

# Geographic Information Analysis & Spatial Data

Lecture #1

---

---

---

---

---

---

---

---

## Outline

- Introduction
- Spatial Data Types: Objects vs. Fields
- Scale of Attribute Measures
- GIS and Spatial Analysis

---

---

---

---

---

---

---

---

## “Spatial Analysis” is a Key Term

- What is SA?: from data to information
  - **Added value!!!!**
- SA is identifiable over 4 areas:
  - Spatial data manipulation
  - Spatial data analysis
  - Spatial statistical analysis
  - Spatial modeling

---

---

---

---

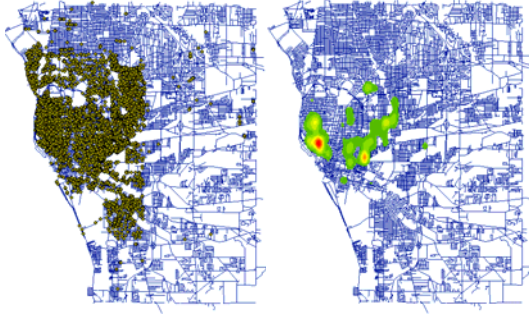
---

---

---

---

## From Data to Information!



Burglary 1996, Buffalo NY

---

---

---

---

---

---

---

---

## Spatial Data Manipulation

- Offers basic GIS tools for data **pre-processing** and preliminary data analysis such as:
  - Merge and split
  - Map overlay
  - Buffering
  - Spatial queries
  - ...



Buffer based on road network (1 km)

---

---

---

---

---

---

---

---

## Spatial Data Analysis

- **Exploratory** Spatial Data Analysis (ESDA)
  - Descriptive, understand data
  - First step in spatial analysis → **Hypothesis Suggesting**

**EDA:** advocates the use of a variety of techniques for examining data more effectively as the first step in statistical analysis and a precursor to more formal and structured data analysis

---

---

---

---

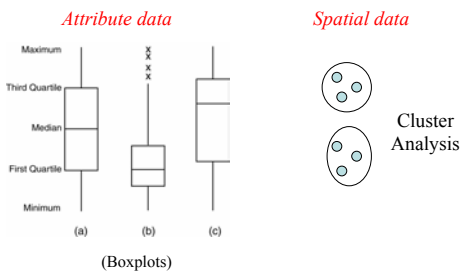
---

---

---

---

# Spatial Data Analysis



---

---

---

---

---

---

---

---

# Spatial Statistical Analysis

- **Confirmatory** Spatial Data Analysis (CSDA)
  - Measures data against an expected model
  - Determines whether data are 'typical' or 'unexpected' relative to a statistical data model → **Hypothesis Testing**
    - e.g. Spatial regression analysis

---

---

---

---

---

---

---

---

# Spatial Modeling

- Construct models to **predict** future outcomes
  - Human: e.g. spatial interaction models
  - Physical: processes of natural processes; e.g. Stormwater run-off models, soil erosion models ...



---

---

---

---

---

---

---

---

## How Are They Interrelated?

- Step 1: data is organized on a map (**manipulation**)
- Step 2: map is examined, explored (**data analysis**)
- Step 3: patterns are noticed and statistically analyzed (**statistical analysis**)
- Step 4: theories about future outcomes are formulated and tested (**modeling**)

---

---

---

---

---

---

---

---

## In Our Class

- **Geographic Information Analysis:** concerned with the **patterns** that arise as a result of **processes** that may be operating in space
  - “the introduction of techniques and methods to enable the representation, description, measurement, comparison, and generation of spatial **patterns**”

---

---

---

---

---

---

---

---

## GIA Is:

- **Represent** spatial data in a functional manner so that we can:
  - Generate and describe spatial **patterns**
  - Measure the statistics of spatial **patterns** presented in the manipulated map material
  - Compare **patterns** within individual and over multiple mapping areas
- Done in computers!!!
  - GIS

---

---

---

---

---

---

---

---

## GIA Covers...

- We will cover:
  - Spatial data manipulation
  - **Spatial data analysis**
  - **Spatial statistical analysis**
  - Spatial modeling

---

---

---

---

---

---

---

---

## Representing **Geography** in Computers?

---

---

---

---

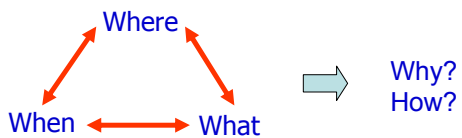
---

---

---

---

## Triad of Geography



- **Where:** locations
- **When:** time
- **What:** things → properties/attributes

---

---

---

---

---

---

---

---

## The Fundamental Problem

- Geographic information links a **place**, and often a **time**, with some **property** of that place (and time)
  - “The temperature at **34 N, 120 W** at **noon local time on 12/2/99** was **18 Celsius**”
- The number of **places (where)** and **times (when)** is vast, potentially infinite
  - The more closely we look at the world, the more detail it reveals: e.g. costal lines



---

---

---

---

---

---

---

---

## Representing Geography in Computers?

---

---

---

---

---

---

---

---

## The Fundamental Problem cont.

- Humans have found ingenious ways of dealing with this problem
  - Simplification, abstraction, generalization, conceptualization ...
  - Simply keep relevant and remove irrelevant!
  - **Data Models!!!**

---

---

---

---

---

---

---

---

## Objects vs. Fields

- Two ways of conceptualizing geographic variation and phenomena:
  - The most fundamental distinction in geographic representation
- Discrete **Objects**
- Continuous **Fields**

---

---

---

---

---

---

---

---

## Objects

- The world as a table-top
- Objects with well-defined boundaries:
  - Countable, persistent over time, perhaps mobile
  - Biological organisms: animals, trees
  - Man-made objects: vehicles, houses, fire hydrants



---

---

---

---

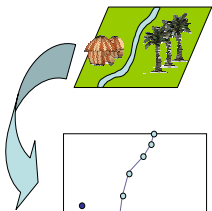
---

---

---

---

## The Vector Data Model



Real World

### Entity Types

**Points:**  $(x,y)$

**Lines:**  $(x_1,y_1) (x_2,y_2) \dots (x_n,y_n)$

**Areas (polygons):**

$(x_1,y_1) (x_2,y_2) \dots (x_1,y_1)$

---

---

---

---

---

---

---

---

## Types of Spatial Objects

- **Points:** object with no length in any direction:  $L_0$
- **Lines:** object with length in one direction:  $L_1$
- **Areas:** object with length in two directions:  $L_2$
- **Volumes:** object with length in three directions:  $L_3$

---

---

---

---

---

---

---

---

## Multiple Representation Problem

- Scale-dependent!!!!



Entity

- **Point Object?**
- **Area object?**
- **Volume object?**

---

---

---

---

---

---

---

---

## Fields

- Properties that vary continuously over space
  - Value is a function of location
  - Property can be of any attribute type, including direction
  - Elevation as the archetype
    - A single value at every point on the Earth's surface



---

---

---

---

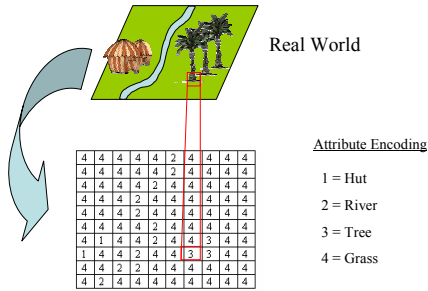
---

---

---

---

# The Raster Data Model



Still objected-based model; a cell → an object

---

---

---

---

---

---

---

---

---

---

# Other Examples of Fields

- Soil properties, e.g. pH, soil moisture
- Population density
  - But at fine enough scale the concept breaks down
- Identity of land owner
  - A single value of a nominal property at any point
- Name of county or state or nation
- Atmospheric temperature, pressure

---

---

---

---

---

---

---

---

---

---

# The IE Building. So What Is It?



---

---

---

---

---

---

---

---

---

---

## Choosing Data Representation

- The choice of viewpoint really depends on what you want to do with the data once they have been entered into the system
  - **Object** view: facility management such as buildings, roads...
  - **Field** view: environmental analysis such as fields of temperature, wind speed, etc
- In some cases, a geographic phenomenon may be represented by either raster or vector coding.
  - e.g. digital elevation can be represented with grid or TIN or contours.
- Some aspects of geographic reality are not well represented in either raster or vector data model
  - e.g. transportation network ( a new data model *networks*)

---

---

---

---

---

---

---

---

## The Fundamental Problem cont.

- The potential number of **properties (what)** is vast
  - In GIS we term them *attributes*
  - Attributes can be physical, social, economic, demographic, environmental, etc.

---

---

---

---

---

---

---

---

## Types of Attributes

- Stevens (1946) devised classification of measurement levels that recognize four levels:
  - Nominal measures
  - Ordinal measures
  - Interval measures
  - Ratio measures

*Low level, **Categorical** data*

*High level, **Numerical** measures, related to fixed unit scale*
- Higher → lower level conversion:  
**Classification**

---

---

---

---

---

---

---

---

## Nominal Measures

- Each value = a category;
  - Categories must be **inclusive** (all objects assignable to a category)
  - Mutually **exclusive** (each object only classifiable as one category)
- Mathematical manipulation is impossible (unless for frequency or area distributions)

---

---

---

---

---

---

---

---

## Ordinal Measures

- Consistent **ranking** of classes according to a certain criterion
- Know the order, not necessarily the difference between levels of order
- Not all mathematical operations can be applied



---

---

---

---

---

---

---

---

## Interval Measures

- Numerical differences or distances are defined on fixed equal units
  - e.g. the difference btw. category 1 and category 2 is 50,000 people
- Lacks a **definable zero** level, thus can only be used to measure absolute differences, **not** magnitudes of difference
- High level measurement can be reduced

---

---

---

---

---

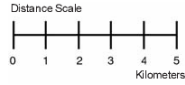
---

---

---

## Ratio Measures

- Have an inherent zero level
  - e.g. Distances
  - Can be mathematically manipulated
  - Higher level measurement
  - Can be converted to lower level such as ordinal




---

---

---

---

---

---

---

---

## Make Any Sense???




---

---

---

---

---

---

---

---

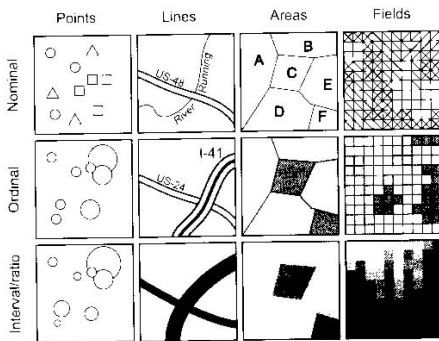


Figure 1.1 Schematic representation of entity-attribute spatial data types.

---

---

---

---

---

---

---

---

## Lecture Organization

- Classified by data types
  - Point data analysis
  - Line data analysis
  - Area data analysis
  - Field data analysis
- Classified by subject
  - SDM: Geog 417G
  - SDA: Geo 520
  - SSA: Geo 520
  - SM
- Multivariate data analysis
- Network analysis & Spatial Interaction Models
- New approaches

---

---

---

---

---

---

---

---

---

---

## Roles of GIS?

---

---

---

---

---

---

---

---

---

---

## GIS for GIA

- GIS usually claim to offer spatial analysis
  - Usually not mean statistical spatial analysis (both exploratory & confirmatory), but spatial data manipulation
- Reason:
  - Special requirement of specific fields
- Situation is changing

---

---

---

---

---

---

---

---

---

---

## Different Views on GIS

Container of maps	General public
Computerized tools for solving geographic problems	Decision makers, community groups
Spatial decision support system	Management scientists, operational researchers
Mechanized inventory of geographically distributed features and facilities	Utility managers, transportation officials, resource managers
Method for revealing patterns and processes in geographic information	Scientists, investigators
Tool to automate time-consuming tasks	Resource managers

---

---

---

---

---

---

---

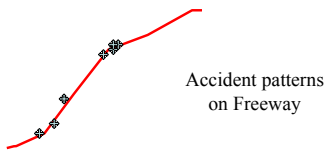
---

---

---

## GIS for GIA: Visualization

- Mapping and Visualization
  - *Scientific visualization*: explore data and information graphically as a means of gaining understanding and insight.
  - *Presentation* → spatial analysis tool




---

---

---

---

---

---

---

---

---

---

## Maps in Spatial Analysis Process

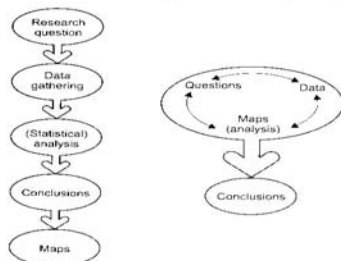


Figure 1.2 Changing role of maps in the analysis process.

---

---

---

---

---

---

---

---

---

---

## GIS for GIA: Visualization

- Data are now produced at rates faster than they can easily be converted into knowledge (data mining)
  - **Visual Data Mining**
- Some of the most exciting discoveries in science have been associated with nonlinear dynamics, or chaos theory, and complex behavior and structure can be readily appreciated when displayed graphically.
  - Visualization: in complex simulation models.

---

---

---

---

---

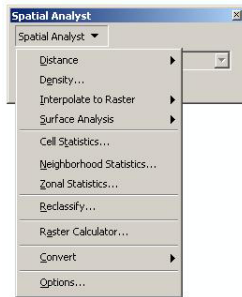
---

---

---

## Linking GIS and GIA

- **GIS** provides the form in which the data are presented for analysis
- **GIA** is primarily concerned with the patterns and processes attained from this analysis provided by GIS



---

---

---

---

---

---

---

---

## Review

- Define and distinguish the 4 areas of spatial data analysis: SDM, SDA, SSA, & SM
- Understand Object & Field views in GIS representation
- Understand the scales of attribute measures
  - Describe the 12 manifestations of the Entity-Attribute spatial data types.
- Understand a working definition of GIA and describe its relationship to GIS

---

---

---

---

---

---

---

---

## Dimensions

- Dimensions of variability (e.g. for a river: velocity, cross-sectional area) normally related to a scale of units, functions of three fundamental dimensions:
  - Mass (M)
  - Time (T)
  - Length (L)
- **Non-dimensional variables'** values are independent of the units involved
  - Proportions of a fixed total

---

---

---

---

---

---

---

---