Geol 552 - Lesson 14

Midterm preparation + mini project 2 introduction

• copy follow_along\midterm_practice.mxd into your U:\ArcGIS folder

Midterm: Tuesday Oct. 18, 9:00-11:00

• Part 1 (15 min):
  – Multiple choice or True-false
  – 14 questions (13 or more will get you an A)
  – open book, open notes

• Part 2 (60 min):
  – Several exercises on the same data set (like in the book)
  – 3 parts (each 14 pts.)
  – Maximum points: 28 pts
  – Data set: shapefiles only (you may save
  – Each part is self-contained and will not require results from any other part
  – Each part will involve several steps (3-5)
  – Open book, notes allowed

• Test this now: connect to \delphi\exam_drop_folder and make a new folder

GIS data - coordinates, attributes, symbolization

• Data files: shapefiles, feature classes (geoDB)
  – layer: features of the same geometry type
  – features: coordinates + attribute(s)
  – coordinates => can be draw on a map
  – does NOT store any symbolization!

• Stand alone tables: only attributes (records)
  – does not store and coordinates* or symbolization!
  – can be joined to features, if a key field exists

• mxd (ArcMap document), lyr (“layer file”)
  – store appearance (symbolization) only
  – no data, only references (links) to data
  – invalid links => red exclamation marks

layer of polygon features

Row = Record
Column = Attribute = Field

<table>
<thead>
<tr>
<th>FID</th>
<th>Shape*</th>
<th>STATE</th>
<th>POP2000</th>
<th>HISP00</th>
<th>POP00</th>
<th>RACE00</th>
<th>HISP00</th>
<th>RACE00</th>
<th>TOTAL</th>
<th>GROWTH</th>
<th>HOUSEHOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Polygon</td>
<td>HI</td>
<td>1182168</td>
<td>1149468</td>
<td>124</td>
<td>35607</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Polygon</td>
<td>WA</td>
<td>486032</td>
<td>583089</td>
<td>22</td>
<td>102431</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Polygon</td>
<td>VT</td>
<td>85295</td>
<td>85295</td>
<td>5</td>
<td>304163</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Polygon</td>
<td>ME</td>
<td>1227338</td>
<td>1227338</td>
<td>38</td>
<td>455312</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Polygon</td>
<td>ND</td>
<td>600068</td>
<td>600068</td>
<td>5</td>
<td>144719</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Polygon</td>
<td>SD</td>
<td>600054</td>
<td>726900</td>
<td>9</td>
<td>259024</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Polygon</td>
<td>WI</td>
<td>455968</td>
<td>455968</td>
<td>5</td>
<td>148039</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Polygon</td>
<td>ID</td>
<td>1065768</td>
<td>1065768</td>
<td>12</td>
<td>360232</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data base joins

- can the tables be joined (key? rule of joining)?
- Which is destination (left) and source (right)?
- how many features do you expect to be in the joined table?
- the input layer with this number of features has to be the destination for the join!

Destination Table

Source Table

**One-to-One Relationship**

**Many-to-One Relationship**

**Destination Table**

**Source Table**

**Many-to-Many Relationship**

**Destination Table**

**Source Table**

In **Data base joins**

- **Destination** vs. **Source**
- **Same type and number of features as destination**
- Destination (left) vs. Source (right)
- **Key field**

USA county data

(per county: NAME, POP and STATE)

- **Find all counties with > 7000 people?**
  - SQL query ("POP" > 20000)
- **Find total number of people in each state?**
  - **summarize**: Group field? Stats field? Type of stat?
  - Group field: STATE ; stats field: POP ; type of stat: SUM
  - What would COUNT_ contain here?
- **Total number of people in US?**
  - statistics for POP

Field statistics:
- all types of stats
- just for one field
- E.g. ALL records (avg. city population?)

Summarize:
- Use one field to make groups (R-click on it!)
- Does some form of stats from all records belonging to each group
- E.g. Per county name (group) average POP
Selections (select-by, queries)

- By Attribute (SQL) - based on table only
- By Location - based on spatial concept only
- Save selection - Right click on layer in TOC - “Selection” - “Create layer from Selected features”
- **Definition query** (not a “light blue" selection, think filter, only the "selection" is shown, rest is hidden)

Spatial Joining Destination/Source

<table>
<thead>
<tr>
<th>City</th>
<th>Dest Poly</th>
<th>Source Poly</th>
</tr>
</thead>
<tbody>
<tr>
<td>City1</td>
<td>10</td>
<td>Poly1</td>
</tr>
<tr>
<td>City2</td>
<td>20</td>
<td>Poly2</td>
</tr>
<tr>
<td>City3</td>
<td>15</td>
<td>Poly2</td>
</tr>
</tbody>
</table>

Q: In which poly is each city? Dest? Simple or Summarized?
How many people in each poly?
How many cities in each poly?

For Dest: cities  Source Polygons

For Dest: polygon  Source: cities

<table>
<thead>
<tr>
<th>City</th>
<th>Inside of</th>
<th>Sum</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>City1</td>
<td>Poly1</td>
<td>10</td>
<td>B</td>
</tr>
<tr>
<td>City2</td>
<td>Poly2</td>
<td>20</td>
<td>B</td>
</tr>
<tr>
<td>City3</td>
<td>Poly2</td>
<td>15</td>
<td>A</td>
</tr>
</tbody>
</table>

Which county has the most cities (and how many)?
Which is the right strategy?
(look at the table on p. 158)

- D: Cities, S: Counties
  - Summarized Inside Join
  - Simple Distance Join
- D: Counties, S: Cities
  - Summarized Inside Join
  - Simple Distance Join

Looking at each hotel, which star is the closest to any hotel?
Which is the right strategy?

- D: Hotels, S: Stars
  - Simple Inside Join
  - Simple Distance Join
- D: Stars, S: Hotels
  - Summarized Inside Join
  - Simple Distance Join
Color each dot according to the closest red triangle (each dot has to know: which is my closest Triangle!)

(I added the partition lines manually...)

Student requests

- Calculating percentages (HW5)
- *2000 features trick (Attribute Table)
- Join Dating table to counties
  - Why FIPS as key?
  - Destination? Number of feature pre and post join?
  - SQL for lost POP from 1990 to 2000?

Miniproject 2 (Iowa data)

- Assemble data (3 - 5 layers) + base map for a theme
- Think of chains of operations from operations: Join, spatial join, summarize, attribute or spatial query, statistics
- Use at least 2 chains
- Possible themes:
  - Tourists/Travellers
  - Geology & environmental issues
  - Transportation
  - Water
  - Population (census)
  - Marketing study (who/where to sell a product)
- Sample Iowa Data: data\Miniproj 2 data
- More: \pub\pub\IowaDNR\IA_state (see categories)