FNRM 3131
Introduction to GIS in Natural Resource Management

Instructors: Andrew C. Jenks
215 Green Hall
ajenks@umn.edu, @acjenks

Paul V. Bolstad
301H Green Hall
pbolstad@umn.edu
Course Topics

- Motivation: why GIS?
- Data Models
- Map Projections and Coordinate Systems
- Data entry - digitizing
- Data entry - GPS
- Data entry - digital data
- Image data
- Tables
- Basic Spatial Analysis
- Raster Analyses
- Spatial Models and Modeling
- Interpolation
- Data Quality
- The Future
What is a GIS?


Two important components - What, and Where
We usually understand GIS to be computer-facilitated system

GIS is NOT only software/hardware

Also includes:
• Trained personnel
• Supporting Institution
• Protocols for use
Why GIS?
Motivation

MISTAKES
It Could Be That the Purpose of Your Life Is Only to Serve as a Warning to Others.
Why GIS?
Why GIS?
Pavement quality inventory and monitoring
Why GIS?

Sustaining Lobster Fisheries in Penobscot Bay

Why is it important?
Fisheries stocks dwindling, lobster an economic mainstay with sustainable harvests
Lobster habitat (green) in outer Penobscot Bay
From Observations and Process Knowledge to Prediction and Action

Distribution of lobster larvae: Stages I and IV
Eastern Gulf of Maine, August 1999

Eric Annis,
University of Maine

Stage I  Stage IV (postlarvae)

Sampling for lobster larvae

http://www.islandinstitute.org/penbaycoll_understandinglife.php
Dynamic Spatial Models

York River

Courtesty Virginia Institute of Marine Science
GIS Software Tools

GIS started at universities as research tools – Harvard, Yale, Minnesota, Clark University
GIS software have evolved to robust (sort of) tools capable of a wide variety of tasks

ESRI (ArcGIS)  Quantum GIS (QGIS)
Microimages  Autocad  MapInfo
ERDAS  Idrisi  Manifold
GRASS  Intergraph
Caveat

Tools for data management

Tools for generic analysis

Science and specific tools often lacking
What is in it for me?

Help us understand your goals and expectations

Please answer the questions on the Moodle site:

Direct link is:  [Questionnaire](https://ay14.moodle.umn.edu/mod/questionnaire/view.php?id=38942)

Or type/paste it in:

Steps for Successful GIS Analysis
(Goals defined, methods exist)

• Choose best data model
• Define bounds, geographic region of interest
• Identify existing spatial data
• Determine coordinate system for analyses
• Develop digital database
• Document database (origin, quality)
• Perform analyses
• Report results
• Update Database
Famous Environmental Thinkers/Writers

• Aldo Leopold
  – *Sand County Almanac* (1949)

• Rachel Carson
  – *Silent Spring* (1962)
Famous Environmental Thinkers/Writers

- **Aldo Leopold**
  - *Sand County Almanac* (1949)
- **Rachel Carson**
  - *Silent Spring* (1962)
- **Ian McHarg**
Ian McHarg
Design With Nature (1969)
Some terms Ian McHarg created:

- Environmental impact assessment
- New community development
- Coastal zone management
- Brownfields restoration
- River corridor planning
- Sustainability & Regenerative design
Overlay Method for Site Analysis

AGRICULTURAL VALUE

AQUIFER VALUE

ECOLOGICAL VALUE

BEST ROUTE
GIS Fundamentals
Real World

Data Model

Data Structure

1.2, 4.7
5.8, 3.6
8.9, 7.2
.
.

phenomena that exist

An abstraction, relevant phenomena and properties

computer representation
Data Model – Define spatial objects in a database,
And represent relationships among objects (connection, adjacency, proximity, influence)
We approximate entities with objects.

This approximation is biased.

**Terms**

Entities - "things" in the real world we represent (*Rivers*, *buildings*, *soil types*, *wetlands*)

Objects - our representation in a data model.
Figure 2-3: Coordinate and attribute data are used to represent entities.
Many objects for a theme, with properties recorded for each object.
Most common data models define *thematic* layers.

Typically, layers, one layer for each distinct view of a theme.
Cartesian Coordinates

2 - Dimensional

3 - Dimensional

origin

0 10 20 30 40 50

X

0 10 20 30 40 50

Y

0 10 20 30 40 50

Y

0 10 20 30 40 50

X

z

y

x
Spherical coordinates

- North pole
- Equator
- Origin (0°, 0°)
- Radius (r)
- Latitude (\(\phi\))
- Longitude (\(\lambda\))
Degree Units

Each 1 minute is split into 60 seconds

Each 1 degree split into 60 minutes

Coordinates are expressed as degrees, minutes, seconds

e.g. 15° 30' 59"

or as decimal degrees, e.g., 15.51539°
Sexagesimal (base sixty)
Sexagesimal (base sixty)

Originated with ancient Sumerians and Babylonians (2000 BC)

60 has 12 factors (1, 2, 3, 4, 5, 6, 10, 12, 15, 30, 60)

Allows many simple fractions also worked well with cuneiform script
We Can Convert!

**DD from DMS**

\[
DD = D + \frac{M}{60} + \frac{S}{3600}
\]

**e.g.**

DMS = 32° 45' 28"

\[
DD = 32 + \frac{45}{60} + \frac{28}{3600} \\
= 32 + 0.75 + 0.0077778 \\
= 32.7577778
\]

**DMS from DD**

D = integer part
M = integer of decimal part \times 60
S = 2nd decimal \times 60

e.g.

\[
DD = 24.93547 \\
D = 24 \\
M = \text{integer of } 0.93547 \times 60 \\
= \text{integer of } 56.1282 \\
= 56 \\
S = 2nd \text{ decimal } \times 60 \\
= 0.1282 \times 60 = 7.692 \\
\text{so DMS is} \\
24° 56' 7.692"
\]
Common Data Models

Vector

- Points
- Line
- Area

(\(x, y\))

Raster

- Points
- Line
- Area
Vectors Define Discrete Features

- Points
- Nodes
- Vertices
- Lines
- Polygons
Points

Lines

Areas
Three Types of Vector Features

Points

Lines

Polygons

Points

<table>
<thead>
<tr>
<th>Point ID</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32.7</td>
<td>45.6</td>
</tr>
<tr>
<td>2</td>
<td>76.3</td>
<td>19.5</td>
</tr>
<tr>
<td>3</td>
<td>22.7</td>
<td>15.8</td>
</tr>
</tbody>
</table>

etc.....

Lines

<table>
<thead>
<tr>
<th>Line ID</th>
<th>Begin Point</th>
<th>End Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>239</td>
<td>1</td>
</tr>
</tbody>
</table>

etc.....

Polygons

<table>
<thead>
<tr>
<th>Polygon ID</th>
<th>Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>11, 12, 52, 53, 54</td>
</tr>
<tr>
<td>B</td>
<td>52, 53, 9, 41, 22, 13</td>
</tr>
</tbody>
</table>
We can get multiple objects from a single entity, e.g.,

lake may be a
• municipal water source
• recreation area
• flood control sink
• wildlife habitat

We typically store different object types (even from the same entity) in different layers.
Representations and Data Structures

• We typically have a one-to-one correspondence between entities (the real world thing) and objects (the computer representation) in a layer – but not always
One-to-One

<table>
<thead>
<tr>
<th>ID</th>
<th>type</th>
<th>area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>16.8</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>22.2</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>18.4</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>20.7</td>
</tr>
<tr>
<td>ID</td>
<td>State Name</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>hawaii</td>
<td></td>
</tr>
</tbody>
</table>
A Many to One Correspondence

Two lakes, one polygon
We typically represent entities as one of four types of objects –

- points
- lines
- areas
- 3-D objects
Representation – Enforced Uniformity

polygon 1: Blissfull Acres
land use: residential
area: 192 ha

polygon 2: Turtle Lake
land use: park
area: 102 ha

polygon 3: Spendmore Mall
land use: commercial
area: 174 ha

coordinates are used to define feature boundaries

attributes define feature properties
Vector Assumptions

Dimensionality – 0, 1^{st}, 2^{nd} points have no dimension, lines have no width
Enforced Uniformity
## Assignment

### Ambiguity

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
</tr>
</tbody>
</table>
Vector Topology

Topology – geometric properties that to not change with shape: Adjacency, connectivity, containment

a) spaghetti

b) topological

c) topological - warped
Why Topology Matters

We frequently deform our data (map projections)

We can better achieve accuracy in analysis

Some calculations are faster
Planar Topology – no overlaps
housing data layer

property line data layer

disallowed overlap
On a blank sheet of paper, draw a topologically correct (not necessarily geometrically correct) rendition of the 7 county metro area (name counties, if you can).

Identify all the nodes.

Create a table corresponding to the geographic data you drew (make the linkages clear).

Include an attribute for county area, in square kilometers, and square miles.
<table>
<thead>
<tr>
<th>NAME</th>
<th>FIPS</th>
<th>sq.km.</th>
<th>sq.miles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anoka</td>
<td>27003</td>
<td>1154</td>
<td>446</td>
</tr>
<tr>
<td>Washington</td>
<td>27163</td>
<td>1096</td>
<td>423</td>
</tr>
<tr>
<td>Hennepin</td>
<td>27053</td>
<td>1570</td>
<td>606</td>
</tr>
<tr>
<td>Ramsey</td>
<td>27123</td>
<td>440</td>
<td>170</td>
</tr>
<tr>
<td>Carver</td>
<td>27019</td>
<td>973</td>
<td>376</td>
</tr>
<tr>
<td>Dakota</td>
<td>27037</td>
<td>1518</td>
<td>586</td>
</tr>
<tr>
<td>Scott</td>
<td>27139</td>
<td>953</td>
<td>368</td>
</tr>
</tbody>
</table>
Can you locate Iran? How about Afghanistan?
Rasters – Fixed Cell Size, Grid Orientation

- **Coordinates of lower-left cell**: (x, y)
- **Cell dimension**: distance between two adjacent cells along the x-axis or y-axis.
Rasters – Discrete or Continuous Features

<table>
<thead>
<tr>
<th>discrete</th>
<th>continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>a a a a a r f f f a a a a a a</td>
<td>645 650 654 658 653 648</td>
</tr>
<tr>
<td>a a a a a r f f f a a a a a a</td>
<td>664 666 670 672 668 659</td>
</tr>
<tr>
<td>a a a a a r f f f a a a a a a</td>
<td>678 682 684 693 689 680</td>
</tr>
<tr>
<td>a a a a a a r f f f a a a a a a</td>
<td>703 708 714 721 719 716</td>
</tr>
<tr>
<td>a a a a a a a a a a a a a a a a</td>
<td>728 732 738 744 745 732</td>
</tr>
<tr>
<td>a a a a a a a a a a a a a a a a</td>
<td>730 739 744 749 748 735</td>
</tr>
</tbody>
</table>

a = agriculture  u = developed
f = forest       r = river
h = highways
a) Vector, one-to-one

attribute table

<table>
<thead>
<tr>
<th>IDorg</th>
<th>class</th>
<th>area</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>16.8</td>
</tr>
<tr>
<td>B</td>
<td>11</td>
<td>22.2</td>
</tr>
<tr>
<td>C</td>
<td>15</td>
<td>18.4</td>
</tr>
<tr>
<td>D</td>
<td>21</td>
<td>16.4</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>3.8</td>
</tr>
</tbody>
</table>
b) **Raster, one-to-one**

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>B</th>
<th>B</th>
<th>B</th>
<th>B</th>
<th>B</th>
<th>B</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>D</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
</tbody>
</table>

### attribute table
*(cell 1 is upper-left corner)*

<table>
<thead>
<tr>
<th>cell-ID</th>
<th>IDorg</th>
<th>class</th>
<th>area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>10</td>
<td>0.8</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>10</td>
<td>0.8</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>10</td>
<td>0.8</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>10</td>
<td>0.8</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>11</td>
<td>0.8</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>11</td>
<td>0.8</td>
</tr>
<tr>
<td>7</td>
<td>B</td>
<td>11</td>
<td>0.8</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>100</td>
<td>E</td>
<td>10</td>
<td>0.8</td>
</tr>
</tbody>
</table>
### c) Raster, many-to-one

<table>
<thead>
<tr>
<th>10</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>11</th>
<th>11</th>
<th>11</th>
<th>11</th>
<th>11</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>21</td>
<td>21</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

### Attribute Table

<table>
<thead>
<tr>
<th>class</th>
<th>area</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>18.4</td>
</tr>
<tr>
<td>11</td>
<td>24.0</td>
</tr>
<tr>
<td>15</td>
<td>21.6</td>
</tr>
<tr>
<td>21</td>
<td>13.6</td>
</tr>
</tbody>
</table>
Rasters as Sampling Points – continuous variables
Raster – The Mixed Pixel Problem

Landcover map – Two classes, land or water

Cell A is straightforward

What category to assign For B, C, or D?
Raster – The Storage Space/Resolution Tradeoff

Decreasing the Cell Size by one-half causes a Four-fold increase in the storage space required
<table>
<thead>
<tr>
<th>2</th>
<th>2</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
<td>null</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>a</th>
<th>a</th>
<th>a</th>
<th>a</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>
Resampling

Values for new layer?

For categorical data: most common, or arbitrary if tied

For numeric data:
- center cell, or
- most common cell, or
- some weighted average

\[ \text{new} = \frac{a+b+a+d}{4} \]
What if the cells aren’t “well behaved”
Resampling - Distance-weighted averaging

bilinear interpolation

What is the value of $Z_{out}$?

$$Z_b = Z_4 + \frac{(Z_3 - Z_4)*d_1}{c}$$

$Z_b = 1.4 + \frac{(4.6 - 1.4)*2.9}{5} = 3.26$

$$Z_u = Z_2 + \frac{(Z_1 - Z_2)*d_1}{c}$$

$Z_u = 4 + \frac{(6 - 4)*2.9}{5} = 5.16$

$$Z_{out} = Z_b + \frac{(Z_u - Z_b)*d_2}{c}$$

$Z_{out} = 3.26 + \frac{(5.16 - 3.26)*2.2}{5} = 4.1$
Changing Resolution (resampling) with Categorical Data

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>C</th>
<th>B</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>D</td>
<td>C</td>
<td>E</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>B</td>
<td>F</td>
<td>H</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>F</td>
<td>C</td>
<td>E</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>F</td>
<td>G</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A or C</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>G</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

| ? | ? | ? |
## Comparisons, raster v.s. vector

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Vector</th>
<th>Raster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positional Precision</td>
<td>Can be precise</td>
<td>Defined by cell size</td>
</tr>
<tr>
<td>Attribute Precision</td>
<td>Poor for continuous data</td>
<td>Good for continuous data</td>
</tr>
<tr>
<td>Analytical Capabilities</td>
<td>Good for spatial query, adjacency, area, shape analyses. Poor for continuous data. Most analyses limited to intersections. Slower overlays.</td>
<td>Spatial query more difficult, good for local neighborhoods, continuous variable modeling. Rapid overlays.</td>
</tr>
<tr>
<td>Data Structures</td>
<td>Often complex</td>
<td>Often quite simple</td>
</tr>
<tr>
<td>Storage Requirements</td>
<td>Relatively small</td>
<td>Often quite large</td>
</tr>
<tr>
<td>Coordinate conversion</td>
<td>Usually well-supported</td>
<td>Often difficult, slow</td>
</tr>
<tr>
<td>Network Analyses</td>
<td>Easily handled</td>
<td>Often difficult</td>
</tr>
<tr>
<td>Output Quality</td>
<td>Very good, map like</td>
<td>Fair to poor - aliasing</td>
</tr>
</tbody>
</table>
No Decision is Final – We Can Convert
Triangular Irregular Network (TIN)

Typically used to represent terrain or other spot-sampled continuous variables
Connect sample points in a network of triangles
Why? – to preserve sample accuracy, save space
TIN Parts

Points – sample locations

Edges – connecting lines

Facets – triangles, “faces”
TIN – Triangle Formation

TIN triangles defined such that

• Three points on a circle
• Circles are empty – they don’t contain another point

These are convergent circles
Multiple Representations
Data and File Structures

Real World Data Model

<table>
<thead>
<tr>
<th>Shape</th>
<th>Area</th>
<th>Sl. Weigh</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygon</td>
<td>280290.98165</td>
<td>580</td>
<td>PEM/SS1Bddg</td>
</tr>
<tr>
<td>Polygon</td>
<td>1630.52135</td>
<td>581</td>
<td>PUBGx</td>
</tr>
<tr>
<td>Polygon</td>
<td>16229.73690</td>
<td>582</td>
<td>U</td>
</tr>
<tr>
<td>Polygon</td>
<td>792.66535</td>
<td>583</td>
<td>PEMC</td>
</tr>
<tr>
<td>Polygon</td>
<td>1218.39765</td>
<td>584</td>
<td>PSS1C</td>
</tr>
<tr>
<td>Polygon</td>
<td>1294.57970</td>
<td>585</td>
<td>U</td>
</tr>
<tr>
<td>Polygon</td>
<td>1009.27755</td>
<td>586</td>
<td>PUBGx</td>
</tr>
<tr>
<td>Polygon</td>
<td>5928.50695</td>
<td>587</td>
<td>PUBGx</td>
</tr>
<tr>
<td>Polygon</td>
<td>1028.50440</td>
<td>588</td>
<td>PEMC</td>
</tr>
<tr>
<td>Polygon</td>
<td>9969.29615</td>
<td>589</td>
<td>PUBGx</td>
</tr>
<tr>
<td>Polygon</td>
<td>6894.33895</td>
<td>590</td>
<td>PUBGx</td>
</tr>
</tbody>
</table>

1.2, 4.7
5.8, 3.6
8.9, 7.2

.
Data and File Structures

Data are stored as binary numbers.

Bits are 0 or 1.

Bytes are 8 bits.

Data (e.g., raster cells) are often references as 1 byte, two byte, etc.

<table>
<thead>
<tr>
<th>binary</th>
<th>decimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000000</td>
<td></td>
</tr>
<tr>
<td>00000001</td>
<td>1</td>
</tr>
<tr>
<td>00000010</td>
<td>2</td>
</tr>
<tr>
<td>00000011</td>
<td>3</td>
</tr>
<tr>
<td>00000100</td>
<td>4</td>
</tr>
<tr>
<td>00000101</td>
<td>5</td>
</tr>
<tr>
<td>00000110</td>
<td>6</td>
</tr>
<tr>
<td>00000111</td>
<td>7</td>
</tr>
<tr>
<td>00001000</td>
<td>8</td>
</tr>
<tr>
<td>00001001</td>
<td>9</td>
</tr>
<tr>
<td>00001010</td>
<td>10</td>
</tr>
<tr>
<td>00001011</td>
<td>11</td>
</tr>
<tr>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>

**Binary Columns**

- **Eights column**
- **Fours column**
- **Twos column**
- **Ones column**

1101

$8 + 4 + 0 + 1 = 13$
1
2
3
4
5
6
7
8
9
10
11
12
Cells Have a Type, Size

**Type** – e.g., Real, unsigned integer, signed integer, text

**Size** – 8 bit, 32 bit, 64 bit, long, short, character width

These control the size of the datasets, and type of data that may be stored.

Mixing types, sizes, often requires some care
Spatial Data Organization Information:

*Direct spatial Reference method: Raster

Raster object information:
* Image format: ESRI GRID
* Number of bands: 1

* Row count: 10632
* Column count: 15519
* Vertical count: 1

* Cell size X direction: 0.313653
* Cell size Y direction: 0.313653

* Bits per pixel: 32
* Pyramid layers: FALSE
* Image colormap: FALSE
* Compression type: Default

We can store a number up to $2^{32}$ or 4,293,967,296 in a cell.
Characters, assign a number to each character value, e.g.,

ASCII

(American Standard Coding for Information Interchange)

65 = A,  66 = B, 67 = C……..
97 = a,  98 = b, 99 = c ………

But

50 = 2
51 = 3
Positive Integers are easy

Positive Integers – “as is”, so interpret all 8 bits for the number
e.g., 8 bit unsigned integer
0 0 0 0 1 1 0 1 is the number 13, as in our previous example

Negative Integers – interpret 1st bit as sign bit,
e.g.,
0 is positive, 1 is negative number, and interpret next bits as number

8 bit signed integer
1 0 0 0 1 1 0 1 is -13
Floating point numbers (real numbers) are a bit more complicated

Generally a sign, exponent, and a fraction part
e.g., a 4-byte floating point number has the first bit as a sign, the next 8 as the exponent, and the next 23 as the fractional part,

S EEEEEEEE FFFFFFFFFFFFFFFFFFFFFFFFF

WHY SHOULD YOU CARE?
When you want this…

You get this:
Data and File Structures

Data often have specific organization to:
- reduce size
- speed access
- ease updates
Data and File Structures
Compression

Reducing size – e.g., raster run-length coding

<table>
<thead>
<tr>
<th>Raster</th>
<th>Run-length codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 9 6 6 6 6 6 6 7</td>
<td>2:9, 5:6, 1:7</td>
</tr>
<tr>
<td>6 6 6 6 6 6 6 6 6</td>
<td>8:6</td>
</tr>
<tr>
<td>9 9 6 6 6 6 7 7 7</td>
<td>2:9, 4:6, 2:7</td>
</tr>
<tr>
<td>9 8 9 6 6 7 7 7 5</td>
<td>1:9, 1:8, 1:9, 2:6, 2:7, 1:5</td>
</tr>
</tbody>
</table>
GIS Hardware
Typical of most computer systems, with a little more storage, a few added tools for input/output.

Coordinate input devices are the primary “specialized” hardware.

Digitizing Tablets

GPS receivers.
Summary

• GIS are systems for the creation, maintenance, analysis, and conveyance of spatial data

• We represent abstractions of our world into spatial and attribute components using data models and data structures

• Two major data models – raster and vector