Hour Exam I Review

GIS = Geographic Information System(s)
- Computerized management & analysis of geographic information
- Group of tools (and people) for collection, management, analysis & display of geographic information
- Computer-based tool for mapping and analyzing things that exist and events that happen
- Others

Historical Development –GIS timeline
- 1963-1977 Innovation
  - Canadian Land Inventory system, Harvard Graphics & S.A. Lab, US Census Bureau, ERTS (Landsat 1)
- 1981-1999 Commercialization
  - ArcInfo, GPS, MapInfo, TIGER, NSDI, MapQuest
- 2000- Exploitation
  - $7 billion industry, >1 million users

A GIS is Composed of Layers

Layers contain Features or Surfaces
- Features are geographic objects represented by a point, line or polygon
  - Polygons (filled or unfilled) for things large enough to have boundaries
  - Lines for things too narrow to be polygons
  - Points for things too small to be polygons

Features are linked to information
- Every Feature (e.g. road) has several Attributes (e.g. name, length) in an Attribute Table.
The Five M's

- Mapping
  - Accuracy, Reproducibility, Portability, Customization
- Measuring
  - Automation, Accuracy
- Modeling
  - Scaling, Verifiability, Analytical Tools
- Monitoring
  - Automation, Flexibility
- Management
  - Storage, Updating, Data Integrity, Security

GIS Advantages:

- Manage & organize vast amounts of geospatial data
- Rapid updating, info. dispersal
- VERIFIABLE methods
- Modeling, hypothesis-testing, PREDICTION
- Automate & customize map production

Make a Map, Graph the World

- What determines spacing of 30° increments of Lat. & Lon.?
- Dimensions and shape of earth (= DATUM)
  - Map Projection
  - Map Scale

Graph shows 30° increments of Lat. & Lon. at ~ 1:385,000,000

The Figure of the Earth

- Models
  - Sphere with radius of ~6378 km
  - Ellipsoid (or Spheroid) with equatorial radius (semimajor axis) of ~6378 km and polar radius (semiminor axis) of ~6357 km
  - Difference of ~21 km usually expressed as "flattening" (f) ratio of the ellipsoid
    - f = difference/major axis = ~1/300 for earth

Horizontal Datums

Datum = ellipse and axis of rotation.

- NAD27 (1927 North American Datum)
  - Clarke (1866) ellipsoid, non-geocentric (local) axis of rotation
- NAD83 (1983 North American Datum)
  - GRS80 ellipsoid, geocentric axis of rotation
- WGS84 (1984 World Geodetic System)
  - WGS84 ellipsoid, geocentric, nearly identical to NAD83
- Other datums in use globally

Datum Rotation Axes and the Geocenter

- Geocenter = center of mass of earth
- Local Datum vs. Geocentric Datum ("GCS" = Geographic Coordinate System = Datum)
  - Local GCS, e.g. NAD27
  - WGS84 datum
  - NAD27 datum
Coordinates have roots in marine navigation

- **Latitude**: measured by vertical angle to polaris (N. Hemisphere) or to other stars and constellations (S. Hemisphere)
- **Longitude**: determined by local time of day vs. standard time (e.g. GMT)
  - Requires accurate clocks; 1 hour difference = 15° of Longitude*

---

**Vertical Datums**

- **Sea Level (MSL), Geoid**
  - **Geoid** = surface of constant gravitational potential that best fits MSL
  - Governed by mass distribution of earth
- **Ellipsoid (HAE = Height above ellipsoid)**
  - Geometric surface
  - Datum used by most GPS receivers

---

**Sea Level (MSL), Geoid**

- Measure ht. of sea surface (via satellites) and connect with coastal surveys on land to get geoid.
- Sea “Level” (geoid) not level; as much as 700 m of relief globally.

---

**Geoid, Ellipsoid and Elevation (z)**

- Earth Surface
- Geoid (~MSL)
- Ellipsoid
- Geoid height = H.A.E.

---

**Laying the earth flat**

- Systematic rendering of Lat. (φ) & Lon. (λ) to rectangular (x, y) coordinates:

```
<table>
<thead>
<tr>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
</tr>
</tbody>
</table>
```

**“Geographic” Display**

- Distance and areas distorted by varying amounts (scale not true); e.g. high latitudes
Projected Display

- E.g. Mercator projection:
  - \( x = \lambda \)
  - \( y = \ln(\tan \phi + \sec \phi) \)

Laying the earth flat

- How?

Projection types:

- Orthographic
- Gnomonic
- Stereographic

Developable Surfaces

- Surface for projection:
  - Plane (azimuthal projections)
  - Cylinder (cylindrical projections)
  - Cone (conical projections)

Cylinder and cone produce a line of intersection (standard parallel) rather than at a point

Tangent or Secant?

- Tangent or secant if they penetrate globe

Tangent or Secant?

- Developable surfaces can be tangent at a point or line, or secant if they penetrate globe

- Secant balances distortion over wider region

- Secant cone & cylinder produce two standard parallels

Projection produces distortion of:

- Distance
- Area
- Angle
- Shape

Distortions vary with scale; minute for large-scale maps (e.g. 1:24,000), gross for small-scale maps (e.g. 1:5,000,000)

Goal: find a projection that minimizes distortion of property of interest
Where's the distortion?

- No distortion along standard parallels, secants or point of tangency.
- For tangent projections, distortion increases away from point or line of tangency.
- For secant projections, distortion increases toward and away from standard parallels.

![Tangent and Secant](image)

What needs to be specified?

- Horizontal Datum
- Origin Coordinates
- Secant Locations
- Origin X, Y Values

![Tangent to Lambert Conformal Conic](image)

Rectangular Coordinate Systems

- Universal Transverse Mercator (UTM)
  - US military developed for global cartesian reference frame.
- State Plane Coordinate System (SPCS)
  - Coordinates specific to states; used for property definitions.
- Public Land Survey System (PLS)
  - National system once used for property description
- Geographic (unprojected)

UTM Coordinate System

- T. M. secant projection is rotated about vertical axis in 6° increments to produce 60 UTM zones.
- Zone boundaries are parallel to meridians.
- Zones numbered from 180° (begins zone 1) eastward and extend from 80° S to 84° N.
- Each zone has a central meridian with a scale factor in US of 0.9996 (central meridian is farthest from secants, meaning scale distortion is greatest here).
- Secants are 1.5° on either side of the central meridian.

UTM Coordinate System

- Locations are given in meters from central meridian (Easting) and equator (Northing).
- (-) Eastings avoided by giving X value of 500,000 m ("false easting") to the Central Meridian
- In S. hemisphere, equator is given "false northing" of 10,000,000 m to avoid (-) Northings.

- N. Hemisphere origin is (500,000m, 0)
- S. Hemisphere origin is (500,000m, 10,000,000m)
UTM Coordinate System

UTM Coordinates for central Austin:

Zone 14
621,000 mE, 3,350,000 mN

Central Meridian (X = 500,000 m)

Central Meridian

99°W

Y = 3,000,000 mN

State Plane Coordinate System (SPCS)

- Developed in 1930’s to provide states a reference system that was tied to national datum (NAD27); units in feet.
- Updated to NAD83, units in meters; some maps still show SPCS NAD27 coordinates.
- Larger states divided into several zones.
- X, Y coordinates are given relative to origin outside of zone; false eastings and northings different for each zone.

Public Land Survey System (PLS)

Step 1

Step 2

Step 3

T2S, R1W, Section 33

Center Sec. 33

X, Y coordinates are given relative to origin outside of zone; false eastings and northings different for each zone.

Summary

- Projections transform geographic coordinates (ϕ, λ) to cartesian (x, y).
- Projections distort distance, area, direction and shape to greater or lesser degrees; choose projection that minimizes the distortion of the map theme.
- Points of tangency, standard parallels and secants are areas of no distortion.
- A conformal map has the same scale in all directions.

Summary (cont.)

- Modern coordinate systems are based on projections that minimize distortion within narrow, conformal zones.
- UTM is a global system using WGS84/NAD83; others are local with varying datums.

Summary (cont.)

- Projection characteristics are classified by:
  - Light source location
    - Gnomonic
    - Stereographic
    - Orthographic
  - Developable surface
    - Plane (azimuthal)
    - Cylinder (cylindrical)
    - Cone (conic)
  - Orientation
    - Normal
    - Transverse
    - Oblique
The Task

- An accurate, registered, digital map that can be queried and analyzed.

Translate:

- Paper Map
- Computer Files of Coordinates
- Entity Info.
- Queriable Database Files
- Relational or Object-Oriented Databases

Relate Spatial Coordinates to Entity Info. “Spatial DBMS” software = GIS software

Conceptual Models

Characterized all features or phenomenon as:

- Discrete objects; e.g. wells, roads, rock bodies, etc. 
  - Object-based models
- Continuous phenomena; e.g. gravity, topography, temperature, snowfall, soil pH, etc.
  - Field-based models

Logical Models

- VECTOR MODEL
  - Discrete objects are represented by points and vectors, continuous fields by irregular tessellations of triangles

- RASTER MODEL
  - Discrete objects and continuous fields are represented by an array of square cells (pixels)

Vector Models

- Graphical
- Topologic/georelational
- T.I.N.
- Network

Graphical Vector Structure

- Contains no explicit information about adjacency, containment or contiguity i.e.
  - Which polygons are adjacent?
  - Which polygons are contained within other polygons?
  - Which lines are connected? Where do lines begin and end?

= “Spaghetti Data Model”
Graphical Vector Model

- Shortcomings for maps:
  - No real world coordinates
  - No identification of individual objects; no way to attach attributes
  - Details of relationships among object (e.g. what's adjacent) not stored. Needed for spatial analysis.

Topological Vector Model

- Store pts. as x,y geographic coordinates
- Store lines as paths of connected pts.
- Store polygons as closed paths

Also explicitly store ....
- Where lines start and end (connectivity)
- Which polygons are to the right and left side of a line (adjacency)

Topological Properties

- Spatial characteristics that are unchanged by transformations like scaling, rotation and translation
  - Non-topological: x, y coordinates, area, distance, orientation
  - Topological:
    - Contiguity — what's adjacent
    - Connectivity — what's connected
    - Containment — what's inside or outside of a region

Topology — Planar Enforcement

- One and only one feature at every x, y location
  - Lines cross at nodes; polygons space-filling, exhaustive, mutually exclusive (no overlaps)
  - Sum of the area of all individual polygons equals the area of extent of all polygons
  - Common boundaries stored only once
  - A PLANAR GRAPH meets these conditions
  - Allows spatial queries for adjacency, containment and rapid what-is-where
  - All raster data is of this sort

Non-Planar vs. Planar Graphs

- Spaghetti
- Topologic

2/20/2018
Why bother with topology?

- Provides a way of error trapping and geometry validation after data entry
  - All lines must meet at a node, all polygons must close, all lines in a network must join
- Permits spatial queries, precise measurements

What kind of queries does topology permit?

- Connectivity
  - What is shortest path between features or locations? (networks, flow)
  - Find all fault trace intersections
- Contiguity
  - What’s adjacent: e.g. Show all granite/limestone contacts
  - Combine all contiguous units with a specific attribute (e.g. age) into a single unit
- Containment (= “Area Definition”)
  - What proportion of an area is underlain by a specific rock type?
  - What is spatial density of specific feature(s)?

ESRI Vector “Models”/Formats

- Topologic:
  - ArcInfo “Coverage”
  - ArcInfo “.EOO” = export format for coverage
  - ArcGIS “Geodatabase”
- Non-Topologic:
  - Shapefile

Some History...

<table>
<thead>
<tr>
<th></th>
<th>Arc/Info</th>
<th>ArcView</th>
<th>ArcGIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Versions</td>
<td>1-7</td>
<td>1-3.2</td>
<td>8.0 -8.3</td>
</tr>
<tr>
<td>Data Model</td>
<td>Coverage</td>
<td>Shapefile</td>
<td>Geodatabase</td>
</tr>
<tr>
<td>O.S.</td>
<td>Unix, PC DOS</td>
<td>PC Windows</td>
<td>PC Windows</td>
</tr>
<tr>
<td>Scripting Language</td>
<td>Arc Macro</td>
<td>Avenue Scripting</td>
<td>Vis. Basic for Appl. (VBA)</td>
</tr>
<tr>
<td>Database Software</td>
<td>Proprietary; Arc Tables</td>
<td>DBase</td>
<td>M.S. Access; ArcSDE for Oracle, etc.</td>
</tr>
</tbody>
</table>

ArcInfo Coverage

- An integrated, homogeneous set of feature classes (pts., lines, polys.) stored together
- Spatial (coordinate) data stored in binary files;
- Attributes and topologic data stored in INFO tables
- Stored within a “Workspace”

Early ESRI Data Models

- Coverages
  - Developed for workstation
  - Arc/Info ~ 1980
  - Complex structure, proprietary format
  - Attributes in INFO tables
- Shapefiles
  - Developed for ArcView ~ 1993
  - Simpler structure in public domain
  - Attributes in DBase (.dbf) tables
Folder/File Organization

Coverage

Texas

Geology

E00

Texas

Shapefile

Texas

Shapefile

Geology

Geology.shp
Geology.shx
Geology.dbf
Geology.prj

Info

... etc., etc., etc.

Sample location (points) Feature Class

Feature Class

- A collection of geographic objects with the same geometry (point, line, polygon) that share the same attributes.

ESRI Geodatabase Model

- Special data structure capable of storing objects with behaviors, not merely graphical shapes with topology and attributes.
- All spatial and attribute data for a feature are stored in a row of a single table.
- A Geodatabase is a top-level container for feature classes, coverages, shapefiles, rasters, et al. (more later) – ALL DATA CAN BE IN ONE CONTAINER AND IS THUS PORTABLE.

GEodatabase model

- Stores geographic coordinates as one of many attribute in a relational database table; no separation between attributes and location data as in earlier models.
- Uses MS Access for "Personal Geodatabase" (single user).
- Uses Oracle, DB2 or other commercial relational databases for "Enterprise GIS" (many simultaneous users).

ArcGIS Geodatabase

- Feature shapes in Geodatabase include:
  - Points, Multipoints (groups of points)
  - Lines
  - Polygons
  - Plus ....
    - Network Junctions (special Nodes)
    - Network Edges
    - For geometric networks
Object Class

- A collection of nonspatial objects stored in a table that share the same attributes (i.e. a simple table)

<table>
<thead>
<tr>
<th>Number</th>
<th>Age_Ma</th>
<th>_1_sigma</th>
<th>Rx_Type</th>
<th>Size_kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>142</td>
<td>1.5</td>
<td>B_schist</td>
<td>3.4</td>
</tr>
<tr>
<td>124</td>
<td>136</td>
<td>2.0</td>
<td>G_schist</td>
<td>1.3</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Object Class (nonspatial table)

Relationship

- A relationship is association or link between two objects in a database.
- A relationship can exist between spatial objects (features in feature classes), non-spatial objects (objects in object classes), or between spatial and non-spatial objects.

<table>
<thead>
<tr>
<th>Number</th>
<th>Age_Ma</th>
<th>_1_sigma</th>
<th>Rx_Type</th>
<th>Size_kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>142</td>
<td>1.5</td>
<td>B_schist</td>
<td>3.4</td>
</tr>
<tr>
<td>124</td>
<td>136</td>
<td>2.0</td>
<td>G_schist</td>
<td>1.3</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Relationship class

E.g. relationship between spatial and non-spatial objects

Shapefile Topology

- Shapefiles don’t store information about adjacency
- Topology is generated on the fly – vertices stored in systematic fashion to deal with containment and adjacency
- Planar enforcement can be broken by editing – not required in structure of shapefile
- But...tools available to maintain planar enforcement when digitizing in heads-up mode

Digitizing is accomplished via:

- Digitizing table or tablet
- “heads-down” digitizing
- Large table available in Rm. 526
- A mouse, on screen
- “heads-up” digitizing
- Aerial photos, other raster or vector sources as base to digitize from
- Software that converts raster to vector
- Vectorization – batch or interactive modes

Rasters are:

- Regular square tessellations
- Matrices of values distributed among equal-sized, square cells

<table>
<thead>
<tr>
<th>565</th>
<th>573</th>
<th>582</th>
<th>590</th>
</tr>
</thead>
<tbody>
<tr>
<td>575</td>
<td>580</td>
<td>595</td>
<td>600</td>
</tr>
<tr>
<td>579</td>
<td>581</td>
<td>597</td>
<td>601</td>
</tr>
<tr>
<td>580</td>
<td>600</td>
<td>620</td>
<td>632</td>
</tr>
</tbody>
</table>
Registration to “world” coordinates

Requires “world file”:
- Specify coords. of upper left corner
- Specify ground dimensions of cell, in same units

Spatial Resolution
- Defined by area or dimension of each cell
- High resolution: cells represent small areas
- Low resolution: cells represent larger areas
- Defined by size of one edge of cell (e.g., “30 m DEM”)
- File size increases with resolution

Resolution constraint
- Cell size should be less than half of the size of the smallest object to be represented (“Minimum mapping unit; MMU”)

Raster Dimension:
- Number of rows x columns
- E.g. Monitor with 1024x768 pixels

Integer Code Attributes
- Code is referenced to attribute via a “look-up table” or “value attribute table” – VAT
- Commonly many cells with the same code
- Different attributes must be stored in different raster layers

Image Raster Attributes
- Single-band: Thematic data
- Black & white: binary (1 bit) (0 = black, 1 = white)
- Grayscale (panchromatic) (8 bit): 0 (black) – 256 (white) or graduated color ramps
- Colormaps: code cells by values that match prescribed R-G-B combinations in a lookup table

VAT

<table>
<thead>
<tr>
<th>Value</th>
<th>Count</th>
<th>Rock Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>21</td>
<td>Marble</td>
</tr>
<tr>
<td>5</td>
<td>37</td>
<td>Gneiss</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>Granite</td>
</tr>
</tbody>
</table>
Image Raster Attributes

- **Multi-band Spectral Data**

  - RGB Composite
  - Band 1, Band 2, Band 3
  - Em Spectrum
  - Band = segment of Em spectrum
  - Map intensities of each band as red, green or blue.
  - Display alone or as composite

Raster File Size

<table>
<thead>
<tr>
<th>Size</th>
<th>Rows x columns x bit-depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 x 8</td>
<td>8 x 8 x 1 = 64 bytes</td>
</tr>
</tbody>
</table>

“Lossy” vs. Lossless Compression

- **Lossy**
  - Techniques that combine similar attribute information to reduce file size (e.g., JPEG, GIF, PNG, MrSID)

- **Lossless**
  - Formats: TIFF, BMP, GRID

Raster or Vector?

- **Raster**
  - Simple data structure
  - Ease of analytical operation
  - Format for scanned or sensed data – easy, cheap data entry
  - But:
    - Less compact
    - Query-based analysis difficult
    - Coarser graphics
    - More difficult to transform & project

- **Vector**
  - Compact data structure
  - Efficient topology
  - Sharper graphics
  - Object-orientation better for some modeling
  - But:
    - More complex data structure
    - Overlay operations computationally intensive
    - Not good for data with high degree of spatial variability
    - Time-intensive data entry

A DBMS provides:

- **Accuracy** - reduce errors during entry by use of established rules, templates
- **Efficiency** - rapid access & retrieval, no redundancy
- **Flexibility** - robust structure for query – e.g., What is where?
- **Security** – access and use can’t corrupt data
- **Easy updating**
Fields are defined:

- **Name** – attribute (column heading)
- **Field Type** – number (long, short), string (letters), or date
- **Length** – no. of characters
- **Precision** - decimal places

<table>
<thead>
<tr>
<th>Type</th>
<th>Label</th>
<th>Type</th>
<th>Length</th>
<th>Precision</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fields types in ArcGIS

- **Short Integer** – 1 to 4 digits (no decimal)
- **Long Integer** – 5 to 9 digits (no decimal)
- **Float** = 1 – 13 digits, decimal (short real)
- **Double** = 14 – 19 digits, decimal (long real)
- **Text** = 1 – 255 characters
- **Date** = 8 character
- **Blob** = binary large object

GIS attribute data models

- **Hierarchical** – pre-1980
- **Relational** – 1980’s, 1990’s; still dominant today
- **Object-oriented** – late ‘90’s; newest, implemented by some GISs – still undergoing R&D

Hierarchical - Limitations

1. Linear structure can’t deal with multiple “memberships”
   - E.g. a single well might be stored under different databases for taxes, production, drilling history, water quality, etc
   - INEFFICIENT

2. Can’t assemble all this data for query in a hierarchical database

Hierarchical - Limitations

2. Can’t deal with exceptions to linear scheme – entities may not belong to next higher class but could instead contain it.
   - E.g. Structure oil well database by:
     
     state, county, oil field, well
     
     What of field that spans several counties with wells that produces from more than one field (pay zone)?

   - i.e. No one-to-many relationships

Relational Database advantages

- Data stored in separate files
  - Easy update, editing, searching without affecting or using all data
- Flexibility
  - Using key(s), can extract and assemble records and attributes to form new tables
  - Subsets of database can be queried by standard means - SQL
### Parts of a Relation

<table>
<thead>
<tr>
<th>Record, tuple</th>
<th>Attribute, field</th>
<th>Primary Key</th>
</tr>
</thead>
</table>

#### Production (barrels/day, cfs)

<table>
<thead>
<tr>
<th>ID</th>
<th>Oil</th>
<th>Gas</th>
<th>Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>53</td>
<td>1200</td>
<td>5</td>
</tr>
<tr>
<td>43</td>
<td>108</td>
<td>2500</td>
<td>15</td>
</tr>
</tbody>
</table>

### Properties of Relations

- Each row has to be unique; no row-to-row dependency
- Row order irrelevant
- Column order irrelevant
- All attribute values must be stored in separate rows (“first normal form”)

### One-to-One Table Join

**Drilling Record**

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Spudded</th>
<th>Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>Exxon #1</td>
<td>2/4/96</td>
<td>6/3/96</td>
</tr>
<tr>
<td>43</td>
<td>Shell #5</td>
<td>3/14/97</td>
<td>6/12/96</td>
</tr>
</tbody>
</table>

**Production (barrels/day, cfs)**

<table>
<thead>
<tr>
<th>ID</th>
<th>Oil</th>
<th>Gas</th>
<th>Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>53</td>
<td>1200</td>
<td>5</td>
</tr>
<tr>
<td>43</td>
<td>108</td>
<td>2500</td>
<td>15</td>
</tr>
</tbody>
</table>

- One record from source table (production) is joined to one record of destination (drilling record) table to create a “View” – virtual combination

### Result of One-to-One Table “Join”

**Joined Production table and Drilling Record:**

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Spudded</th>
<th>Completed</th>
<th>Oil</th>
<th>Gas</th>
<th>Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>Exxon #1</td>
<td>2/4/96</td>
<td>6/3/96</td>
<td>53</td>
<td>1200</td>
<td>5</td>
</tr>
<tr>
<td>43</td>
<td>Shell #5</td>
<td>3/14/97</td>
<td>6/12/96</td>
<td>108</td>
<td>2500</td>
<td>15</td>
</tr>
</tbody>
</table>

- View can’t be edited – destination table can be

### One-to-Many Join

**Drilling Record**

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Spudded</th>
<th>Completed</th>
<th>Field_ID</th>
<th>Name</th>
<th>Discovered</th>
<th>Total_Oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>Exxon #1</td>
<td>2/4/96</td>
<td>6/3/96</td>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>43</td>
<td>Shell #5</td>
<td>3/14/95</td>
<td>6/12/96</td>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>72</td>
<td>Amoco #3</td>
<td>4/8/88</td>
<td>4/8/89</td>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>55</td>
<td>BP #2</td>
<td>6/8/90</td>
<td>8/8/91</td>
<td>2</td>
<td>Wildcat</td>
<td>2/2/82</td>
<td>85640</td>
</tr>
</tbody>
</table>

**Oil/Gas Fields**

<table>
<thead>
<tr>
<th>Field_ID</th>
<th>Name</th>
<th>Discovered</th>
<th>Total_Oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Longview</td>
<td>1/20/56</td>
<td>13000564</td>
</tr>
<tr>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>3</td>
<td>Anahuac</td>
<td>4/11/73</td>
<td>3587889</td>
</tr>
</tbody>
</table>

- One record from source table (production) is joined to many records of destination table

### One-to-Many Join Result

**Drilling Record**

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Spudded</th>
<th>Completed</th>
<th>Field_ID</th>
<th>Name</th>
<th>Discovered</th>
<th>Total_Oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>Exxon #1</td>
<td>2/4/96</td>
<td>6/3/96</td>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>43</td>
<td>Shell #5</td>
<td>3/14/95</td>
<td>6/12/96</td>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>72</td>
<td>Amoco #3</td>
<td>4/8/88</td>
<td>4/8/89</td>
<td>2</td>
<td>Katy</td>
<td>2/3/48</td>
<td>85640</td>
</tr>
<tr>
<td>55</td>
<td>BP #2</td>
<td>6/8/90</td>
<td>8/8/91</td>
<td>2</td>
<td>Wildcat</td>
<td>2/2/82</td>
<td>85640</td>
</tr>
</tbody>
</table>

- View can’t be edited – destination table can be
GIS’ are Spatial Databases

- Coverage and Shapefile models
  - Spatial information stored in spatial attribute files, attributes in relational database table
  - Feature ID is key
  - Spatial information can’t participate in relational database advantages

- Geodatabase model
  - All information, spatial and aspatial, are stored together in a relational database

Nature of the problem:

- Data source registration may differ by:
  - Rotation
  - Translation
  - Distortion

General problem is then:

Source \((x, y)\) \(\rightarrow\) Destination \((X', Y'; UTM)\)

**Displacement Link**

Control Points

\((0,0)\)

\((0,1)\)

\((1,0)\)

\((1,1)\)

\((501000, 3725000)\)

\((498100, 3715000)\)

**“Warp”**

Geometric Transformations

- Affine transformation constants:
  \[ X'_1 = Ax_1 + By_1 + C \]
  \[ Y'_1 = Dx_1 + Ey_1 + F \]

A, E = scale factors
B, D = rotation terms
C, F = translation terms

- With six unknowns, need *minimum of three points* (yielding 6 equations).

Geometric Transformation of Raster Data

- **Solution**: “Resampling” – Create and fill a new matrix of empty destination cells with values from source raster. Tag remaining cells as “no data”.

Affine Transformations

- “Goodness of Fit” given by RMS error:

\[
\text{RMS error} = \sqrt{\frac{e_1^2 + e_2^2 + e_3^2 + e_4^2}{4}}
\]
Implications of Resampling

- Cell size and number of rows and columns may change on projection and/or georeferencing.
- Minimize problems by georeferencing to a desired projection, not to unprojected vector data.
- Raster datasets must be in same projection and coordinate system for display and analysis.

“Georeferencing” vs. “Spatial Adjustment”

- Georeferencing
  - Best fit of all source control points to all destination control points – transformation ("Warping") of data for overall best fit.
  - Alignment of data to map coordinates
  - R.M.S. error given

- “Spatial Adjustment”
  - More versatile; can “Warp”, also “Rubbersheet” and “Edgematch”
  - Adjustment by latter two is piece-wise fitting; point by point matching but no overall warping.