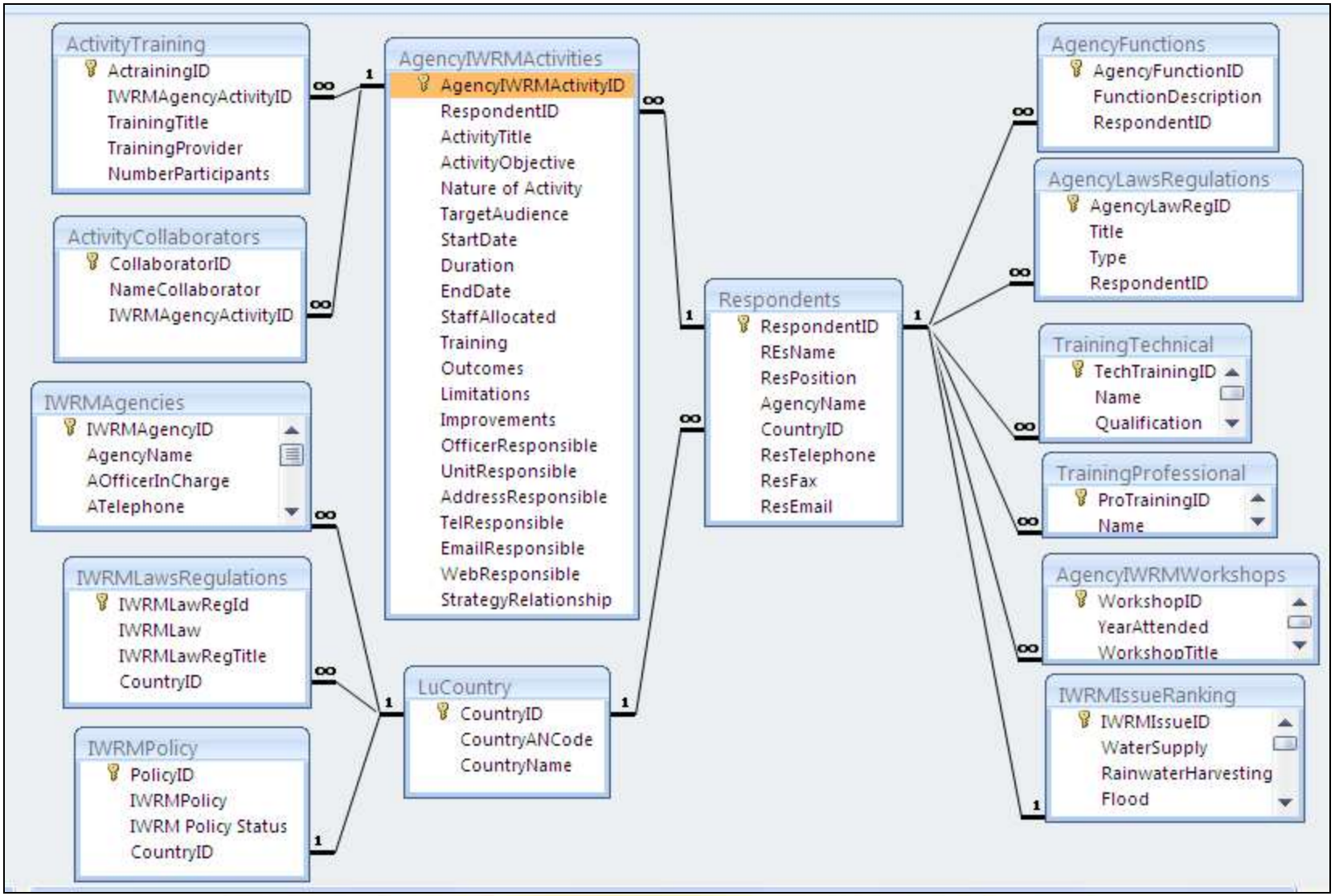
ID	type
1	a
2	b
3	b
4	a

type	color	size	age
a	blue	big	old
b	dun	tiny	old

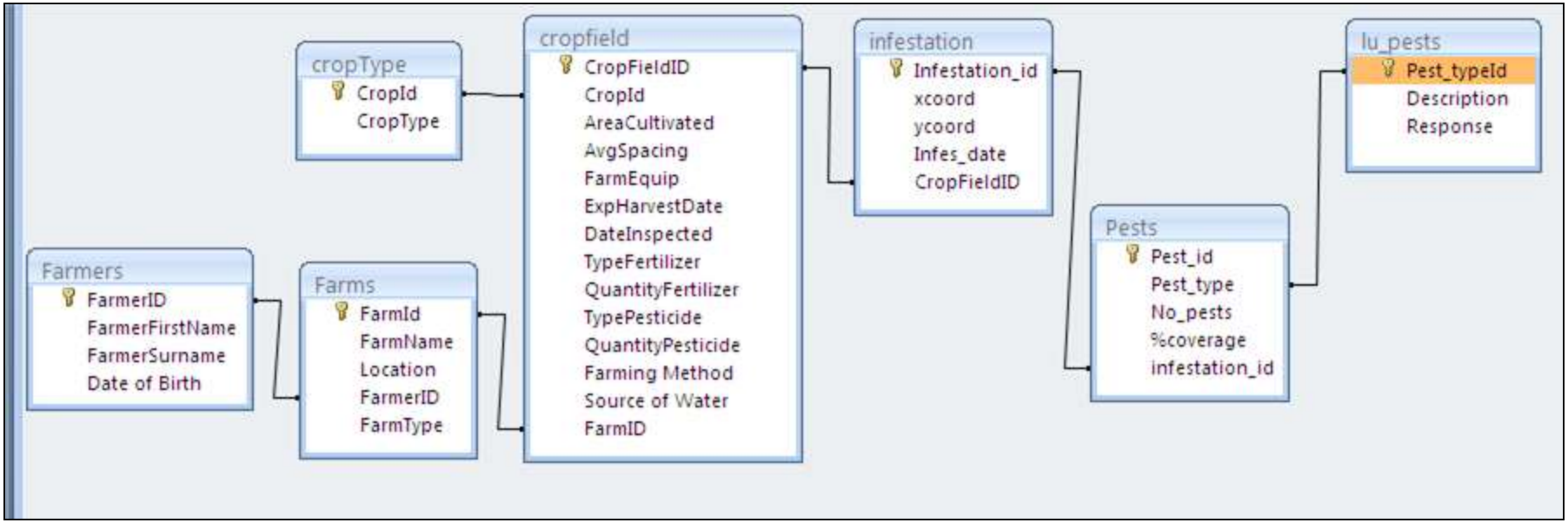
join →

ID	type	color	size	age
1	a	blue	big	old
2	b	dun	tiny	old
3	b	dun	tiny	old
4	a	blue	big	old

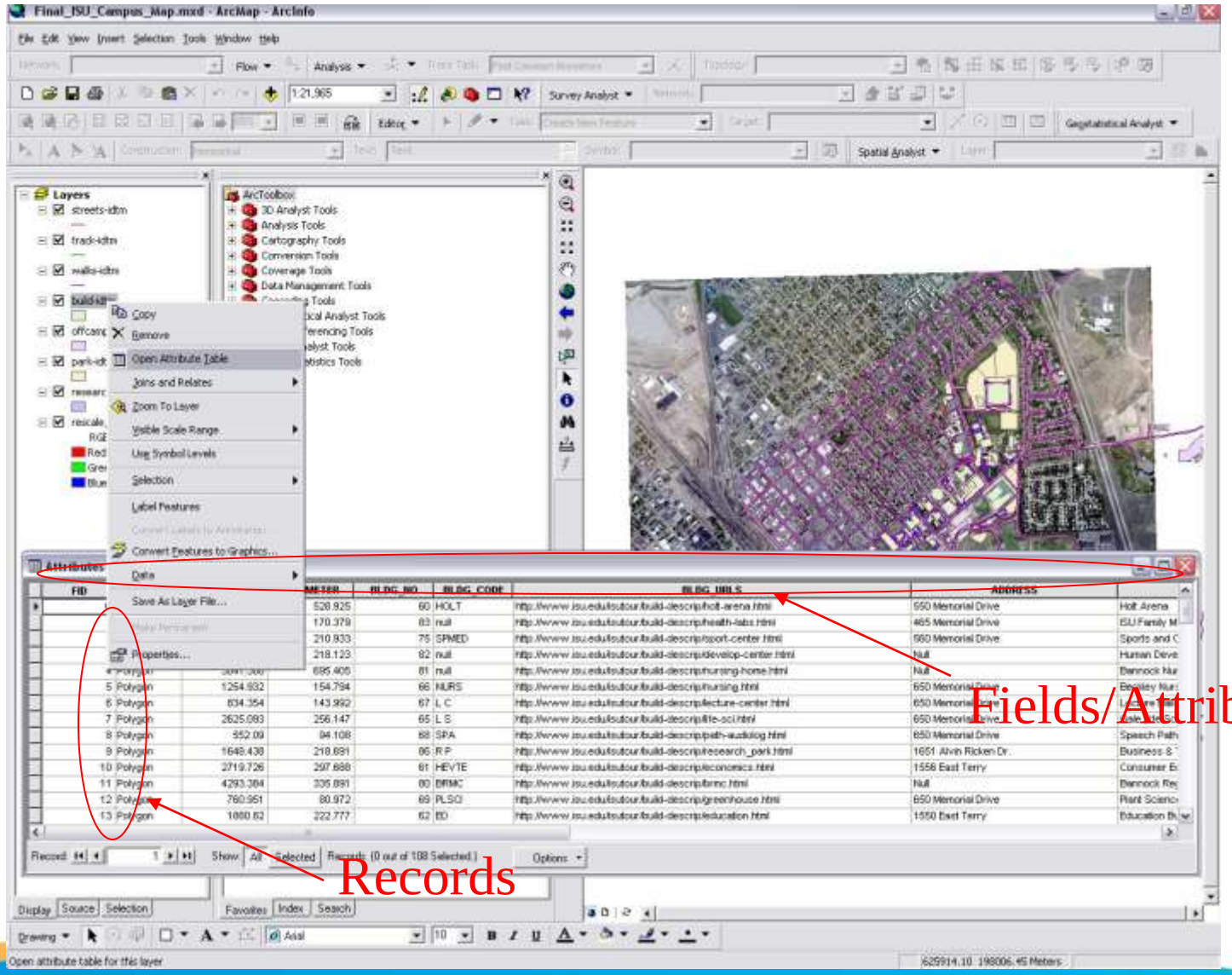
IWRM EAR


ActivityTraining
ActrainingID
IWRMAgencyActivityID
TrainingTitle
TrainingProvider
NumberParticipants
ActivityCollaborators
CollaboratorID
NameCollaborator
IWRMAgencyActivityID
IWRMAgencies
IWRMAgencyID
AgencyName
AOfficerInCharge
ATelephone
IWRMLawsRegulations
IWRMLawRegId
IWRMLaw
IWRMLawRegTitle
CountryID
IWRMPolicy
PolicyID
IWRMPolicy
IWRM Policy Status
CountryID
AgencyIWRMActivities
AgencyIWRMActivityID
RespondentID
ActivityTitle
ActivityObjective
Nature of Activity
TargetAudience
StartDate
Duration
EndDate
StaffAllocated
Training
Outcomes
Limitations
Improvements
OfficerResponsible
UnitResponsible
AddressResponsible
TelResponsible
EmailResponsible
WebResponsible
StrategyRelationship
LuCountry
CountryID
CountryANCode
CountryName
Respondents
RespondentID
REsName
ResPosition
AgencyName
CountryID
ResTelephone
ResFax
ResEmail
AgencyFunctions
AgencyFunctionID
FunctionDescription
RespondentID
AgencyLawsRegulations
AgencyLawRegID
Title
Type
RespondentID
TrainingTechnical
TechTrainingID
Name
Qualification
TrainingProfessional
ProTrainingID
Name
AgencyIWRMWorkshops
WorkshopID
YearAttended
WorkshopTitle
IWRMIssueRanking
IWRMIssueID
WaterSupply
RainwaterHarvesting
Flood

Pest Monitoring EAR



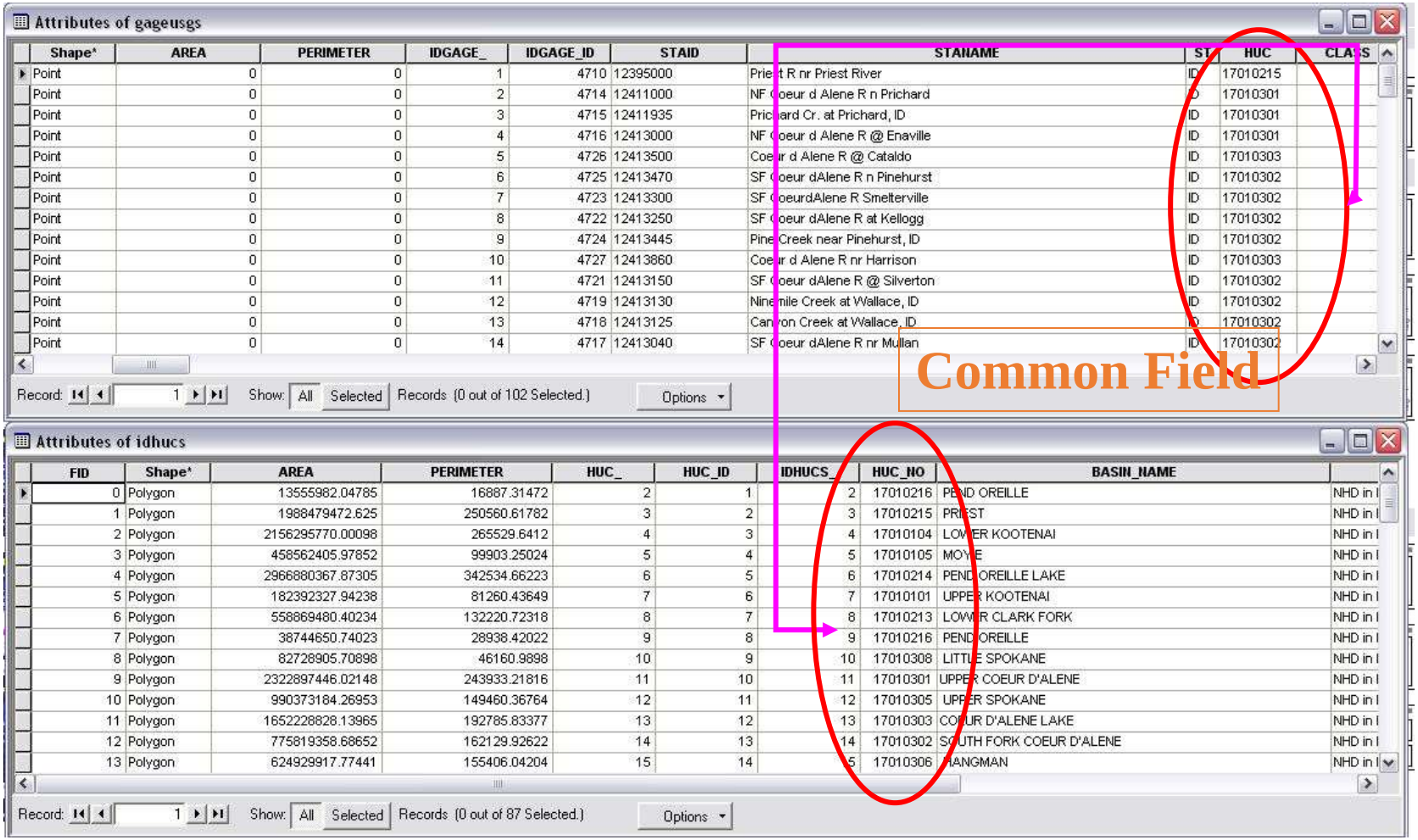
Attribute Tables: Database in ArcGIS



Some Facts about Attribute Table

- columns : fields/attributes
- rows : records
- Automatic Fields:
  - FID and Shape fields automatically created
  - During Digitizing:
    - Length of lines
    - Area & Perimeter of Polygons
- You can add new fields and values to an existing attribute table
- Attribute tables are saved in workspace as *.dbf file
- You can export your new updates in attribute file in many formats

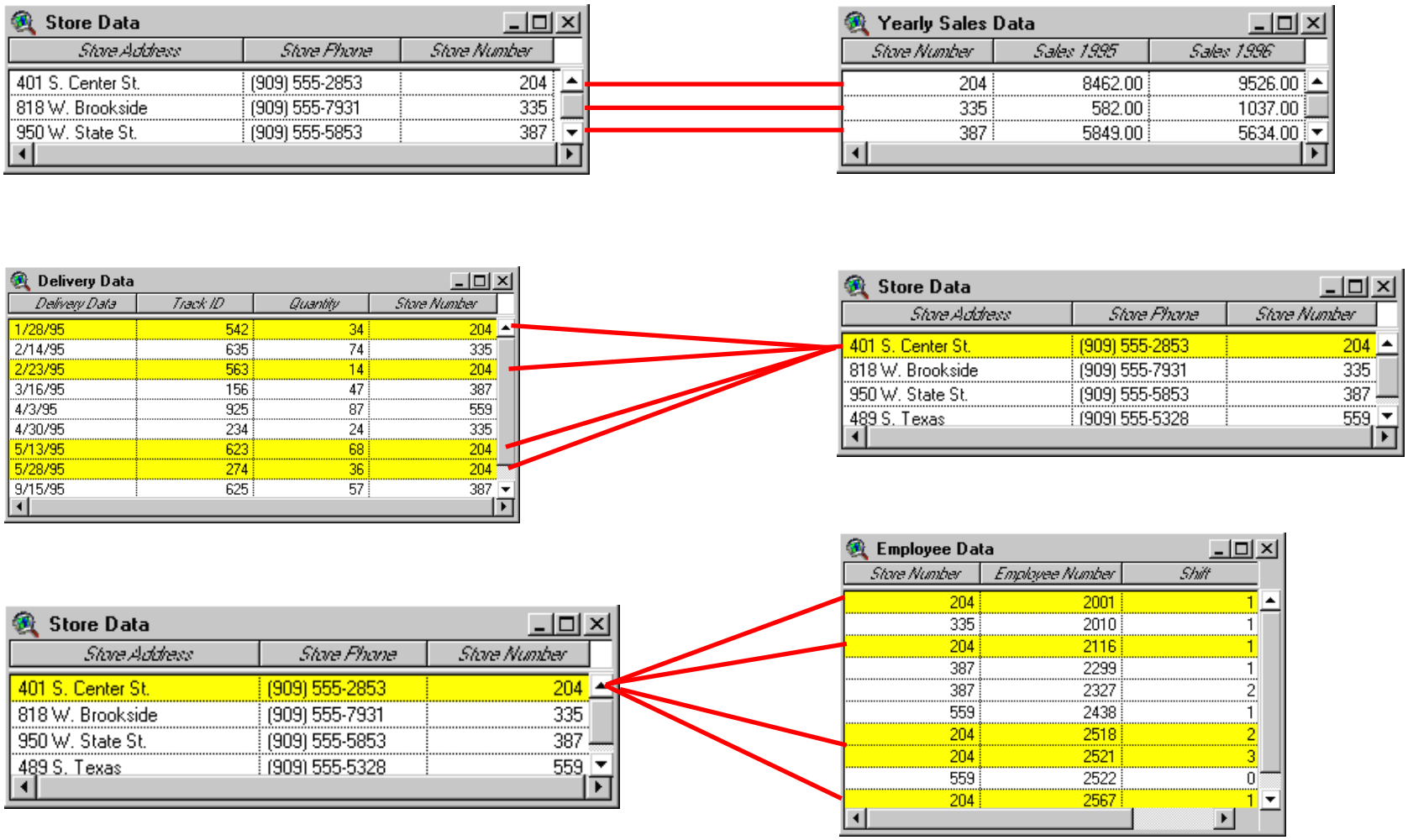
Linking Tables



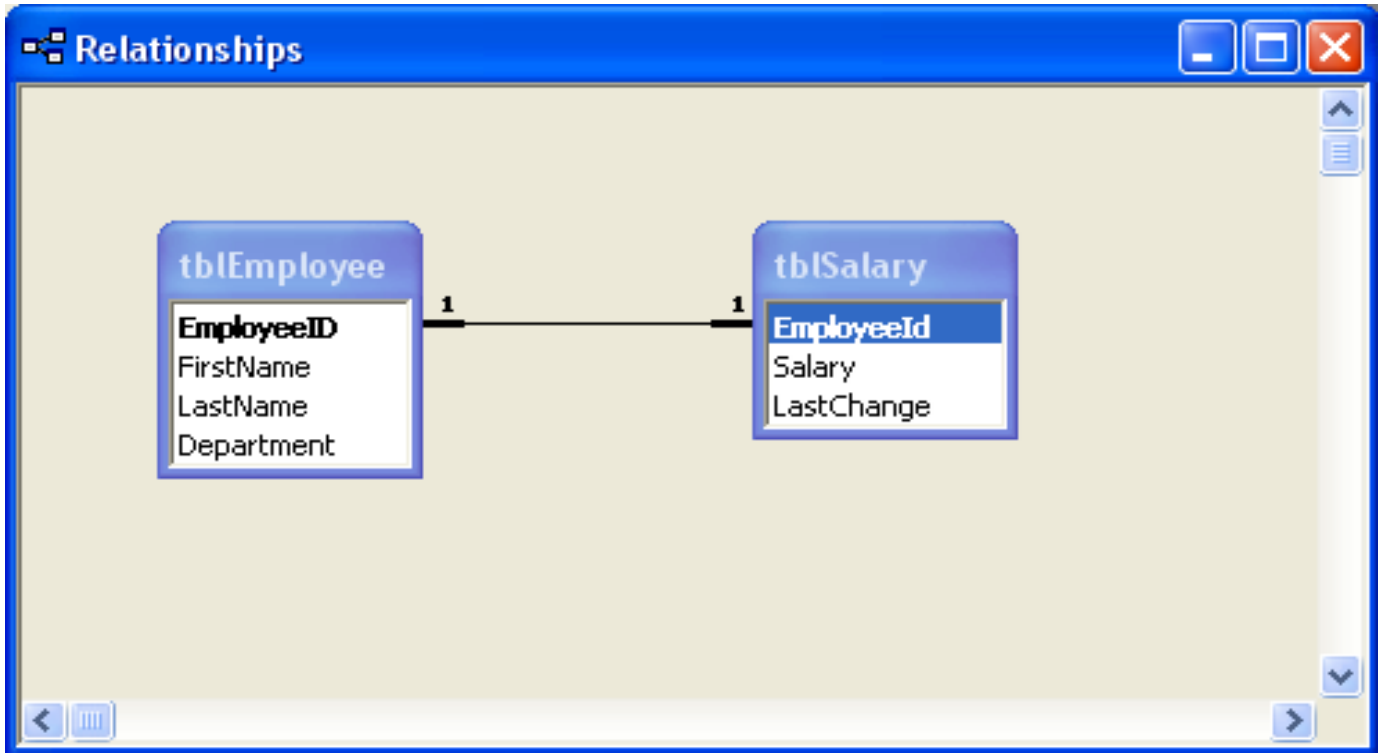
Relate Table

- Relate function lets you associate data (as we discussed) with the layer
- The data is not appended into the layer as in the case of join
- But, the data layer of the associated table can be accessed when you work with the other table
- Establishing a relate is useful when there is 1-to-many and many-to-many association between layer and the related data

Table Record Relationships Cardinality

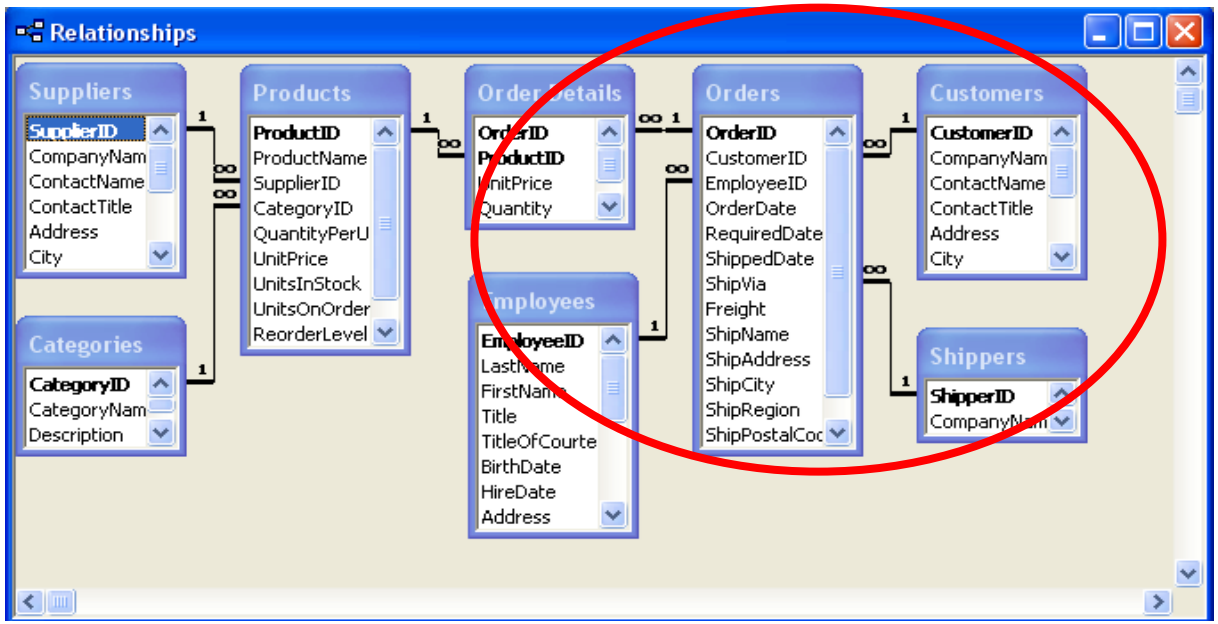


One-to-One Relationships



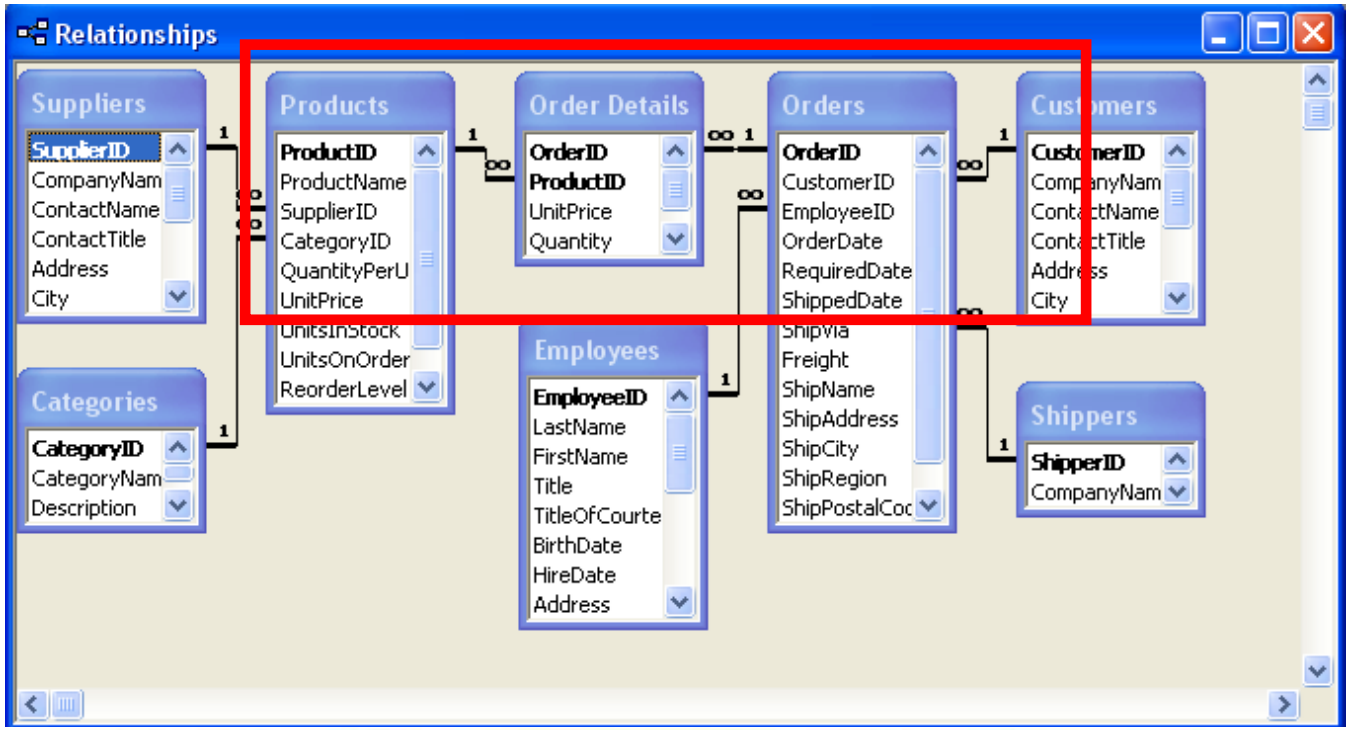
One-to-Many Relationships

- ✔ Most common type of relationship
- ✔ Related between primary and foreign keys



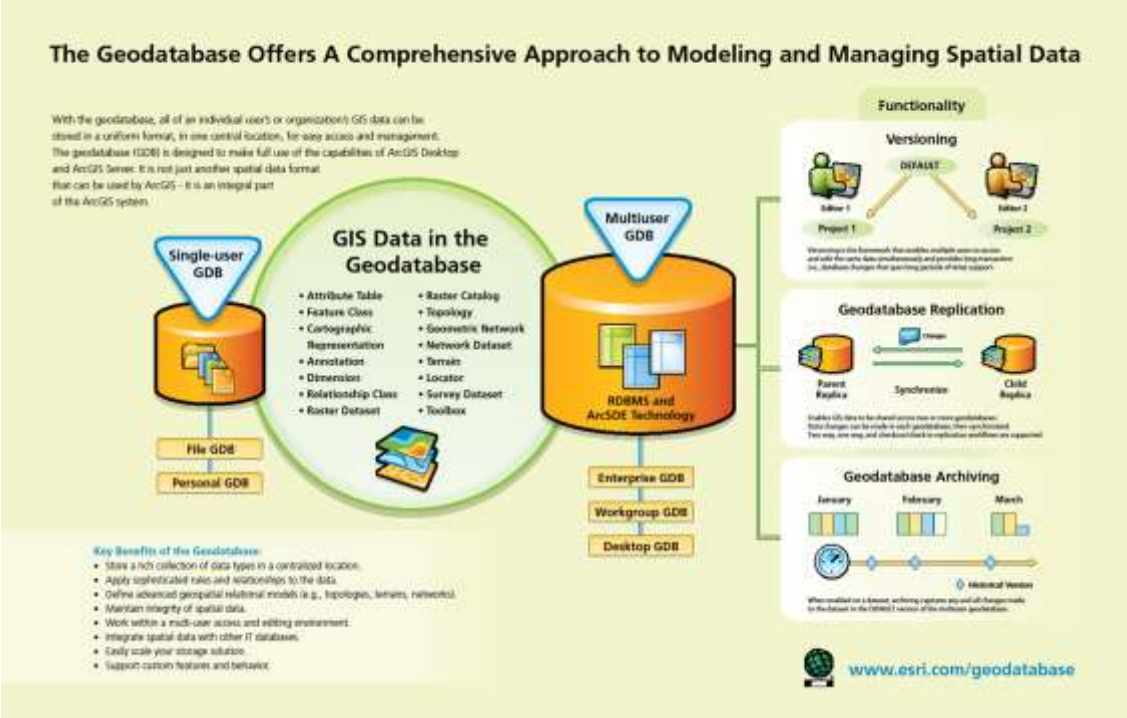
Many-to-Many Relationships

- ✔ Not directly supported between tables
- ✔ Use a junction table to relate
- ✔ One order, many products
- ✔ One product, many orders



Geodatabases in ArcGIS

- A structured set of data
- managed and stored as one unit
- generally associated with a software to update and query the data
- geodatabase = an object-oriented data model introduced by ESRI



Geodatabases

- Hosted inside of a relational database management system
- Storage for features classes, feature datasets, non-spatial tables,…
- Relationships (spatial, non-spatial) - > advantage over shapefiles

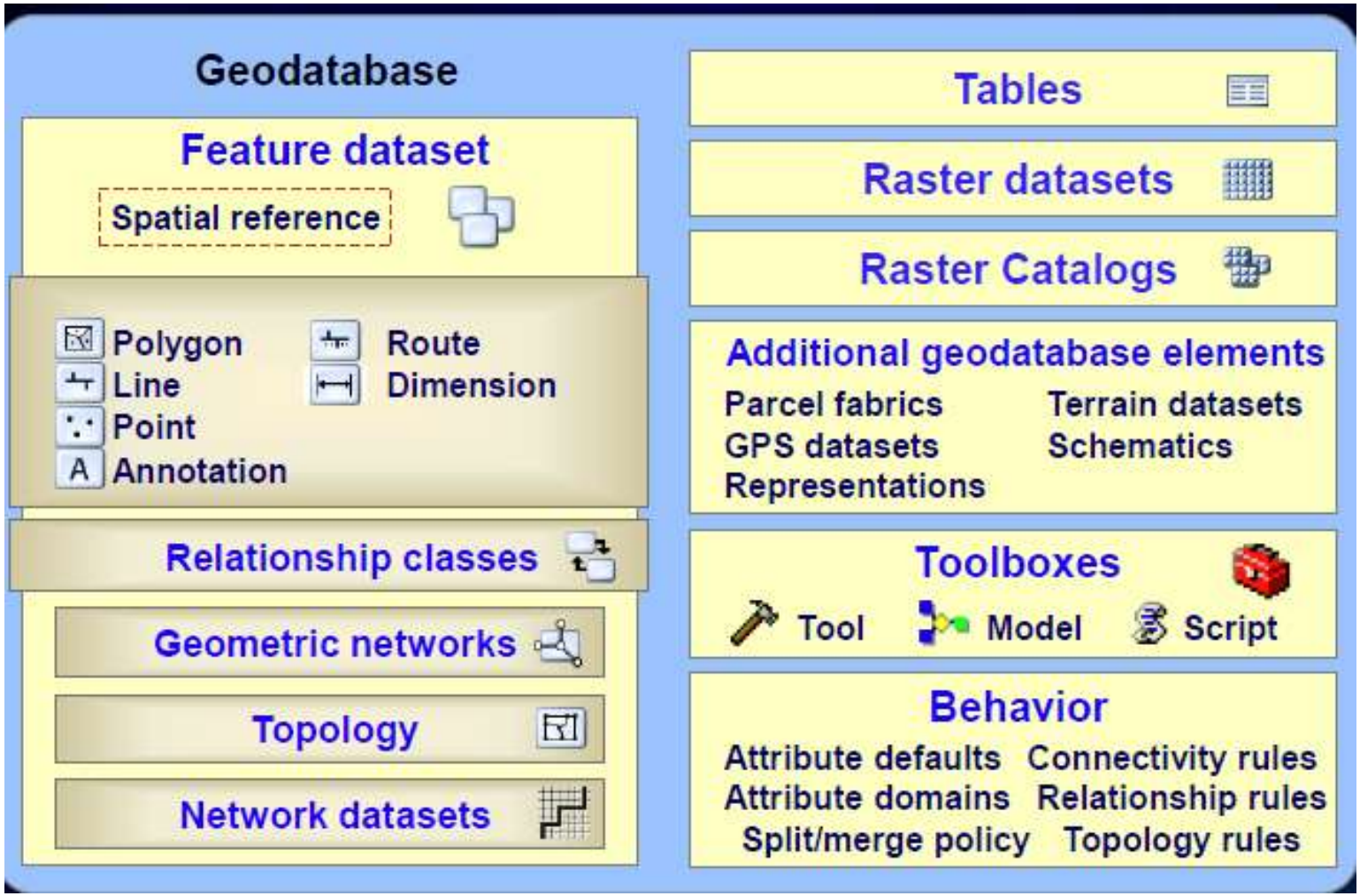
Geodatabases

- **File geodatabase**
  - stores datasets in a folder of files
  - each dataset a file up to 1 TB in size
  - can be used across platforms
  - can be compressed and encrypted for read-only, secure use
  - ESRI's recommended choice

Geodatabases

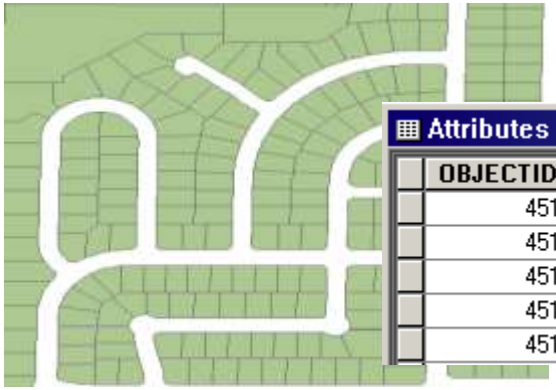
- **Personal geodatabase**
  - stores datasets in a Microsoft Access .mdb file
  - storage sizes between 250 and 500 MB
  - limited to 2GB
  - only supported on Windows
- **ArcSDE geodatabase**
  - stores datasets in a number of optional DBMSs:
    - IBM DB2, IBM Informix , Microsoft SQL Server , Oracle, or PostgreSQL
  - unlimited size and users

Geodatabases



Geodatabases

- Features and Feature classes - collections of features (spatial objects) with the same type of geometry and the same attributes
  - Holds geometry attributes
    - rows features, table feature class

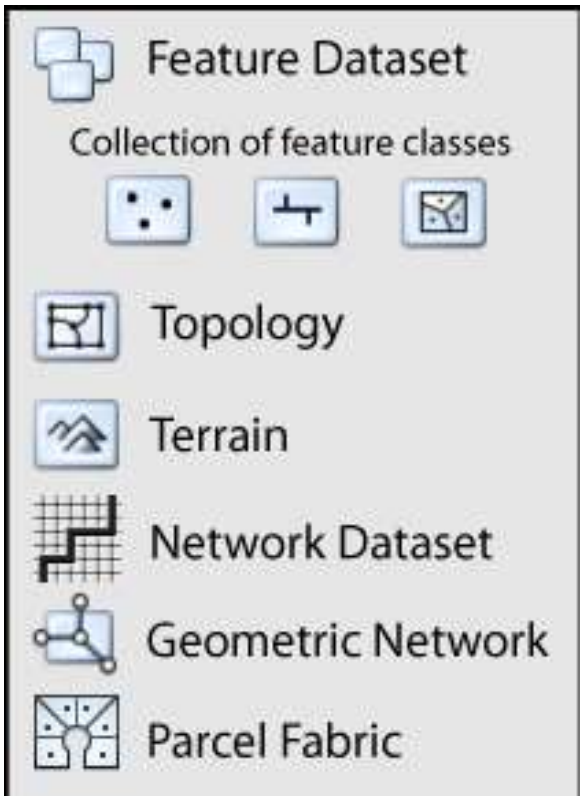


Attributes of Parcels

OBJECTID*	SHAPE*	PARCEL_ID*	ZONE_CODE*	SHAPE_Length	SHAPE_Area
4513	Polygon	67970	W	544.053559	9259.209935
4514	Polygon	67971	W	158.545394	774.602847
4515	Polygon	67973	R60M	400.003008	7499.965473
4516	Polygon	67974	B1	236.126101	2905.890606
4517	Polygon	67982	B1	550.458538	17499.011493

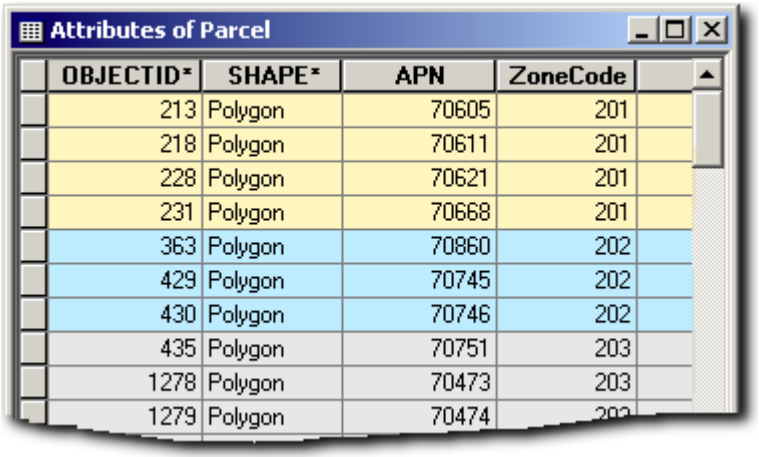
Geodatabases

- Feature datasets - collections of feature classes that share the same spatial reference
  - Topological rules must be inside Feature datasets



Geodatabases

- Subtypes
  - Categorize objects or features into groups
    - Share same attributes
  - Defined at the class level
  - Select a field to base the subtype on
    - Short or long integer field
    - Can have different default values and domains
    - Can define behavior rules



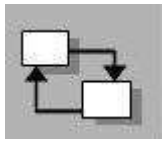
OBJECTID*	SHAPE*	APN	ZoneCode
213	Polygon	70605	201
218	Polygon	70611	201
228	Polygon	70621	201
231	Polygon	70668	201
363	Polygon	70860	202
429	Polygon	70745	202
430	Polygon	70746	202
435	Polygon	70751	203
1278	Polygon	70473	203
1279	Polygon	70474	203

Geodatabases

- Relationship classes – define relationships between objects (one-to-one, one-to-many, many-to-many)
- Related objects communicate (cascade delete)

Parcel

OID	Zone	Area
28	Commercial	10000
794	Residential	5000
858	Residential	6050



Buildings

OID	Parcel_ID	Building Type
1	28	Office Building
2	794	Townhouse
3	794	Townhouse
4	858	Condo

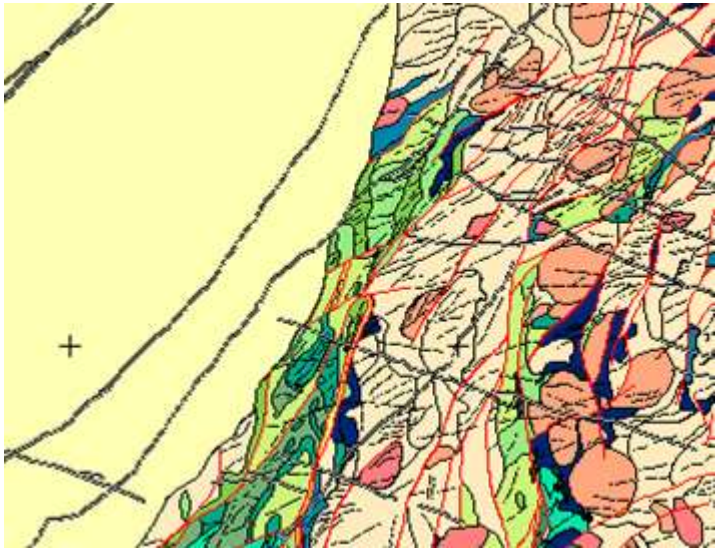
Geodatabases

- Object behavior
  - Control default value and acceptable values for any attributes (Domains)
  - Subdivide objects into similar groups (Subtypes)
  - Control the spatial relationships (Topology)

- Annotation
  - Placing text and graphic on a map
  - Linked to features or not
  - Can store text as well as other graphics

Geodatabases

- Topology
  - Constraints on the shared geometry
  - Define data integrity
  - Control on editing tools
  - Validation of features
  - Ensures quality of your data

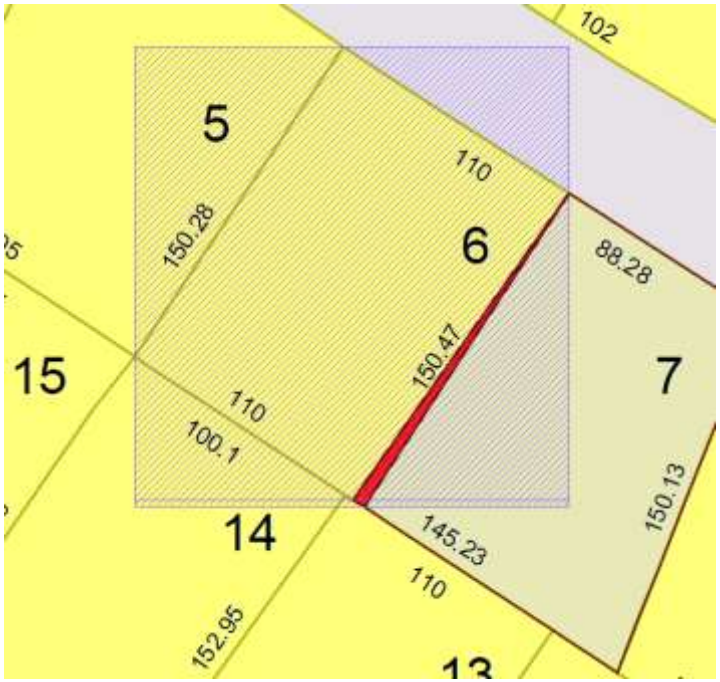


Geodatabases

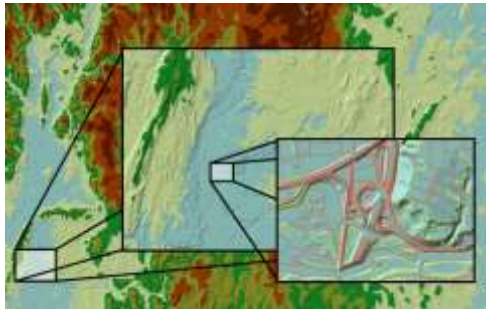
- Create topology in a feature dataset
  - Participating feature classes
  - Cluster tolerance
  - Ranks and rules
- Define rules
  - Rules evaluated during validation
- Errors
  - Managed as a part of the topology
  - Exceptions
  - Examine and fix errors in ArcMap

Geodatabases

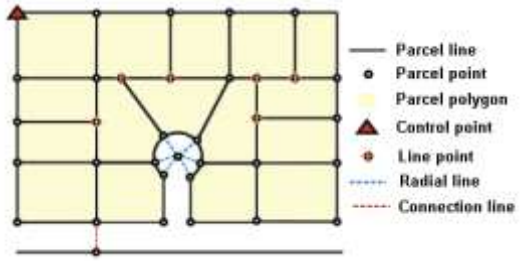
- Editing with topology
  - Editing creates dirty area
    - Can be symbolized
- Errors found during validation
  - Have properties
    - What rule is concerned
    - Which features are involved
- Options
  - Ignore error
  - Mark as exception
  - Fix



Geodatabases



- Other elements included
  - Terrains
    - Point datasets for modeling 3D surfaces
    - Requires 3D Analyst extension
  - Representations
    - Properties of feature classes storing symbologies
      - Rules and overrides
  - Parcel fabrics
    - Designed for parcel data management



Geodatabases

- Directly loadable data types
  - dBase (.dbf)
  - Text with comma (.csv) or tab-separated values (.txt)
  - Microsoft Access (.mdb)
  - Microsoft Excel (.xls)

What made all of these possible?

Access to

- **Technology (currency)**
- **Data (currency, resolution, accuracy)**
- **Policy (relevance, monitoring)**
- **Human resources (skills and benefits)**
- **Funding (driven by applications and results)**
- **Political support (motivation for success)**

Thank you!

- ----------END----------