Chapter III

An Introduction to GIS
(All Things Spatial)

The last chapter provided an introduction to two of the key components in a GIS: getting the data in, and then manipulating them to answer questions. This chapter considers how these data can be visualized and analyzed. The analysis section in this chapter will be fairly simplistic, as this topic is revisited again in Chapters V and VI with the introduction of increasingly more sophisticated measures. The chapter will end with a discussion of an important topic integral to urban area analysis but which is often missed when discussing the GIS: What exactly is a neighborhood and how should it be spatially defined?

Visualizing the Data

The most widely used role of a GIS in public health (or for that matter in any government agency) is as a visualization tool. Data of a GIS is spatial in nature, and spatial data basically means maps. You can use the GIS to make maps, and simply put, people understand maps far better than tables of numbers.

An “introduction to cartography” lecture in a GIS course will often present the reasons why maps displaying the same general information can vary. For example, consider a map showing citywide infant mortality. Why would maps of such a specific topic differ? One reason is the skill of the mapmaker; for example, if the map displays graduated colors (a choropleth map), the choice of colors,
data classification type and breaks chosen, and the geographic aggregation mapped will all result in different surfaces. Aggregation problems have already been discussed in Chapter II with another issue being addressed at the end of this chapter — namely, how to spatially define a neighborhood.

The second reason for map variation is the range of data available. For example, the relationship (or lack thereof) between the mapmaker and the Vital Statistics Department will influence how spatially disaggregated these data are. Address level data opens other forms of visualization, including hot spot displays and infant mortality rate surfaces (these will be discussed in Chapter V). If the data is only released by zip code, the resulting types of map display are limited. It should also be pointed out that if a vital statistics department only releases data at the zip code level, sizeable error can occur if they rely on the birth and death certificate, do not geocode the addresses, and then aggregate these up to the zip code level. In other words, the address listed on the certificate does not match the zip code listed. Table 1 displays a summary table for infant death data for the Baton Rouge Healthy Start Region over a 3-year period. The *zip named* column lists the frequency of that zip code being listed on the birth certificate (linked to the death certificate). The *geocode* column lists how many times a death was aggregated to a zip code by finding where the listed street address falls onto a zip code map. The *unmatch* column lists how many of these addresses listing one of the named zip codes could not be geocoded. The *% error* column shows the difference between using the named zip codes, as compared to actually mapping the addresses (unmatched records have been subtracted from the *zip named* column in the calculation). The resulting error is large (obviously we are dealing with small numbers), yet resulting maps of infant mortality will look quite different.

If the mapmaker is connected to a health organization, such as a Healthy Start program, data accessibility will again be improved, resulting in more spatially disaggregated information and more quality controls in data recording, resulting in greater confidence and accuracy in the findings.

The third reason for map variation results from the software and hardware available to the mapmaker. Ten years ago (early 1990s) the software costs and computing power needed to run most GIS packages (usually a workstation), would limit GIS use to academic departments. Today all costs have dropped to the point of widespread availability. A suitable PC, with printer, GPS unit, and

<table>
<thead>
<tr>
<th>Zip Named</th>
<th>Geocode</th>
<th>Unmatch</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>70805</td>
<td>28</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>70806</td>
<td>17</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>70807</td>
<td>20</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>70811</td>
<td>12</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>70812</td>
<td>13</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>
Related Content

Challenges of Semantic 3D City Models: A Contribution of the COST Research Action TU0801
www.irma-international.org/article/challenges-of-semantic-3d-city-models/138264/

Location Management in PCS Networks Using Base Areas (BAs) and 2 Level Paging (2LP) Schemes
www.irma-international.org/chapter/location-management-pcs-networks-using/70516/

Practical Approaches to Spatial Estimation of Disaster-Affected Populations
www.irma-international.org/article/practical-approaches-spatial-estimation-disaster/45129/

The Development of the Environmental Impact Assessment Process in Botswana
www.irma-international.org/chapter/the-development-of-the-environmental-impact-assessment-process-in-botswana/187738/

Three-Dimensional Information Retrieval (3DIR): Exploiting 3D Geometry and Model Topology in Information Retrieval from BIM Environments
www.irma-international.org/article/three-dimensional-information-retrieval-3dir/171614/