INTRODUCTION TO GIS

Introduction

• GIS -
• GIS –
• GIS –
What is a GIS

- Geographic
  - of or relating to geography
    - the study of the physical features of the earth, its atmosphere, and of human activity as it affects and is affected by these

- Information
  - knowledge gained through study, communication, research, instruction, etc.

- System
  - a set of connected things or parts forming a complex whole
  - a set of principles or procedures according to which something is done; an organized scheme or method.

wikipedia
What is a GIS

• Spatial Data
  – Data that has inherent geographic or locational component to it
    • Coordinates of a corner
    • Zip codes
    • Boundaries
    • Area codes

• Non Spatial Data
  – Data without a geographic or locational component to it
  – Many times we can associate this data with Spatial Data
    • Iron Pipe located at a corner
    • +4 of a zip code
    • County
    • Line number for a phone
Some GIS Terminology

- Map
  - An all encompassing term
- Features
  - Items that make up the map
- Layers
  - Like Features are organized into individual layers
- Attributes
  - The non spatial data describing the features

Geospatial Data

- How data is organized in GIS software packages
  - Information in a map is usually divided into like themes or layers
  - Each layer is saved as an individual File
    - (ignoring geodatabases)
  - Example layers
    - Roads, Rivers, Lakes, Land use
- All layers fall within 1 of two types of Spatial Data models
Spatial Data Models

• Two general types of Spatial Data utilized by GIS packages
  – Vector
    • Points, lines, polygons
    • 0,1,2 dimensions respectively
  – Raster
    • A grid of cells or pixels
    • 2 dimensional
    • Represents an area

Spatial Data Models

• Vector Spatial Data Model Types
  – Points
  – Lines
    • Made up of at least 2 connected points
  – Polygons
    • Note, polygons must close
  – No single data layer can contain more than 1 of the above vector types
  – Point is the primitive unit
    • Can not be broken down further and still be vector
Spatial Data Models

- Vector Spatial Data Model
  - Topology
    - Spatial relationship between features and how they connect to one another
    - Important for any networks
      - Utilities
      - Roads
    - Separates a line from a polygon
  - Different file types have different levels of Topology (more on this later)
Spatial Data Models

• Vector Spatial Data Model
  – Attributes
    • An unlimited number of attributes allowed for each feature
    • Like attributes are organized into columns in the Attribute table
      – Simple to add new columns (fields)
      – Type of column defines what you can do with it
    • Each layer has its own Attribute Table

Spatial Data Models

• Vector Spatial Data Model
  – Attribute Table
    • Each feature’s attributes arranged into a single row in the table
      – 1 Feature = 1 Row in the attribute table
    • Columns types (like attributes) are defined when created
      – String
      – Float
      – Integer
      – Blob
Spatial Data Models

• Raster
  – A grid of pixels or cells representing an area
    • Each cell in the raster represents an area
  – Every pixel contains a single numerical value
  – Raster can have only 1 type of attribute
    • Elevation
    • Concentration
    • Land cover
    • -etc
Spatial Data Models

• Raster
  – Never any gaps in the raster
  – May be referred to as a surface
  – Simple, innate topology
  – Very useful for modeling
Spatial Data Models

• Raster
  – Values in raster cells are **always** numbers
  – Values may function as codes depending on the data the raster represents
    • Example landcover:
      – 21: Developed, Open Space
      – 22: Developed, Low Intensity
      – 23: Developed, Medium Intensity
      – 41: Hardwood forest
      – 90: Wetlands
    – Codes *should* be defined in the metadata

Spatial Data Models

• Raster
  – Imagery is usually comprised of multiple rasters
    • True color –
      – 1 raster storing brightness values for Red
      – 1 raster storing brightness values for Blue
      – 1 raster storing brightness values for Green
    • False color –
      – 1 raster storing brightness values for Near Infrared
      – 1 raster storing brightness values for Red
      – 1 raster storing brightness values for Green
Spatial Data Models
Spatial Data Models

• Raster
  – Landsat has 8 bands (1 raster for each band)
  – Hyperspectral imagery comprised of 200+ rasters

Spatial Data Models

• While Vector and Raster Data can be used in a map, they can not interact with each other
  – There’s always exceptions

• To combine data from one type with another one must convert
  – Vector -> Raster = Rasterization
  – Raster -> Vector = Vectorization
Common File Types

• Raster File Types
  – Tiff
  – IMG
  – GRID

• Vector File Types
  – Shapefiles
  – Coverages
  – Geodatabases
Common File Types

• Raster File Types
    • Not as common
      – JP2000 becoming more common for imagery
    • Highly compressed
      – May affect image quality
  – GRID – ESRI raster format
    • Complicated file structure
    • Fairly common
    • Somewhat compressed
    • Usually imports without issues

Common File Types

• Raster File Types
  – Many - many more
    • BIL
    • MrSID
    • DRG
    • BMP
    • RST
    • ...
    • ASCII
Common File Types

• Raster File Types
  – World File
    • OK, not really a raster
    • Used to georeference some rasters
      – *.tif and *.jpg most common
    • Simple text file accompanying the raster
      – *.tfw or *.jpw

Common File Types

• Vector File Types
  – Shapefile
    • Open format
    • Very Common
    • Poor Topology
    • Made up of multiple files
      – Not all are necessary
    • Must have these three
      – *.shp
      – *.dbf
      – *.shx
Shapefiles File Types

- .sbn and .sbx—The files that store the spatial index of the features.
- .fbn and .fbx—The files that store the spatial index of the features for shapefiles that are read-only.
- .ain and .aih—The files that store the attribute index of the active fields in a table or a theme’s attribute table.
- .atx—An .atx file is created for each shapefile or dBASE attribute index created in ArcCatalog. ArcView 3.x attribute indexes for shapefiles and
- .ixs and .mxs—Geocoding index for read/write shapefiles.
- .prj—The file that stores the coordinate system information (used by ArcGIS).
- .xml—Metadata for ArcGIS—Stores information about the shapefile.

Common File Types

- Vector File Types
  – Coverages
    • ESRI proprietary format
    • Fairly uncommon anymore
    • Excellent Topology
    • Complicated file structure
    • Difficult to use in most software
    • Strict rules for naming
Common File Types

• Vector File Types
  – ESRI export format *.e00
    • A safe way to move a coverage layer
    • More likely to find this than coverage itself
    • Has to be converted back to a coverage to be used
    • Essentially a giant text file

• Vector File Types
  – Geodatabase
    • Introduced by ESRI
    • Importing to other software can be a problem
    • Can store multiple layers
    • Topology is dependent on the user
    • Personal Geodatabase
      – Microsoft Access format *.mdb
Database Applications

• In order for a piece of software be termed a GIS package it requires a database management system
  – Connection must be made between spatial features and associated non spatial data

• An example: Shapefiles
  – *.shp – Feature geometry and Identifier
  – *.dbf – non spatial data – the attribute table

Database Applications

• Making the connections between shapes and data
  – ID field
    • Key field
    • Not editable in Software
    • Never use for a Join*
    • Types
      – PID
      – FID
      – OID
Database Applications

• Some Terminology
  – Fields – Columns
  – Tables – Relations
  – Joins – Relationships
  – Query – Subsetting attributes
  – SQL – Structured Query Language

• Joins
  – Connecting multiple tables through like attributes
  – Reduces redundancies
  – Types
    • 1 : 1
    • 1 : many
    • Many : 1
    • Many : many
Database Applications

- Joins

{Database image}

Database Applications

- Joins

{Database image}
Database Applications

• Joins

Applications

• Spatial Query
  – Locating features from one layer based upon their spatial relationship to features in another layer
    • Distance
    • Intersection
    • Contains
Applications

• Topography to Raster
  – Converting individual elevation points or contours to a raster

• Networks
  – Topology is very important
    • What is connected to what?
    • Direction
    • Timing
Applications

- Networks

Applications

- Record Keeping
Applications

• Aspect and Slope from a DEM
  – Aspect and Slope
    • Rate of change in X and Y directions determined for each cell based upon input elevation of the surrounding cells
  – Slope
    • $rise\_run = \sqrt{\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2}$
  – Aspect
    • $aspect = 57.29578 \times \arctan2\left(\frac{dz}{dy}, -\frac{dz}{dx}\right)$
Applications

- Reclass
  - Also referred to as Recoding
  - Changing some or all of the values in a raster to new values
    - Remember, has to be numbers since it is a raster
Applications

• Simple Hydrologic Analysis
  – Flow Accumulation
  – Flow Direction
  – Watershed Delineation
Applications

• GIS 335
Class Project
Applications

• Georeferencing
  – Establishing the location of a 2d or 3d object in real space
    • Giving real world coordinates to something that doesn’t have it
  – 1855 Map of Erie County
    • Medium 1 map ; 125 x 93 cm
    • Call Number G3803.E6 1855 .G4
    • Library of Congress Geography and Map Division
    • Digital Id g3803e.la000494
    • http://hdl.loc.gov/loc.gmd/g3803e.la000494 Library of Congress Catalog Number 2012593658

Metadata

• Data about data
  – Incredibly Important
  – Should be included with all data
  – Often overlooked by GISers
  – Something surveying community has all over GISers
Metadata

• Data about data
  – Attributes
  – Coordinate systems
  – How data was collected/created
  – Accuracies
    • Positional
    • Attributes
  – Point of contact
  – ...

Metadata

• Standards
  – FGDC
    • Content Standards for Digital Geospatial Metadata
      – CSDGM
    • International Organization for Standardization
      – ISO 19115-1:2014
Metadata

• Examples
  – CUGIR – Tompkins Building Outlines
  – Maine GIS – Conserved Lands (10/28/2014)
  – Click USGS – NJ Lidar

Questions?

• Russ Aicher
  - Assistant Professor
  - Survey Technology Program Coordinator
  - Paul Smith’s College
  - 518-327-6950
  - raicher@paulsmiths.edu