UNIVERSITY OF GUJRAT

Web GIS

Lecture 02

Introduction to GIS

Dr. Usman Mazhar

usmanmazhargeo@outlook.com

In Today's Class

1. Define GIS
2. components of GIS
3. Differentiate spatial and non-spatial data
4. Understand raster vs vector data models
5. GIS Methods
6. Identify real-world GIS applications

1. GIS

- GIS is a technology that is used to create, manage, analyze, and map all types of data. (ESRI)

- A Geographic Information System (GIS) is a computer system that analyzes and displays geographically referenced information. It uses data that is attached to a unique location. (USGS)

- GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there). This provides a foundation for mapping and analysis that is used in science and almost every industry. GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication, efficiency, management, and decision-making.

What does GIS do?

Data management

- GIS is a foundational system of record. You can store and integrate information from business systems and authoritative sources, magnifying data’s usefulness.

Mapping and visualization

- From digital maps and dashboards to satellite imagery, 3D, and real time—GIS brings data to life, helping us understand problems and solve them.

Spatial analysis

- Most data has a location component—everything happens somewhere. With spatial analysis tools, we find hidden relationships and generate new insights from data.

Communication

- Maps and dashboards communicate complex ideas quickly. Science and data build common understanding, supporting collaboration and problem-solving.

2. Components of GIS

- Hardware
- Software
- Data
- People
- Methods

Hardware

- The physical devices that run GIS software, including computers, servers, scanners, and GPS units for data collection.
  - Computers
  - GPS devices
  - Servers

Software

- Applications (e.g., ArcGIS, QGIS) that provide tools to store, analyze, and visualize spatial data, typically using raster (images) and vector (points, lines, polygons) formats.
  - ArcGIS
  - QGIS

Data

- The most crucial component, consisting of spatial data (location) and attribute data (descriptive characteristics).
  - Spatial data
  - Attribute data

People

- Users, ranging from technical specialists who design and maintain the system to planners and managers who use it for analysis.
  - GIS Analysts
  - Developers
  - Planners

Methods/Procedures

- The techniques, workflows, and best practices applied to ensure accuracy in data modeling, analysis, and output.
  - Spatial analysis techniques
  - Workflows

3. Differentiate spatial and non-spatial data

- Spatial data (geospatial data) represents information linked to specific geographic locations on Earth, such as coordinates, shapes, and sizes (points, lines, polygons).
- Non-spatial data (attribute data) is descriptive information independent of location, such as names, types, or statistics.
- Spatial data is multi-dimensional and contextual, while non-spatial data is often one-dimensional and descriptive

Key Differences

- Definition: Spatial data describes where something is, while non-spatial data describes what it is.
- Format: Spatial data includes coordinates (x,y), shapes, and maps. Non-spatial data includes text, numbers, and tables.
- Analysis: Spatial data is analyzed using spatial relationships (e.g., buffering, overlay). Non-spatial data is analyzed using conventional statistical methods.

- Examples:
- Spatial: Latitude/longitude of a building, a road network map, parcel boundaries.
- Non-spatial: Name of building owner, number of lanes on a road, tax value of a parcel.

4. raster vs vector data models

Vector data

- Vector data represents geographic features using precise coordinates and is ideal for representing discrete, well-defined, or sharp-boundary objects.
- Vector data structures represent specific features on the Earth's surface, and assign attributes to those features. Vectors are composed of discrete geometric locations (x, y values) known as vertices that define the shape of the spatial object. The organization of the vertices determines the type of vector that we are working with is a point, line or polygon

Points

Each point is defined by a single x, y coordinate. There can be many points in a vector point file. Examples of point data include: sampling locations, the location of individual trees, or the location of cities (depending on the scale of a map).

Lines

- Lines are composed of many (at least 2) points that are connected. For instance, a road or a river stream may be represented by a line. This line is composed of a series of segments, each “bend” in the road or river stream represents a vertex that has defined x, y location.

Polygons

- A polygon consists of 3 or more vertices that are connected and closed. The outlines of survey plot boundaries, lakes, oceans, and states or countries are often represented by polygons. However, depending on the scale, if we ‘zoom-in’ a lot, we may represent rivers as polygons instead of lines since the water volume of a river stream covers a specific area (i.e. river width of 5 or 10 meters)

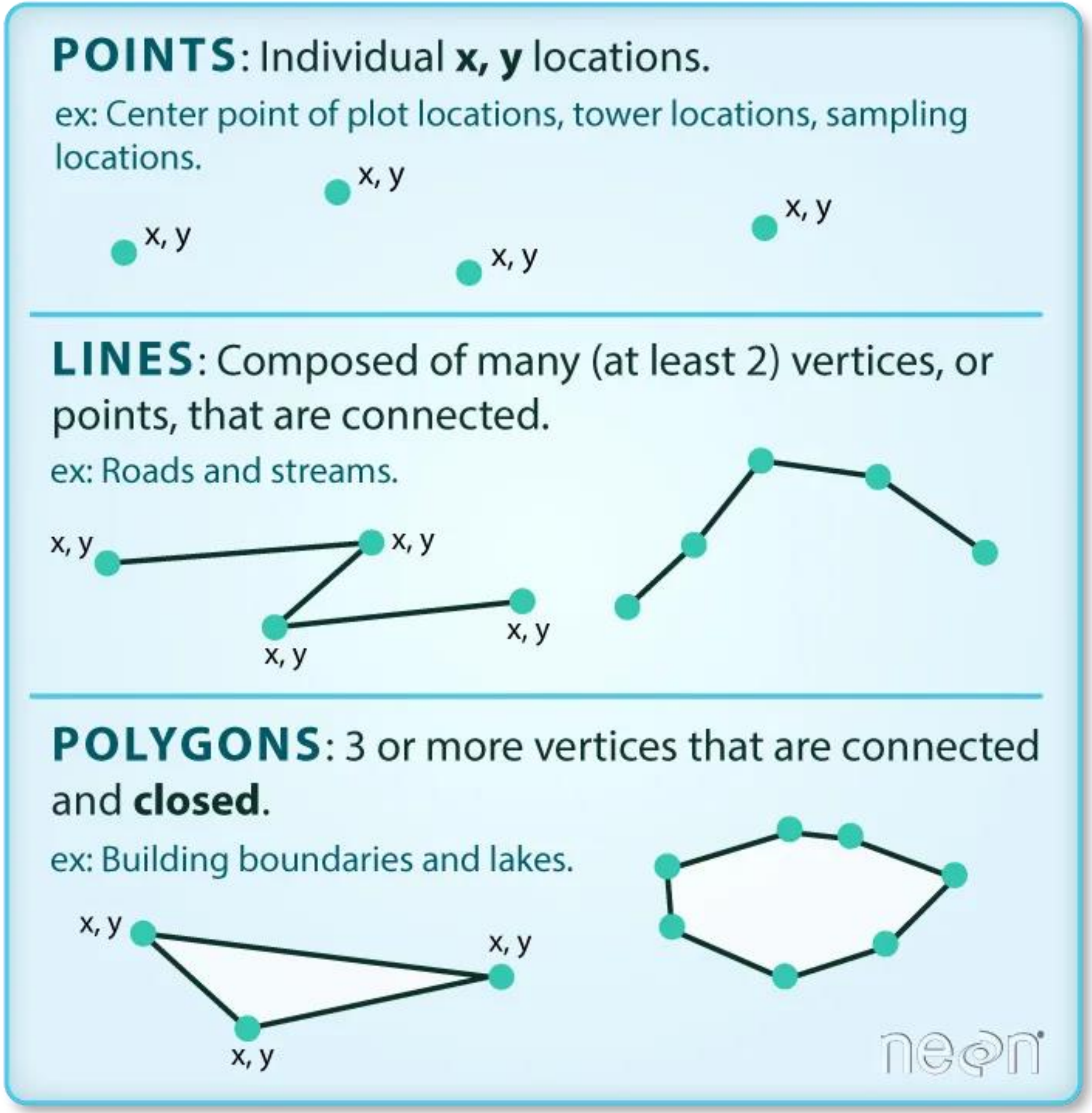

Image Source: National Ecological Observatory Network (NEON)

Raster data

- Raster data is any pixelated (or gridded) data where each pixel is associated with a specific geographical location. The value of a pixel can be continuous (e.g. elevation) or categorical (e.g. land use).
- A geospatial raster is only different from a digital photo in that it is accompanied by spatial information that connects the data to a particular location. This includes the raster's extent and cell size, the number of rows and columns, and its coordinate reference system (or CRS).

Some examples of continuous raster data include:

- Precipitation maps.
- Maps of tree height derived from LiDAR data.
- Elevation values for a region.

Some rasters contain categorical data where each pixel represents a discrete class such as a landcover type (e.g., “forest” or “grassland”) rather than a continuous value such as elevation or temperature. Some examples of classified maps include:

- Landcover / land-use maps.
- Tree height maps classified as short, medium, and tall trees.
- Elevation maps classified as low, medium, and high elevation.

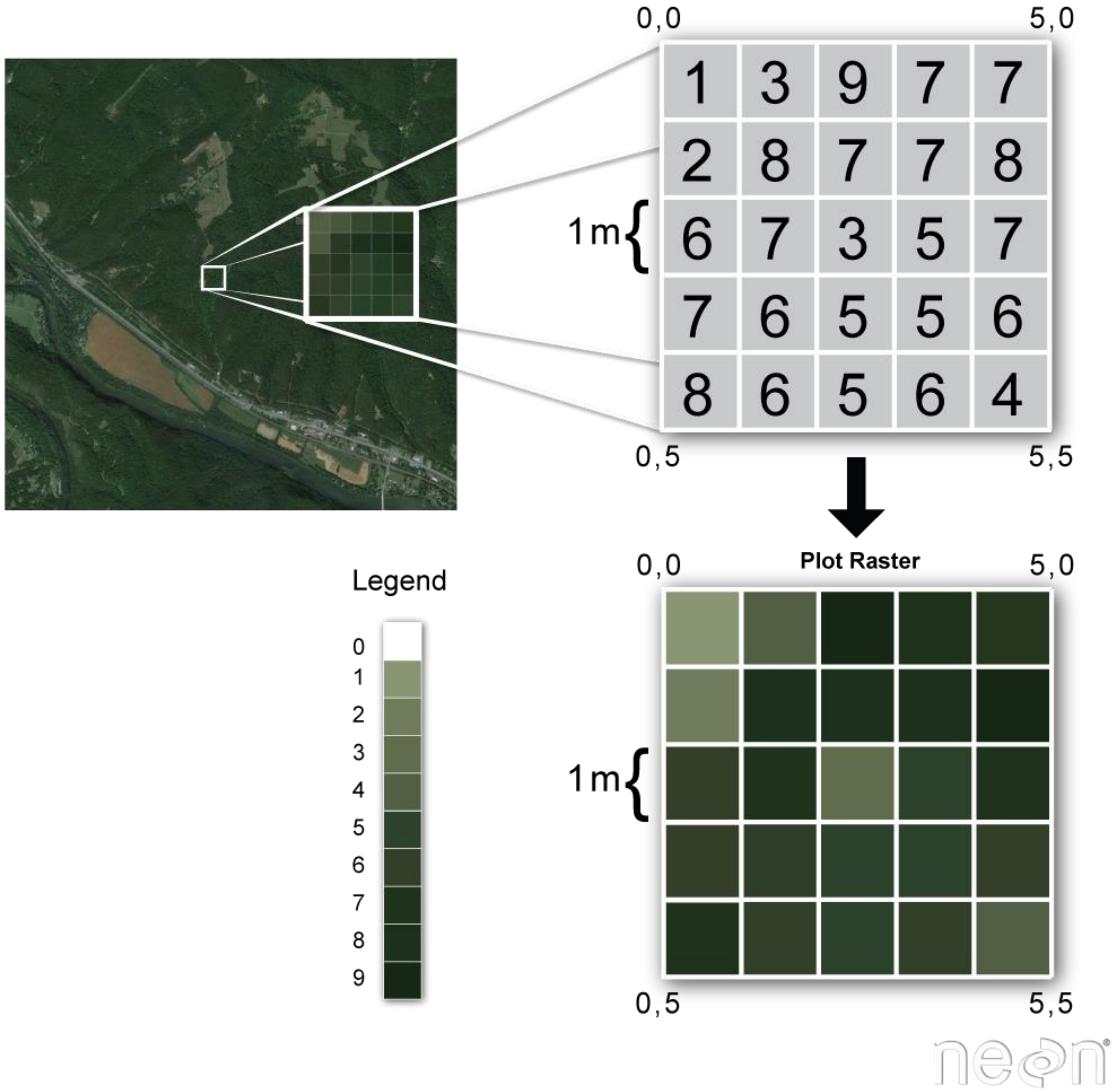

Image Source: National Ecological Observatory Network (NEON)

Point To Ponder

- What are the attributes of Raster Data

- What are the formats of Vector and Raster data

Other Categorizations

Attribute Data

- While spatial data describes the location, it is often paired with "attribute data" (or tabular data), which describes the characteristics of that spatial feature.

Geostatistical/Areal Data

- In spatial statistics, data is often broken down into point patterns (locations of events), geostatistical data (continuous variables measured at sample locations), and lattice/areal data (aggregated data within regions).

5. GIS Methods

- GIS operations/methods comprise data input, management, and spatial analysis to manipulate and analyze geographic information
- GIS methods provide the means to extract meaningful insights from spatial data, aiding decision-making processes.

Core GIS Methods

Buffer Analysis

- Buffer analysis is a fundamental spatial analysis technique used in GIS. It involves creating a zone or an area around a geographic feature, such as a point, line, or polygon, to represent a specified distance or range from that feature. This zone is referred to as a buffer zone.

Shortest Path Analysis

- Shortest path analysis is a fundamental operation in GIS and graph theory. It involves finding the shortest path or route between two points or locations within a network, considering factors such as distance, travel time, cost, or other attributes associated with the network.

Cluster Analysis

- Cluster analysis is a statistical technique used to group a set of objects in such a way that objects in the same group, called clusters, are more similar to each other than to those in other groups. The goal of cluster analysis is to identify inherent patterns or structures in the data without prior knowledge of any group memberships.

K-means Clustering

- K-means clustering is one of the most popular and widely used clustering algorithms in data science and machine learning. It's an iterative algorithm that partitions a dataset into K distinct, non-overlapping clusters. The "K" in K-Means represents the number of clusters the algorithm seeks to create, and it's a hyperparameter that needs to be specified by the user.

Nearest Neighborhood Analysis

- Nearest Neighbor Analysis (NNA) is a spatial analysis technique used in GIS to identify patterns and relationships within spatial data. It focuses on determining the proximity of features to one another based on their spatial locations.

Overlap Analysis

- Overlap analysis, also known as overlay analysis or spatial intersection analysis, is a GIS technique used to identify and analyze areas where different spatial datasets intersect or overlap. The primary goal of overlap analysis is to understand the spatial relationships between different geographic features and to derive insights from the intersections of these features.

Spatial Interpolation

- Spatial interpolation is a technique used in GIS and spatial analysis to estimate the values of unknown locations within a study area based on known values from surrounding locations. It’s particularly useful when dealing with irregularly distributed data points or when trying to create continuous surfaces or maps from point data.

6. Real-world GIS applications

- Urban Planning & Smart Cities: Zoning, land-use planning, infrastructure design, and managing city services.
- Environmental Management: Monitoring climate change, habitat tracking, analyzing environmental impact, and conservation planning.
- Disaster Management: Mapping flood zones, assessing wildfire risks, and planning emergency evacuation routes.

- Transportation & Logistics: Route optimization, fleet tracking, and logistics, reducing fuel consumption.
- Utilities Management: Mapping and maintaining water pipes, electric lines, and telecommunications infrastructure.
- Agriculture: Precision farming, crop monitoring, and managing land resources.

- Public Health: Tracking disease outbreaks, mapping health facility locations, and analyzing environmental health risks.
- Real Estate & Business: Site selection for new businesses, demographic analysis, and property valuation.
- Defense & Security: Intelligence gathering, terrain analysis, and logistics management.

Thank you!