THE APPLICATION OF AUTOMATED STATISTICAL MAPPING TO HEALTH STATISTICS

Henry Tom
Computer Programmer
Computer Graphics Staff
Geography Division
U.S. Bureau of the Census

Statistical maps provide a spatial summary capable of portraying and analyzing statistics. The recent emergence of automated statistical mapping has increased these capabilities notably with greater flexibility. This development affords health planners a basic input for decisionmaking; however, effective applications depend upon a fundamental understanding of computer mapping. To further such an understanding, my purpose is to explain how computer mapping concepts/techniques are integrated and applied to spatial display, generation of cartographic artwork for publication, and spatial analysis of digital cartographic files at the Bureau of the Census.

I. Census Computer Mapping

Census computer mapping is based on two inputs; a digital cartographic file and a data file. The former is comprised of map areas translated into a form computers can recognize and replicate, while the latter consists of data values for those same areas. A digital cartographic file usually consists of locational coordinates for sets of points; each set of points delineating an areal boundary. Locational coordinates are produced through digitizing, i.e., automated recording of x/y positions for areal boundary points on a locational coordinate system. Each area in the digital cartographic file is identified by a geocode, generally a unique numerical identifier, e.g., a Federal Information Processing Standards (FIPS) county code. The data file is composed of values and their geocodes for the variable to be mapped. It is assumed a common set of geocodes is used for both the digital cartographic file and data file; a significant requirement in the computer mapping program.

The computer mapping program matches geocodes from both input files to correctly associate data values to their areas. Depending upon mapping requirements, the scaling, translating, rotating, conversion of coordinates, classing of data values and additional processing may be performed within the computer mapping program. The map drawing instructions producing the actual map are called from a graphics software package. Graphics software is normally specific to the designated peripheral output device. In essence, these subroutines interpret map drawing commands into the appropriate instructional code for a given device.

The generation of Census computer mapping output occurs in either an on-line or off-line mode. At the Bureau, on-line refers to immediate map display, while the off-line process diverts maps to a storage device such as a magnetic tape for subsequent display. The on-line mode involves either a high speed line printer or graphics display terminal, which may have a hard copier attached. High speed line printer maps result from overprinting alphanumeric characters. A graphics display terminal produces map images on a cathode ray tube (CRT). In the off-line mode, a "plot tape" is created according to the graphics software for a flatbed pen plotter or computer output on microfilm (COM) unit. The flatbed pen plotter at the Bureau has a pen attached to a electromagnetic drawing head which moves across the underside of a flat metal top over the plotter bed. A COM unit is a light protected compartment using an intense light beam to indirectly expose film or photosensitive paper. It can produce microfiche, microfilm or 310mm film positives. Such peripheral output devices produce the Bureau's computer-generated maps.

II. Spatial Display

At the Bureau, automated spatial display is expressed in the form of choropleth or dot maps. Of the two, choropleth maps are most common because they maintain areal boundaries. This is important since the Bureau collects, processes, tabulates and displays data according to a geographic hierarchy of politically or statistically defined areas. Through black/white patterns of color coded data, choropleth maps depict the spatial variation between defined areas. For Census computer mapping, the availability of data for many variables at each level of the geographic hierarchy permits numerous maps to be created by replacing input data in relation to the same digital cartographic file. Two such files used for computer mapping are the DIMECO and Urban Atlas Files.

The DIMECO File is a digital cartographic file containing boundary points for all U.S. counties. It has been used as the digital base for many Census computer maps of the United States by county. Such maps furnish a strategic perspective for revealing state and regional patterns within the national context; thereby, identifying regions for further study and assessment. Urban Atlas Files are digital cartographic files of the 1970 Standard Metropolitan Statistical Areas (SMSA's) census tracts. Mapping at the census tract level provides a tactical viewpoint of major urban areas. The existence of this set of files for over 230 SMSA's presents an opportunity for large scale automated mapping not previously available. One of these files was used to preserve the spatial orientation of census tracts while concurrently increasing the size of inner city tracts and decreasing the size of peripheral city tracts. This was accomplished by computer mapping the locational coordinates of census tract boundary points through an exponential function. Beyond the ability to portray information, maps also enhance the analytical capability of their users.

In cartographic analysis, the presentation of statistics on maps often reveals geographic patterns which might otherwise not be apparent; furthermore, statistical maps may also help to clarify and reduce the complexity of relationships between variables distributed over the same area. Mapping this areal association is a cartographic means for identifying geographic areas common to the variables mapped together and for indicating possible correlations between them. The rapid generation of cartographic artwork for color maps through computer mapping prompted Census efforts which lead to displaying the spatial interrelationship of phenomena through the "cross-map". Cross-maps result from overlaying two single variable color choropleth maps, each with its own color coded data. The overlayed colors produce new colors representing the degree of spatial association between two variables and displays their common spatial patterns of interrelationship. The primary value of computer-generated cartographic artwork; however, is for the offset printing of color maps, an activity which the Bureau performs frequently.

III. Computer-generated Cartographic Artwork

To understand how computer-generated cartographic artwork is prepared and used in the offset printing of color maps, both the conventional and computer process are briefly described. The objective of both processes are screened composite negatives, which are necessary to produce printing plates.

In the conventional cartographic process, screened composite negatives are created progressively through a map negative, peelcoat and open window negative. Initially, a base map outline is photographed to create a map negative, i.e., dark lines are transparent and light areas black. This map negative is used to transfer the map image onto a sheet of peelcoat by placing the map negative over the light sensitive peelcoat, exposing it to an intense light source and developed. The adhesive surface of the peelcoat now has an etched map image. Selected map areas on this layer are then manually peeled off to create an open window negative. An open window negative defines map areas to receive the color representing a data interval. In addition, once an open window is screened with dot patterns, it will regulate the shade of color. Screened composite negatives are required because the colors for areas on the printed map are combinations of varying shades of printing colors. Therefore, a screened composite negative for each of the four primary printing colors of red, yellow, blue and black is created. Each one is formed by successive exposures (composite) of screened open windows for all map areas containing a percentage of that color before the screened composite negative is developed. Prior to platemaking for offset printing, a color proof of the screened composite negatives is made to insure accurate overlaying registration and to verify color fidelity.

The computer mapping process of preparing cartographic artwork for offset printing relieves much of the photographic, chemical and manual peeling operations associated with conventional cartographic methods. The computer mapping process differs from conventional procedures by creating open window negatives directly from computer-generated film positives (reverse of open window negatives) created on the COM unit. The digital cartographic file defining map areas is used to generate a scan-line file. This file defines the optical horizontal shading of map areas during the COM operation. Within the COM unit, commands from the plot tape direct movements of the light beam to expose areas according to the scan-line files established for each map area. The actual output of the COM unit is a 310mm (8½ x 11 inch) film positive. After photographic developing, the areas which are to receive a particular color are black, while the remaining areas of the 310mm frame are transparent. A contact negative of the film positive is then made, developed and results in an open window negative. Similarly, as in the conventional cartographic process, screened composite negatives are then created. Thus, the computer mapping process significantly reduces the time, effort and costs required to produce cartographic artwork for offset printing color maps. Computer mapping also includes non-mapping applications, many of which are used to complement automated mapping output. This combination is generally representative of output from the spatial analysis of Census digital cartographic files.

IV. Spatial Analysis

The structures of Census digital cartographic files, in geometric terms, refer to the type of geographic entity represented. Thus point, line segment and polygon structures pertain to a geographic place, linear feature and area. These files, depending on their structure, are useful for health planning in a number of ways.

In a point structure, locational coordinates identify specific geographic places. The Census MEDList-X file contains locational coordinates for population centroids of approximately 262 thousand enumeration districts and their 1970 Census population and housing counts. This file used in conjunction with RADII5, a Census computer mapping program, produces a tabular and graphic output indicating the number of people or houses within a specified radii of a designated geographic point. It can compute up to nine radii at a time and the resulting circle can be partitioned up to 32 sector totals. Applications of this file and program include studies on possible pollution sites, population density, civil defense, munitions transport and railroad derailment. Point structures can also be used to compute a weighted geographic centroid. For example, given the locational coordinates of geographic points and their associated variable weight at each point, the locational center of gravity can be determined. This would be appropriate for locating a regional health center within a number of counties from several adjoining states or to determine where the center of gravity for some disease distribution might be.

A line segment structure refers to a pair of nodes and the line segment joining them. These nodes have locational coordinates and geographic information is encoded for both sides of the line segment. Good examples are the Geographic Base File/Dual Independent Map Encoding Files (GBF/DIME-Files), which constitute a computerized version of the Census Bureau's Metropolitan Map Series (MMS) for the urban cores of most Standard Metropolitan Statistical Areas (SMSA's). Within these files, nodes are formed by street intersections and the line segment occuring between nodes have right/left block faces. This type of digital cartographic structure has the ability to distinguish what areas and other geographic information is adjacent to a line segment. A GBF/DIME-File is useful in geocoding addresses, i.e., matching records of an address file with a GBF/DIME-File on the basis of ZIP code, street name and address range. When a match occurs, the address is assigned all the geographic identifiers from the GBF/DIME-File related to the address range on either the right or left side of the street segment, as the case may be. Hence, health data collected and associated with an address can be assigned to areal units ranging from Census blocks to States. This is one way in which health data can be made compatible with Census areal data for subsequent analysis.

As previously described, polygon structures have locational coordinates for sets of points; each set of points defining an area. The Urban Atlas Files have such a structure. These structures are useful for measuring the area enclosed by a digital boundary file.

Area measurement has applications for health related density studies. A polygon structure can be formed by chaining line segments, such as those found in Census GBF/DIME-Files. Since street addresses can be geocoded, address matching in terms of encoded geographic information can determine if an address is located within some user defined health district. In the case of naturally occuring phenomena, for which locational coordinates are available, their locations inside or outside a polygon (area) can be checked through point-in-polygon routines.

The aforementioned are but a few applications of digital cartographic files. New applications are evolving as more people begin to use these files for their own purposes.

V. Prospects

The prospects for computer mapping are good. This is due to an increasing acceptance of computers in cartography matched with a growing recognition of the value of computer mapping by professionals working with spatially related data. It is evidenced by more user conferences and supported through the presence of modern and superior computer technology; moreover, the 1980 Decennial Census will furnish new data and digital cartographic files. Finally, when computer mapping, don't let the means become the ends, yet, don't let the means be without ends. The propagation and advancement of computer mapping requires a distinct blend of imagination, ingenuity and knowledge integrated within a theoretical/operational context distinguishing computer mapping as a proper field of study and research.