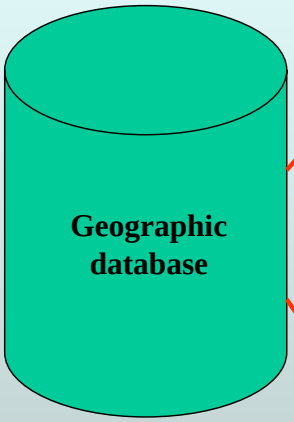
LESSON 2

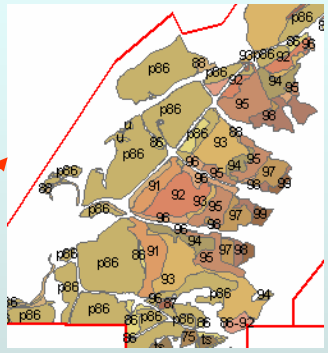
Working with Data

Geographic Database

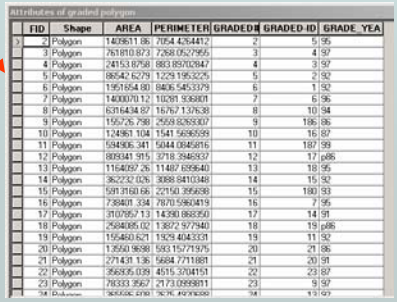
A geographic database is the core of a GIS and its completeness and accuracy affects all the applications it supports. The database is a collection of information describing the spatial and descriptive attributes of real-world entities.



Spatial



Attribute

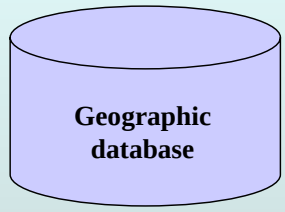


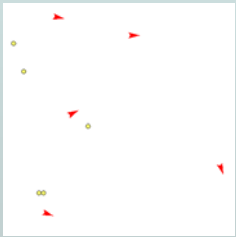
Attributes of graded polygon

FID	Shape	AREA	PERIMETER	GRADEDI	GRADED-ID	GRADE_YEA
2	Polygon	1409611.86	7054.4264412	2	5	95
3	Polygon	761810.873	7268.0527955	3	4	97
4	Polygon	24153.8758	883.89702847	4	3	97
5	Polygon	86542.6279	1229.1953225	5	2	92
6	Polygon	1951654.80	8406.5453379	6	1	92
7	Polygon	1400070.12	10281.936801	7	6	96
8	Polygon	6316434.87	16767.137638	8	10	94
9	Polygon	155726.798	2559.8269307	9	186	86
10	Polygon	124961.104	1541.5696599	10	16	87
11	Polygon	594906.341	5044.0845816	11	187	99
12	Polygon	809341.915	3718.3946937	12	17	p86
13	Polygon	1164097.26	11487.699640	13	18	95
14	Polygon	362232.026	3088.8410348	14	15	92
15	Polygon	5913160.66	22150.395698	15	180	93
16	Polygon	738401.334	7870.5960419	16	7	95
17	Polygon	3107857.13	14390.868350	17	14	91
18	Polygon	2584085.02	13872.977940	18	19	p86
19	Polygon	155460.621	1929.4043331	19	11	92
20	Polygon	13550.9698	593.15771975	20	21	86
21	Polygon	271431.136	5684.7711881	21	20	91
22	Polygon	356935.039	4515.3704151	22	23	87
23	Polygon	78333.3567	2173.0999811	23	9	97
24	Polygon	[illegible]	[illegible]	24	13	97

Abstracting real-world entities

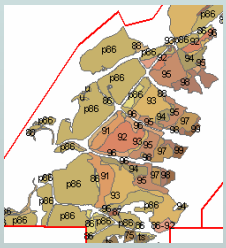
•Real-world entities can be abstracted into three basic shapes





Points
(drillholes/photo pts.)

Lines/Arcs
(roads/streams)

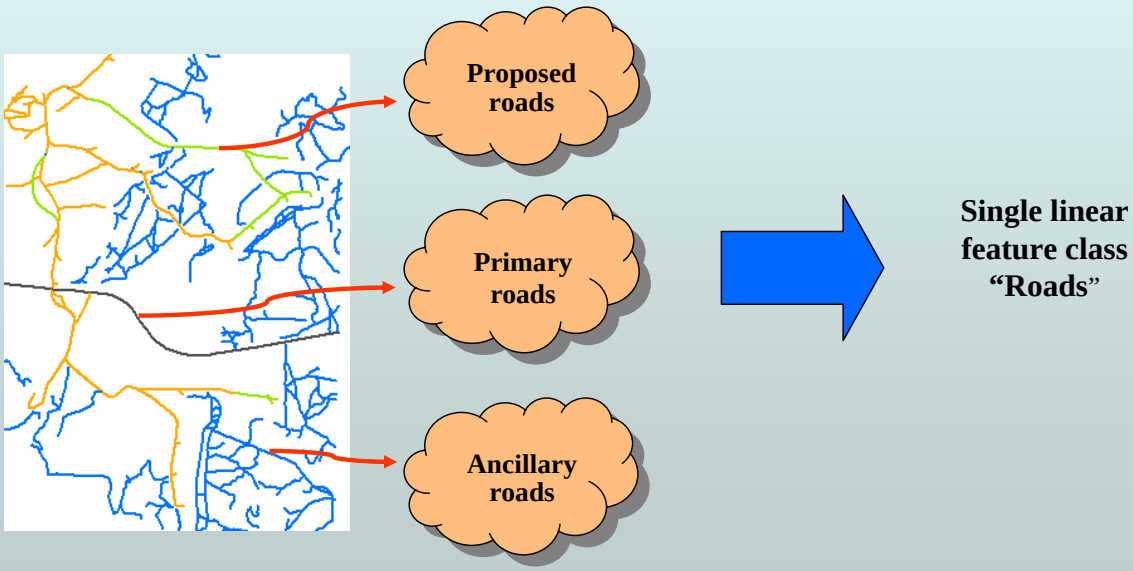


Areas/Polygons
(Land condition/Permit boundaries)

Feature Classes

(Geodatabase)

- **Group objects into homogeneous units**
- **Can create point, line, or area feature classes**



Feature Data Set
(Geodatabase)

- **Must have same coordinate system**
- **May have same geographic extent**
- **Should group feature classes based on topology or coincident geometry**

Features	Feature classes	Feature Data Set
Primary roads Ancillary roads Proposed roads	**Roads**	**XYZ Mine**
Permit boundaries	**Permits**	
Federal control State control Private control	**Coal Ownership**	

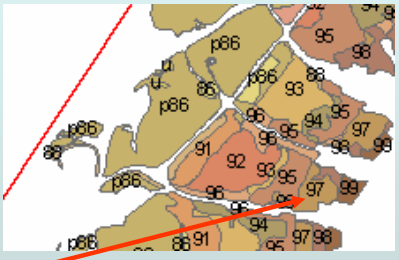
An ArcInfo coverage stores a collection of feature classes and can be called a file based feature dataset (as opposed to a geodatabase feature dataset)

Linking Features and Attributes

- **Feature classes have an associated table**
- **Each feature has a record in the table**
- **A unique identifier links feature and attributes**

Attributes of graded polygon

FID	Shape	AREA	PERIMETER	GRADED#	GRADED-ID	GRADE_YEA
2	Polygon	1409611.86	7054.4264412	2	5	95
3	Polygon	761810.873	7268.0527955	3	4	97
4	Polygon	24153.8758	883.89702847	4	3	97
5	Polygon	86542.6279	1229.1953225	5	2	92
6	Polygon	1951654.80	8406.5453379	6	1	92
7	Polygon	1400070.12	10281.936801	7	6	96
8	Polygon	6316434.87	16767.137638	8	10	94
9	Polygon	155726.798	2559.8269307	9	186	86

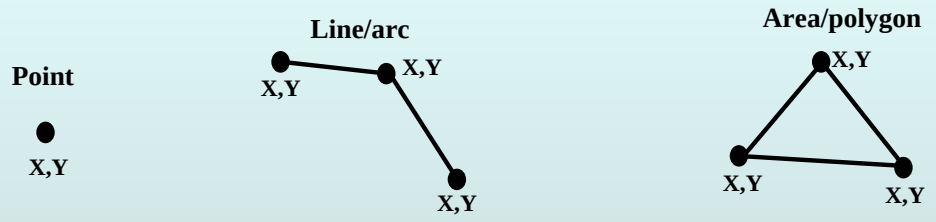


FID = 4

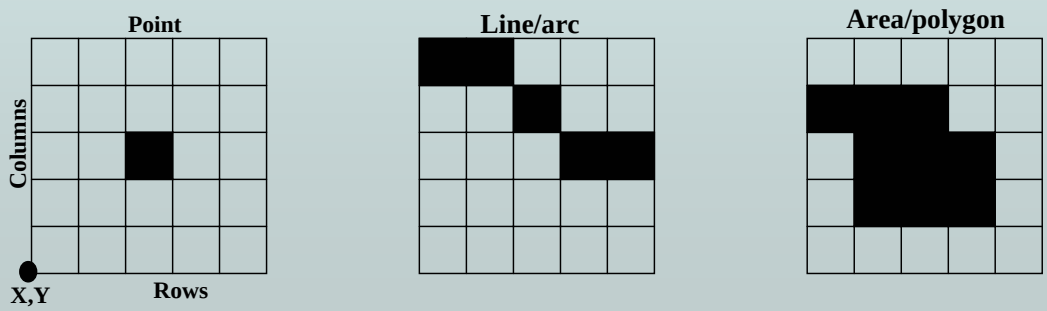
(Feature Identifiers)

Storing Abstracted Objects

- **Two primary methods for digital storage**
  - **Vector formats discretely identify shape coordinates**



  - **Raster formats assign square cells to real world entities**

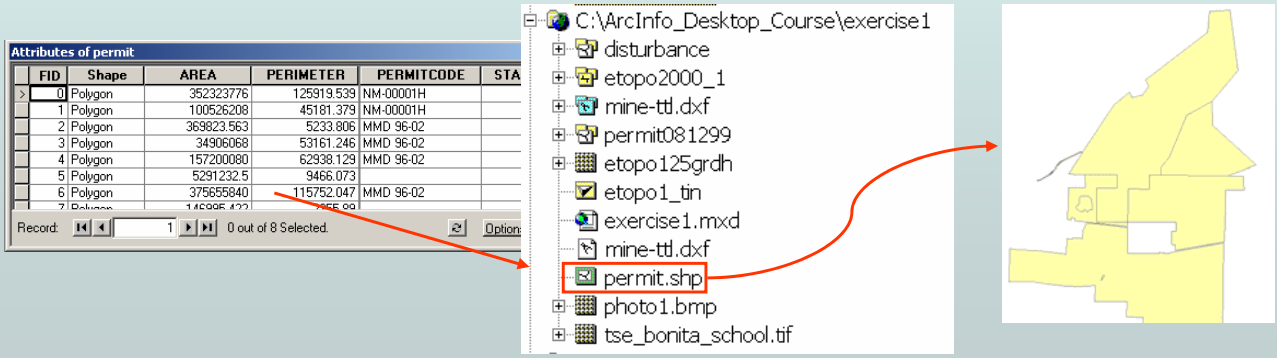


Vector Data Formats

- **Shapefile - ArcView 3.x format**
- **Coverage - ArcInfo 8.x and before**
- **Geodatabase - ArcGIS/ArcSDE**
- **CAD files - AutoCAD**
- **Event tables - ASCII/TEXT format**

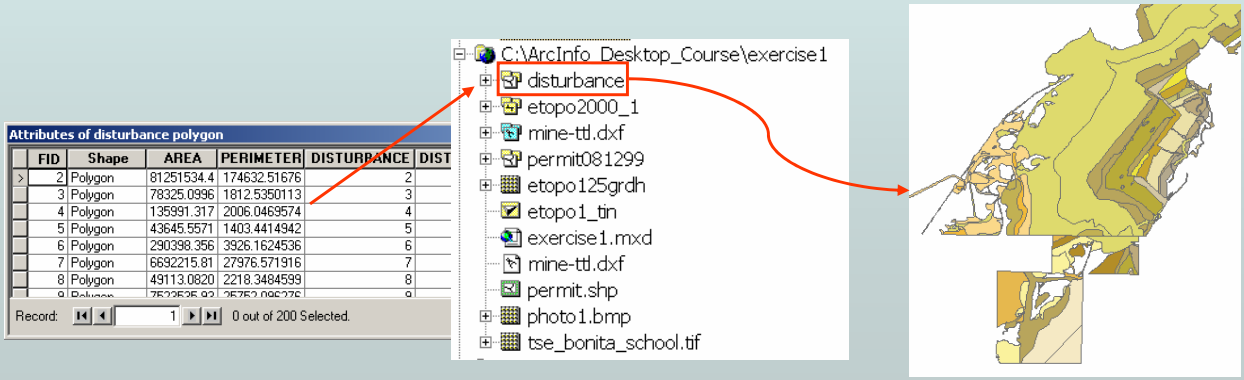
Shapefile

- **Single feature class - can only contain one feature class**
- **Attributes stored in dBASE table: *.dbf e.g. soils.dbf**
- **Each shapefile consists of at least three files: *.shp, *.shx, and *.dbf.**
- **If the shapefile has a defined coordinate system then the spatial reference information will be stored in a *.prj file**
- **In addition to the above four files other files may be created and/or used by ArcInfo as needed**



Coverage

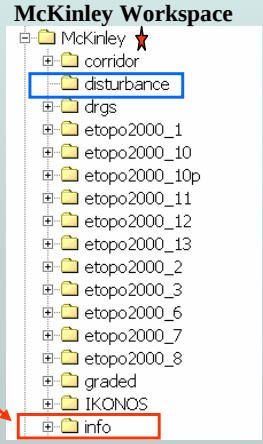
- **Stored as a directory containing both the locational data files and the descriptive data files**
- **A collection of feature classes**
  - **Stored in a coverage directory**
  - **Line and/or polygon topology**
- **Attributes stored in a separate INFO table**
- **INFO attribute tables associated to line feature classes are named *.AAT**
- **INFO attribute tables associated to polygon or point feature classes are named *.PAT**
- **Coverages store the topological information (length, area, perimeter, adjacency, and connectivity) as part of the feature attribute table.**



Coverage Organization

- **Coverages must be stored in an ArcInfo workspace**
  - **Workspaces contain an INFO subdirectory**
  - **The INFO subdirectory stores the info tables**
- **Manage coverages and workspaces with ArcInfo tools only**
  - **Operating system commands do not honor the coverage to INFO link**

McKinley Workspace



Attributes of disturbance polygon

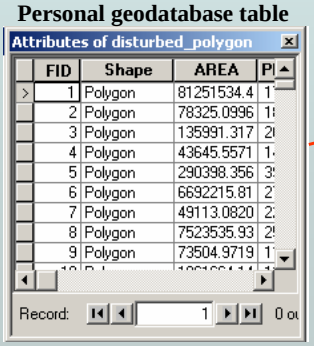
FID	Shape	AREA	PERIMETER	DISTURBANCE	DISTURBANCE-ID	REGSTAT	DISTPERIOD
2	Polygon	81251534.4	174632.51676	2	206		u
3	Polygon	78325.0996	1812.5350113	3	1	pp	99
4	Polygon	135991.317	2006.0469574	4	2	pp	91-92
5	Polygon	43645.5571	1403.4414942	5	3	pp	93
6	Polygon	290398.356	3926.1624536	6	4	pp	99
7	Polygon	6692215.81	27976.571916	7	5	pp	92-94
8	Polygon	49113.0820	2218.3484599	8	6	pp	94

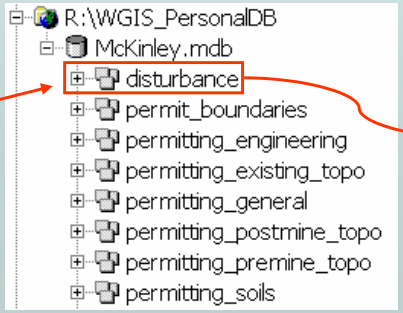
Record: 1 0 out of 200 Selected. Options

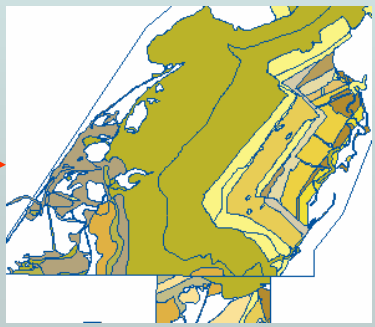
Geodatabase

- **Stores feature classes or collections (feature datasets)**
  - **Line topology only (8.2 will support all topology types)**
- **RDBMS table stores single feature class and attributes**
- **Can store feature and attribute behavior**
- **ArcInfo 8.x supports two physical implementations of the geodatabase: a *personal geodatabase* and an *ArcSDE geodatabase***
  - **Personal geodatabase is designed for smaller-scaled projects - Microsoft Access database - stored in .mdb format**
  - **ArcSDE designed for larger databases - several formats e.g. Oracle, SQL Server etc.**

Personal geodatabase table



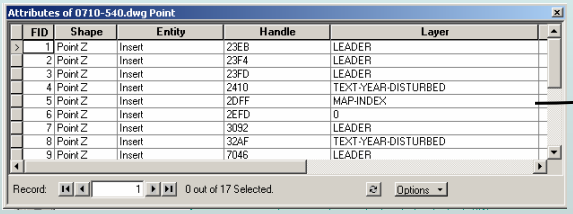




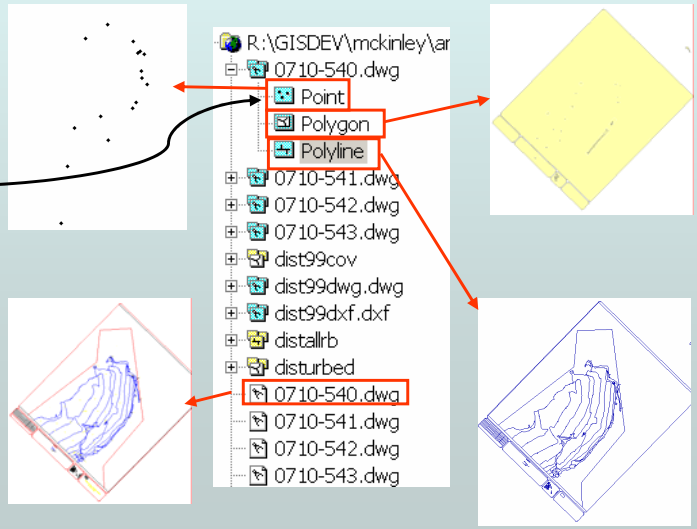
CAD

- **Computer Aided Design files (DXF, DWG, DGN)**
- **Can display and query the data as it exists in its native CAD format**
- **Can convert to a coverage or geodatabase feature class**
- **Each individual feature class within the CAD file has an associated attribute table**
- **When viewing a CAD file it will be displayed by two icons**
  - **Single white icon**
  - **Three overlapping blue icons**

CAD file (read-only attribute tables)



FID	Shape	Entity	Handle	Layer
1	Point Z	Insert	23EB	LEADER
2	Point Z	Insert	23F4	LEADER
3	Point Z	Insert	23FD	LEADER
4	Point Z	Insert	2410	TEXT-YEAR-DISTURBED
5	Point Z	Insert	2DFF	MAP-INDEX
6	Point Z	Insert	2EFD	0
7	Point Z	Insert	3092	LEADER
8	Point Z	Insert	32AF	TEXT-YEAR-DISTURBED
9	Point Z	Insert	7046	LEADER



- **Accessing CAD sources may be done individually by choosing the feature class to display or you can view the entire CAD file feature class collection all at one. If you view all feature classes at once, you cannot change the symbology nor view the associated attribute table.**

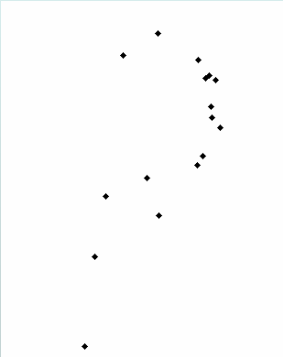
Event Tables

- **Create point or line feature classes from attribute table containing spatial information**
- **Contains either raw coordinate information or raw measures that you can import into INFO, dBASE, or Geodatabase tables**

1	X-COORD	Y-COORD	Drillhole ID
2	156784.6	1698255.3	M99-43
3	156934.2	1698622.9	M99-32
4	156233.7	1698874.1	M99-12
5	155872.7	1699023.5	M99-49

Create

New feature class

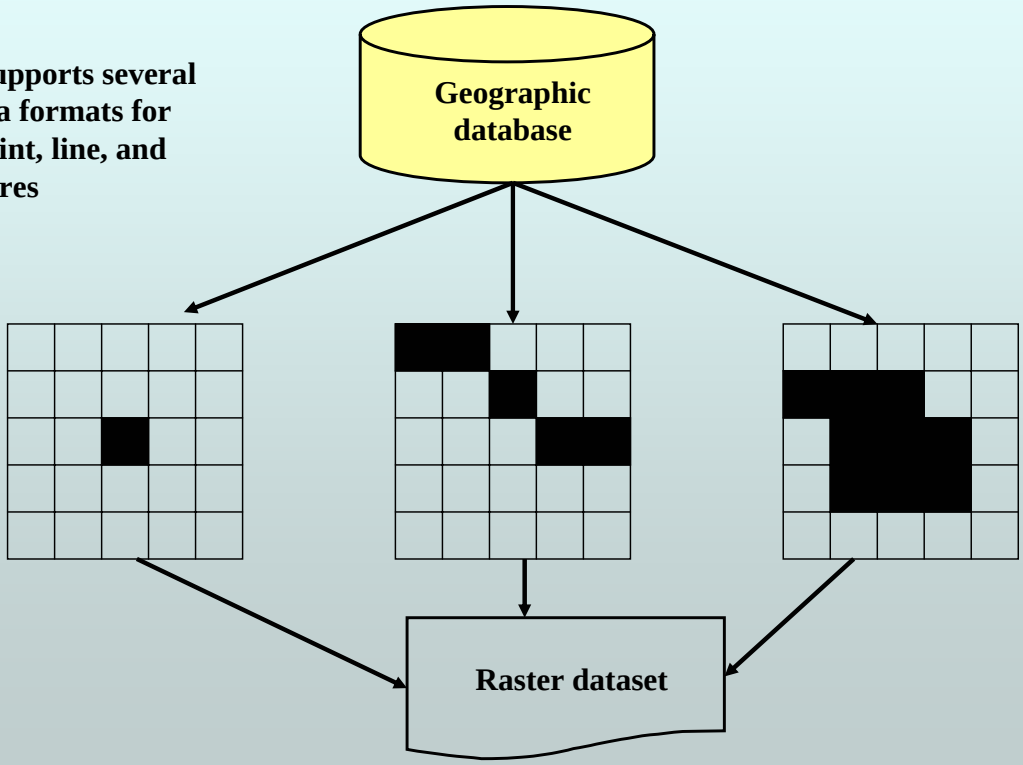


Summary of Vector Options

	Shapefile	Coverage	Geodatabase	CAD
Single feature class	✓	✓	✓	✓
Feature class collections		✓	✓	✓
Topology		✓	✓	
Allows custom features			✓	

Raster Data Formats

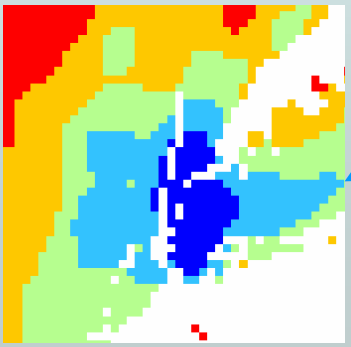
- **Image**
- **Grid**
- **ArcInfo supports several raster data formats for storing point, line, and area features**

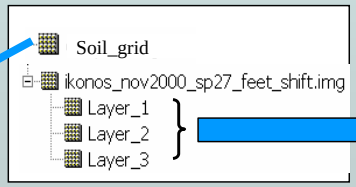


Images and Grids

- Rows and columns of equal sized cells
  - Each cell stores a value
  - Detail depends on cell size
- Images (tiff, bmp, sid, jpeg, ERDAS)
- Grids (the ArcInfo native raster format)
  - Discrete grid - data usually stored as integer - may have attribute files e.g. Land use
  - Continuous grid - data usually stored as floating point - cannot have attribute files - e.g. Elevation

Discrete grid





Image

Exercise 2A
Spatial Data Formats

- **Start ArcCatalog and ArcMap**
- **Create a new folder connection**
- **Display data in ArcCatalog and ArcMap**
- **Examine shapefiles, coverages, geodatabases, CAD, and raster formats**

Working with Tabular Data

- **Table Structure**
  - **Records/rows and fields/columns**
  - **Column types can store numbers, text or dates**
  - **User able to assign column names**
- **Some tables have a preset number of columns. For example, a polygon coverage has four standard columns: area, perimeter, coverage#, and coverage-id.**

Columns

Rows

Attribute values

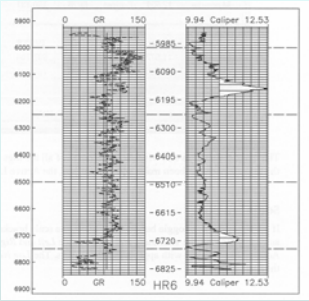
Attributes of disturbance polygon

FID	Shape	AREA	PERIMETER	DISTURBANCE	DISTURBANCE-ID	REGSTAT	DISTPERIOD
15	Polygon	43168650.4	79981.134490	15	34	pp	77-86
16	Polygon	4907117.01	25114.016756	16	14	pp	94
17	Polygon	55776044.6	101812.25060	17	15	in	77-86
18	Polygon	22040.8609	969.48055643	18	16	pp	96
19	Polygon	6952.51505	512.38890474	19	17	pp	96
20	Polygon	193005.318	5757.1051095	20	18	pp	94
21	Polygon	464810.426	8008.5183339	21	19	pp	96
22	Polygon	25417.7478	881.40569013	22	20	pp	99

Field Types for Tabular Data

•**Different field types store different kinds of values**

Drillhole - MK99-12
Date Drilled - 5/12/99
Collar Elevation - 6590.7 ft.
Depth Drilled - 235 ft.
Collar Easting - 178,246
Collar Northing - 1,684,512

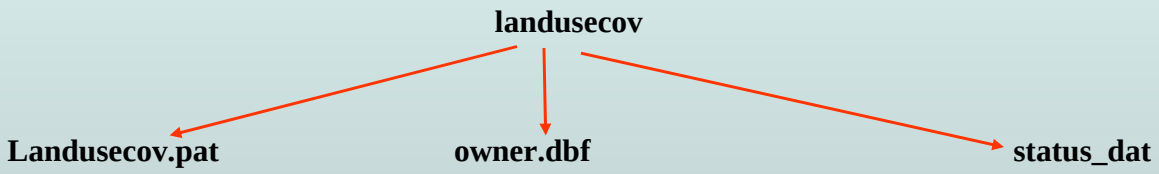


Text	Date	Float	Short	Long	Long	Blob
MK99-12	5/12/1999	6590.7	235	178246	1684512	

•**Choose the right field type for the right value**

ArcInfo Tabular Formats

- **Each ArcInfo spatial format has a native tabular format**
  - **Coverage - INFO**
  - **Shapefile - dBase**
  - **Geodatabase - RDBMS**
- **Some spatial formats can link with multiple tabular formats**



- **ArcInfo has tools for converting between formats**

Associating Tables to Feature Classes

- **Can store attributes in feature table or separate table**
- **Associate tables with common column key values**
  - **Primary keys have unique values**
  - **Foreign keys connect to primary keys and may have duplicates**
- **Must know table relationships (cardinality)**

Feature attribute table **PK** **FK** **Additional attribute table**

Attributes of seeding polygon

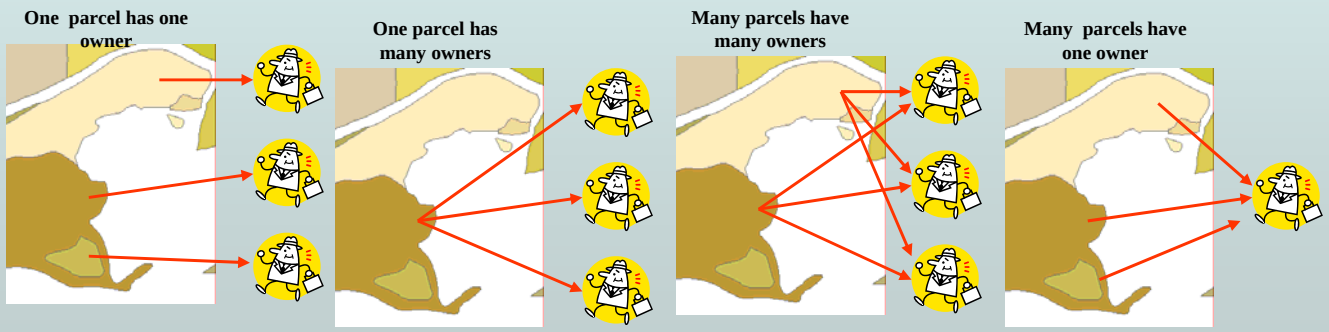
FID	Shape	AREA	PERIMETER	SEEDING#	SEEDING-ID	POLYID
2	Polygon	130357	1570.139	2	1	1
3	Polygon	174358	12194.32	3	2	8
4	Polygon	339035	4764.925	4	3	7
5	Polygon	956412	4911.167	5	4	2
6	Polygon	133040	1584.291	6	5	6
7	Polygon	174057	2976.086	7	6	414
8	Polygon	49857.	1097.885	8	7	397
9	Polygon	109226	1469.264	9	8	2
10	Polygon	91416.	1623.233	10	9	11

Attributes of seed_date

POLYID	SEED_YR	MIX_NUMBE
1	1997	97-1A
2	1998	98-1
3	1999	99-3
4	1997	97-1
5	1997	97-2
6	1993	2
7	1996	96-1
8	1995	5
9	2000	3

Table Relationships

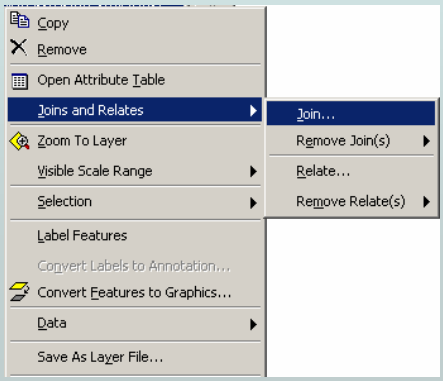
- How many "A" objects are related to "B" objects
- Four types of cardinality
  - One-One
  - One-Many
  - Many-Many
  - Many-One
- User must know cardinality before connecting tables
- Knowing which relationship type you have will prevent potential record omission errors



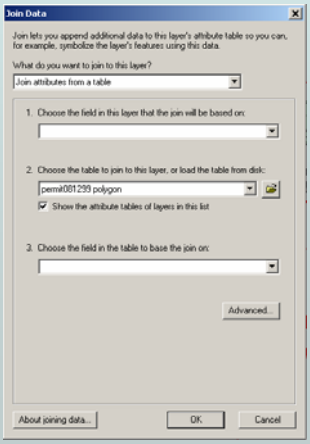
Connecting Tables with Joins

- **Physical connection between two tables**
- **Works with one-to-one or many-to-one relationships**
- **Can be created either in ArcMap or ArcToolbox**
- **Output tables can be either static or dynamic**
  - **Static - output tables do not automatically reflect edits made to input tables**
  - **Dynamic - if input table is in a geodatabase then can choose between creating a new user-named static geodatabase table or creating a dynamic virtual table that will reflect changes in the input table**

ArcMap



Define connection



Can query, symbolize or analyze the joined attributes in the new table

Relationship Classes

- **Tables are virtually linked, not physically connected**
- **Strengths:**
  - **No new feature class created**
  - **Once created the connection is active until removed**
  - **Dynamic association - can see changes to parent tables**
  - **Can edit, query symbolize the data in either table**
  - **User has more flexible cardinality associations available**

Origin table

Attributes of seeding polygon

FID	Shape	AREA	PERIMETER	SEEDING#	SEEDING-ID	POLYID
2	Polygon	130357	1570.139	2	1	1
3	Polygon	174358	12194.32	3	2	8
4	Polygon	339035	4764.925	4	3	7
5	Polygon	956412	4911.167	5	4	2
6	Polygon	133040	1584.291	6	5	6
7	Polygon	174057	2976.086	7	6	414
8	Polygon	49857.	1097.885	8	7	397
9	Polygon	109226	1469.264	9	8	10
10	Polygon	91416.	1623.233	10	9	11

Destination Table

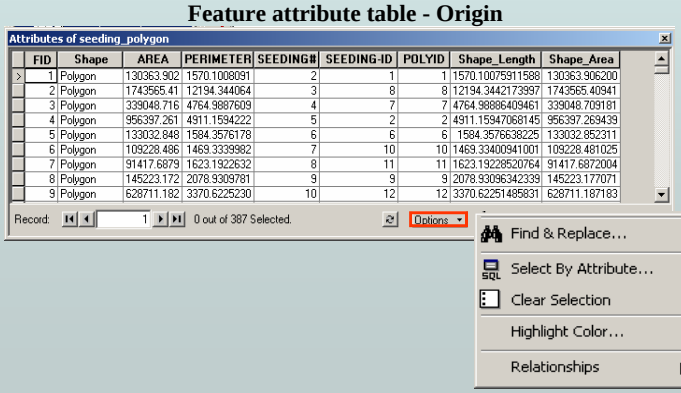
Attributes of seed_date

POLYID	SEED_YR	MIX_NUMBE
1	1997	97-1A
2	1998	98-1
3	1999	99-3
4	1997	97-1
5	1997	97-2
6	1993	2
7	1996	96-1
8	1995	5
9	2000	3

Connecting Tables with Relationship Classes

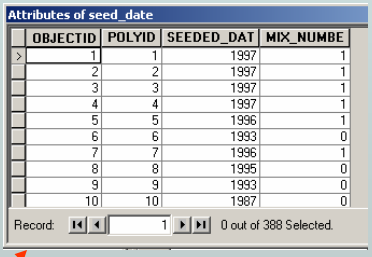
- **Can use with either Coverage or geodatabase files**
- **Define the link in ArcCatalog**
  - **Name of relationship**
  - **Type of relationship: simple or composite**
  - **Cardinality: one-one, one-many or many-many**
- **Use the relationship in ArcMap**

Feature attribute table - Origin



FID	Shape	AREA	PERIMETER	SEEDING#	SEEDING-ID	POLYID	Shape_Length	Shape_Area
1	Polygon	130363.902	1570.1008091	2	1	1	1570.10075911588	130363.906200
2	Polygon	1743565.41	12194.344064	3	8	8	12194.3442173997	1743565.40941
3	Polygon	339048.716	4764.9887609	4	7	7	4764.98886409461	339048.709181
4	Polygon	956397.261	4911.1594222	5	2	2	4911.15947068145	956397.269439
5	Polygon	133032.848	1584.3576178	6	6	6	1584.3576638225	133032.852311
6	Polygon	109228.486	1469.3339982	7	10	10	1469.33400941001	109228.481025
7	Polygon	91417.6879	1623.1922632	8	11	11	1623.19228520764	91417.6872004
8	Polygon	145223.172	2078.9309781	9	9	9	2078.93096342339	145223.177071
9	Polygon	628711.182	3370.6225230	10	12	12	3370.62251485831	628711.187183

Destination table



OBJECTID	POLYID	SEEDED_DAT	MIX_NUMBE
1	1	1997	1
2	2	1997	1
3	3	1997	1
4	4	1997	1
5	5	1996	1
6	6	1993	0
7	7	1996	1
8	8	1995	0
9	9	1993	0
10	10	1987	0

Exercise 2B
Tabular Formats

- **Start ArcCatalog and ArcMap**
- **Explore tabular formats**
- **Merge tables using joins and relationship classes**