MAP Digitisation and Data Ingest

Vinay.S
EWRG, CES, IISc
Overview

• Paper to Digital Map
• Digitising
• Field Data Collection
• Data ingest
Paper Map and Digital Map

• Paper Map
• Maps in the form of hard copies

• Digital Maps
• Maps as soft copies
  – Those can be viewed directly on computers.
  – Modification with respect to scale.
  – Zoom/Move to a particular portion.
Scanning

- Scanners are the devices that convert analog data to digital grid based images.
- Used to capture the geographical data in a very high resolution raster format that are further processed to develop vector data.
- Light source is used to scan and record the maps in digital format.
Scanning

• Digital images are made of grids with a reflectance values

• Scanned images would contain smudges and defects of the original map

• Needs enhancement (contrast, brightness etc.) if the image quality is low.
Paper Map to Digital Map

Paper Map → Conversion → Digital Map

Scanning → Printing
Need for conversion

Most of the ancillary data's are available in analog versions, conversion is necessary to:

1. To prepare GIS database.

2. Vectorisation as Point, Line and Polygon for further analysis. *Example*: Change detection

3. Paper maps are susceptible to physical damage.
Digitising

- Process of representing features as points, lines and polygons.

Example: Light towers, Bus stops, Schools etc

Example: Road network, Stream Network etc.

Example: Forests, political boundaries etc
Reading Coordinates in toposheet

1: 50000

- Scanned toposheets have no geographical reference i.e., they are in the rectangular coordinate system
Features in a toposheet

**Line Features:** Contours, Roads, Railway lines, Streams, Drainages, etc

**Polygon Features:** Tanks, Forests, Reservoirs, towns and villages
Digitising Features

• Digitisation features is based on the scale of the work
• Example: A building on a small scale map is represented as a point, where as on a large scale map a building is represented as polygon.
• For linear features, end points are sufficient to represent, rather than having multiple points in between.
• Features must obey the topology rules
Topology Rules

• Points shall not overlap one over the other

• Linear features shall have end nodes only other than intersections

• Features shall be connected to each other.

• Line features shall not have dangles or overshoots

• Polygons must not overlap each

• Polygons must not have gaps

• Sliver polygons should not exist

...............etc
Field Data Collection

- Information to be collected - about the feature of interest.
- Example: **Trees**: location of the tree, DBH of the tree, canopy area of the tree, tree species-family...etc.

- The coordinate information about the location (point), length of the road (line), areal extent (polygon) are collected using GPS
GPS

• GPS – Global Positioning System
• Gives the x, y and z coordinates about a point
• Uses the satellites to locate the position (triangulation).
• Application
  – Navigation
  – Location of a Feature
  – Velocity measurements
  – Altitude Measurements
Using GPS in Field data collection

• In the field GPS is used to identify the location of the feature.
• Waypoints are used to mark the locations and store in GPS.
• For each (known) waypoint, attribute data including the description of the feature/location are identified.
• GPS needs to be pre calibrated, uncertainty exists in closed areas where signals are low.

Example:

<table>
<thead>
<tr>
<th>Latitude (Degree)</th>
<th>Longitude (Degree)</th>
<th>Elevation (meters)</th>
<th>Id</th>
<th>Description</th>
<th>Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.9694</td>
<td>77.5935</td>
<td>910</td>
<td>A1</td>
<td>Kanteerava Stadium</td>
<td>40000</td>
</tr>
</tbody>
</table>
Data Ingest

• Data ingestion is the process of obtaining, importing, and processing data.

• Process: involves altering individual files by editing their content and fit into a larger document.

• Concept of database is used in the process of joining the datasets using the primary key and foreign key.
Database

• Database is a collection of structured data, independent of any particular application.

  – Data: consists of unorganised facts and entities
  – Information: Organised data
  – Knowledge: information used to process, and draw conclusions

• Database management systems (DBMS) are computer programs used to organise and manage the database.
Data Ingest Examples

• Recorded field data to the GPS (Spatial) data.

• Attaching Population database to Village layer (based on the village id)

• Attaching the attributes about the road
  – Attaching photo of pot holes on the road with the location and description on web

• Rain Gauge stations with rainfall data at each station
Database ingest

<table>
<thead>
<tr>
<th>OBJECTID</th>
<th>Rain gauge id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>270201</td>
</tr>
<tr>
<td>2</td>
<td>270401</td>
</tr>
<tr>
<td>3</td>
<td>270402</td>
</tr>
<tr>
<td>4</td>
<td>270403</td>
</tr>
<tr>
<td>5</td>
<td>270603</td>
</tr>
<tr>
<td>6</td>
<td>270606</td>
</tr>
<tr>
<td>7</td>
<td>270903</td>
</tr>
<tr>
<td>8</td>
<td>270907</td>
</tr>
<tr>
<td>9</td>
<td>270902</td>
</tr>
<tr>
<td>10</td>
<td>270904</td>
</tr>
<tr>
<td>11</td>
<td>270908</td>
</tr>
<tr>
<td>12</td>
<td>270909</td>
</tr>
<tr>
<td>13</td>
<td>270910</td>
</tr>
</tbody>
</table>

Rainfall database at each rain gauge stations

Joining two database

Spatial Database

Non Spatial Database

<table>
<thead>
<tr>
<th>RGS_id</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>Septembe</th>
<th>October</th>
<th>Novembe</th>
<th>December</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td>270201</td>
<td>24.9</td>
<td>8.3</td>
<td>23.8</td>
<td>33.4</td>
<td>174.6</td>
<td>1117.7</td>
<td>1271.3</td>
<td>827.8</td>
<td>397.9</td>
<td>206.0</td>
<td>72.7</td>
<td>34.2</td>
<td>4192.7</td>
</tr>
<tr>
<td>270401</td>
<td>6.6</td>
<td>0.3</td>
<td>15.6</td>
<td>15.9</td>
<td>124.5</td>
<td>986.6</td>
<td>1125.7</td>
<td>701.8</td>
<td>350.6</td>
<td>171.5</td>
<td>65.0</td>
<td>32.2</td>
<td>3596.1</td>
</tr>
<tr>
<td>270402</td>
<td>4.7</td>
<td>0.0</td>
<td>9.1</td>
<td>9.3</td>
<td>157.3</td>
<td>1122.0</td>
<td>1588.6</td>
<td>1090.9</td>
<td>400.6</td>
<td>147.2</td>
<td>69.3</td>
<td>8.3</td>
<td>4607.2</td>
</tr>
<tr>
<td>270403</td>
<td>17.9</td>
<td>0.0</td>
<td>7.3</td>
<td>27.3</td>
<td>166.1</td>
<td>1006.8</td>
<td>1123.6</td>
<td>799.2</td>
<td>360.9</td>
<td>188.3</td>
<td>65.1</td>
<td>21.2</td>
<td>3783.6</td>
</tr>
<tr>
<td>270603</td>
<td>6.1</td>
<td>22.6</td>
<td>26.8</td>
<td>22.6</td>
<td>138.0</td>
<td>1074.0</td>
<td>1234.2</td>
<td>725.8</td>
<td>378.6</td>
<td>172.9</td>
<td>66.9</td>
<td>32.4</td>
<td>3900.9</td>
</tr>
<tr>
<td>270601</td>
<td>17.9</td>
<td>9.4</td>
<td>34.0</td>
<td>24.0</td>
<td>135.6</td>
<td>944.5</td>
<td>1035.6</td>
<td>632.1</td>
<td>321.9</td>
<td>208.5</td>
<td>86.0</td>
<td>20.0</td>
<td>3469.6</td>
</tr>
<tr>
<td>270903</td>
<td>28.3</td>
<td>0.0</td>
<td>30.6</td>
<td>29.5</td>
<td>127.2</td>
<td>1220.3</td>
<td>2210.6</td>
<td>1541.8</td>
<td>468.6</td>
<td>197.7</td>
<td>117.6</td>
<td>37.1</td>
<td>6009.2</td>
</tr>
<tr>
<td>270907</td>
<td>9.3</td>
<td>5.6</td>
<td>31.3</td>
<td>33.4</td>
<td>72.1</td>
<td>615.8</td>
<td>1058.6</td>
<td>652.3</td>
<td>208.9</td>
<td>145.5</td>
<td>68.5</td>
<td>19.2</td>
<td>2920.4</td>
</tr>
<tr>
<td>270902</td>
<td>8.3</td>
<td>11.0</td>
<td>46.4</td>
<td>22.8</td>
<td>97.2</td>
<td>527.8</td>
<td>1034.9</td>
<td>617.7</td>
<td>223.4</td>
<td>181.8</td>
<td>69.8</td>
<td>21.5</td>
<td>2862.6</td>
</tr>
<tr>
<td>270904</td>
<td>22.7</td>
<td>10.7</td>
<td>37.7</td>
<td>52.6</td>
<td>62.3</td>
<td>283.1</td>
<td>534.3</td>
<td>320.1</td>
<td>137.5</td>
<td>146.5</td>
<td>98.6</td>
<td>25.6</td>
<td>1731.7</td>
</tr>
<tr>
<td>270908</td>
<td>33.2</td>
<td>9.9</td>
<td>39.3</td>
<td>67.0</td>
<td>91.3</td>
<td>374.2</td>
<td>806.5</td>
<td>531.3</td>
<td>160.7</td>
<td>129.0</td>
<td>73.5</td>
<td>17.8</td>
<td>2333.7</td>
</tr>
<tr>
<td>270909</td>
<td>12.3</td>
<td>12.1</td>
<td>16.4</td>
<td>25.2</td>
<td>71.1</td>
<td>530.7</td>
<td>857.6</td>
<td>516.9</td>
<td>203.0</td>
<td>131.2</td>
<td>66.9</td>
<td>21.5</td>
<td>2465.0</td>
</tr>
<tr>
<td>270910</td>
<td>5.0</td>
<td>14.3</td>
<td>18.5</td>
<td>31.6</td>
<td>65.8</td>
<td>511.8</td>
<td>925.5</td>
<td>502.6</td>
<td>186.2</td>
<td>132.9</td>
<td>61.7</td>
<td>21.5</td>
<td>2477.3</td>
</tr>
</tbody>
</table>
Thank you