Answerkey – Set A – GS2.401 – Spatial Informatics

Q1. Which one of the following is referred to as the first recorded example of spatial analysis?

- **Answer: C. Cholera outbreak in London**

Explanation: The first recorded example of spatial analysis is the study conducted by Dr. John Snow during the 1854 cholera outbreak in London. He mapped the locations of cholera cases and water pumps, identifying a connection between the outbreak and a contaminated water source, which led to significant developments in public health and spatial analysis.

Q2. Which of these is not a vector analysis function?

- **Answer:** A. Proximity analysis

Explanation: It is a raster, cell or object neighbourhood at equal intervals

Q3. Shape factor is highest for which of the following?

- **Answer: C. Starfish footprint**

Explanation: The shape factor is a measure of the complexity of a shape. A starfish footprint, with its irregular and extended shape, would have a higher shape factor compared to a square or circular footprint, which are more regular and compact.

Q4. Finding the nearest branch of the bank in which I have an account from IIITH campus, is what type of query?

- **Answer: C. Spatio-attribute query**

Explanation: This query involves both spatial and attribute components. You are looking for the nearest (spatial) bank branch that matches a specific attribute (the bank where you have an account). Therefore, this is a spatio-attribute query.

Q5. Which of the following can be represented as spatial object/s, either in Vector and Raster data models?

- **Answer:**
    - **A. Disease Occurrence**
    - **B. Pest infestation in agricultural fields**
    - **C. Rice Mill**
    - **D. Snow Avalanche**
    - **E. Over the road railway intersections**

Explanation: All of the options listed can be represented as spatial objects:

- **A. Disease Occurrence:** This can be mapped as points (vector) where cases are recorded or as a density surface (raster).
- **B. Pest infestation in agricultural fields:** This can be represented as polygons (vector) or as a continuous surface indicating the severity of infestation (raster).
- **C. Rice Mill:** Can be represented as a point or polygon (vector) indicating the location or area of the mill.
- **D. Snow Avalanche:** Can be mapped as areas only as a raster representing the likelihood or frequency of avalanches.
- **E. Over the road railway intersections:** Can be represented as points (vector) where intersections occur.

Q6. S in GIS refers to

- **Answer:**
  - **A. Systems**
  - **D. Services**
  - **E. Sciences**

Explanation:

- **A. Systems:** Geographic Information System (GIS) refers to a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data.

  D. Services: GIS provides various services related to spatial data analysis, such as mapping services.
- **E. Sciences:** GIS is often associated with the science of studying geographic phenomena.

Note: B. Scenarios is not typically what "S" in GIS refers to.

Q7. Which of the following can be represented as both Vector and Raster data models?

- **Answer:**
  - **F. Roads with widths**
  - **G. Property owned by individuals or organizations**
  - **H. A buffer of the air pollution around a road**
  - **J. National Forests**

Explanation:

- **F. Roads with widths:** Can be represented as lines or polygons (vector) or as raster cells indicating road presence and width.

- **G. Property owned by individuals or organizations:** Typically represented as polygons (vector), but could also be represented in raster form, with each cell indicating ownership.
- **H. A buffer of the air pollution around a road:** This can be represented as a buffer polygon (vector) or as a raster surface showing pollution levels.
- **J. National Forests:** Can be represented as polygons (vector) or as raster data indicating forested areas.

Note: I. Electricity power lines are generally represented as lines in vector data and less commonly in raster form, making them less suitable for both representations.

Q8. You are new to a city and want to know all the points of interest in your neighborhood. Do you think spatial analysis will help in such a case? [2+ (1.5x2) = 5m]

Answer: Yes, spatial analysis will be highly beneficial in this case. Spatial analysis allows you to visualize, explore, and understand the spatial relationships between different points of interest (POIs) within your neighborhood.

Data to Collect:

1. **Spatial Data:**
   - **Location of POIs:** Geographic coordinates (latitude and longitude) of all points of interest such as parks, schools, hospitals, restaurants, grocery stores, etc.
   - **Neighborhood Boundary:** The boundary of your neighborhood as a polygon.

     Use this to identify if the shape/extent of the neighbourhood is recognised by the responder. Radial (like a buffer) or road network distance or semantic by name of the locality ==> this gets 0.5marks
   - **Road Network:** Streets and pathways within the neighborhood.
2. **Non-Spatial Data:**
   - **Attributes of POIs:** Information about each POI, such as the type of facility (e.g., hospital, school), operating hours, services offered, etc.
   - **Demographic Data:** Information about the population within the neighborhood, such as age distribution, income levels, etc.

Storage Plan: Spaghetti model, DIME or Arc/Node - which of these storage data structures?

- **Spatial Data:**
  - Use a **Geodatabase** (such as PostGIS or Esri Geodatabase, geotiff) to store spatial data.
  - POIs can be stored as points in a spatial layer, with their geographic coordinates.
  - The neighborhood boundary and road network can be stored as polygon and line layers, respectively.

- **Non-Spatial Data:**
    - Store attribute information in a **relational database** or as attribute tables linked to the spatial data. For example, the attributes of POIs can be stored in a table that is linked to the points in the spatial layer via a unique identifier.

Queries/Analysis Methods: Only state the method 0.5marks; description is the 1mark

1. **Spatial Query:**
    - **Buffer Analysis:** Create a buffer around your location (or neighborhood boundary) to find all POIs within a certain distance (e.g., 1 km). This will help identify which POIs are nearby.
2. **Overlay Analysis:**
    - **Spatial Overlay:** Use spatial overlay techniques to intersect the POIs layer with the neighborhood boundary to filter out only those POIs that fall within your neighborhood. This ensures that you are only considering relevant points of interest.(Topological)
3. **Map/ Cartographic Query :** Identifying Find, locate (Slides)
4. **Spatio-attribute query**

Q9. Briefly explain why.

a) Minkowski distance can be used as a generalized distance measure: [2m = 0.5 for yes/no; 1.5 for formula and explanation]

- **Explanation:**
    - Minkowski distance is a generalized distance metric that encompasses both Euclidean and Manhattan distances. It is defined by the equation:

D(x,y) = (Summation of (((x_i – y_i) ^ (p))) over i belonging to (1,n)) ^ (1/p)

- By adjusting the parameter p, you can derive different types of distances:
  - **When p=1:** It becomes the Manhattan distance, which measures the distance between points along axes at right angles.
  - **When p=2:** It becomes the Euclidean distance, the "straight line" distance between two points.
- **Generalization:** Because Minkowski distance can adjust to different p values, it provides a flexible framework for measuring distance in different scenarios, making it a powerful tool in spatial analysis.

b) Continuous Spatial objects cannot be captured using vector data model: [2m = 0.5 continuous; vector as a format = 0.5m; explanation = 1m; optional/grace if someone uses an example to respond]

- **Explanation:**
  - **Continuous Spatial Objects:** These are phenomena that vary smoothly over space, such as elevation, temperature, or pollution levels.
  - **Vector Data Model:** This model represents data using discrete geometric shapes (points, lines, and polygons). It is ideal for discrete objects like buildings, roads, and boundaries.
  - **Limitation:** The vector model is not well-suited for representing continuous data because it relies on discrete, often arbitrary **boundaries**. Continuous phenomena, by nature, do not have distinct boundaries, making it difficult to accurately capture them in a vector format.
  - **Alternative:** Continuous spatial data is better represented using the **raster data model,** where the phenomenon is mapped across a grid of cells, with each cell containing a value representing the phenomenon's intensity or level.

Q10. On a Map, can we represent a campus like IIIT Hyderabad as a point feature and also as a polygon feature? [3m = 0.5 for keyword scale and its meaning; then 1mx2 for small scale vs large scale aspects; 0.5 for the right explanation]

Answer: Yes, IIIT Hyderabad can be represented both as a point feature and as a polygon feature on a map, depending on the context and scale of the map.

- **Point Feature:**
  - **When to Use:** When the map is of a small scale, such as showing the whole state/city in a map, where detailed representation of individual buildings or campus boundaries is unnecessary. The entire campus is represented as a single point, usually at the center of the campus.
  - **Why:** This simplifies the map and prevents clutter, making it easier to read at smaller scales.
- **Polygon Feature:**

- **When to Use:** When the map is at a large scale, such as a neighborhood or detailed campus map, where the exact boundaries and layout of the campus are important. The campus is represented as a polygon showing its exact shape and boundaries.
- **Why:** This provides a detailed and accurate representation of the campus, allowing for more precise spatial analysis within or around the campus.

Q11. A dataset that shows the annual average solar radiation available for each locality in the city is given. You are given the task of prioritizing the areas where solar panels can be deployed based on the radiation levels. Describe how you will make a map for each year showing the areas where the solar panel deployment will be done in that year? [3m = 2m for the attribute division approach + 1m for map output (binary colour option for each year) or reclassify]

Answer:

- **Step 1: Data Analysis**
  - **Analyze the Dataset:** Begin by analyzing the dataset to understand the distribution of solar radiation across the city. Identify areas with the highest levels of solar radiation as they would be the most effective for solar panel deployment.
- **Step 2: Classification and Prioritization**
  - **Classify Localities:** Divide the city into classes based on the levels of solar radiation, using methods like Natural Breaks or Standard Deviation to categorize localities into "High," "Medium," and "Low" solar radiation zones. --> attribute generalisation method
  - **Prioritization:** Given the limited budget and time, prioritize the "High" radiation areas for the first year, followed by "Medium" radiation areas in subsequent years.
- **Step 3: Create Annual Deployment Maps**
  - **Year 1 Map:** Highlight the "High" solar radiation areas. These areas will be marked for immediate solar panel deployment.
  - **Year 2-5 Maps:** Each subsequent year, create a map showing the next set of localities, focusing on "Medium" radiation areas. Continue this until all prioritized areas are covered.
- **Step 4: Consider Constraints**
  - **Additional Factors:** Consider other factors like proximity to infrastructure, cost of installation, and potential energy savings. Adjust the prioritization accordingly if some high-radiation areas are not feasible for immediate deployment.
- **Step 5: Update and Review**

- **Annual Review:** Each year, review the progress and adjust the maps based on any new data or changes in priorities. This might involve updating the priorities based on changes in budget, technological advancements, or shifts in energy demand.

Answerkey – Set B – GS2.401- Spatial Informatics

Question 1: Not present in class,not sure;

B. Percentage of Building fully collapsed over the region

Q2. Which of these is not a geometric primitive?

- **Answer: D. Curves**

Explanation: In geometry, **geometric primitives** are the basic shapes or entities from which more complex shapes are constructed. The most common geometric primitives are:

- **A. Point:** Represents a location in space with no dimensions.
- **B. Line:** Represents a straight path between two points.
- **C. Polygon:** Represents a closed shape with three or more sides.

Curves are not considered geometric primitives; they are more complex shapes that can be constructed from these basic primitives.

Q3. The path that aircrafts take to fly is called:

- **Answer: A. Geodesic distance**

Explanation: The **geodesic distance** is the shortest path between two points on the surface of a sphere, such as the Earth. Since aircraft often fly long distances over the curved surface of the Earth, the geodesic path represents the most efficient route. It accounts for the curvature of the Earth, unlike Euclidean distance, which assumes a flat surface.

Q4. Identifying whether the road I am traveling on is passing over a water body is what type of query?

- **Answer: B. Spatial Query**

Explanation: A **spatial query** is a type of query that involves the spatial relationships between objects in a geographic space. In this case, the query is about the spatial relationship between the road and the water body, specifically whether the road intersects or passes over the water body. This makes it a spatial query, as it is concerned with the location and intersection of these geographic features

Q5. Which of the following can be represented as spatial object/s, either in Vector and Raster data models?

Answer:

- **A. Individual House Plots**
- **B. Flood and its impact**
- **D. Crop sowing area**
- **E. Forest fire**

Explanation:

- **A. Individual House Plots:** Can be represented as polygons (vector) to show boundaries or as raster data to indicate the presence or absence of plots over a grid.
- **B. Flood and its impact:** Can be represented as a polygon (vector) showing the affected area or as a raster indicating the extent and severity of the flood.
- **D. Crop sowing area:** Can be represented as polygons (vector) showing different crop areas or as raster data indicating the extent of crop coverage.
- **E. Forest fire:** Similar to floods, can be represented as polygons (vector) to show affected areas or as raster data to indicate the spread and intensity of the fire.

Note: C. Under the road railway crossing is typically represented as a point or line (vector), and it does not translate well into raster data, so it is less applicable in both vector and raster formats.

Q6. When preparing a Map layout, it is always important to have these on the map:

Answer:

- **C. Legend having symbols/colours/etc**
- **D. Scale of the map**
- **E. Title indicating the thematic content displayed**

Explanation:

- **C. Legend having symbols/colours/etc:** Essential for interpreting the symbols and colors used in the map.
- **D. Scale of the map:** Necessary to understand the real-world distances represented on the map.
- **E. Title indicating the thematic content displayed:** Clearly defines the subject and purpose of the map.

Note: A. Direction with a South Arrow is not typically used; instead, a **North Arrow** is standard on maps to indicate direction. Therefore, including a South Arrow would not be appropriate.

B. An inset map indicating where on the Globe it is: Helps the viewer understand the broader context of the map's location. As this is optional and some may do it or not