

Mapping for the Masses: GIS Lite and Online Mapping Tools in Academic Libraries

Kathleen W. Weessies
and Daniel S. Dotson

ABSTRACT

Customized maps depicting complex social data are much more prevalent today than in the past. Not only in formal published outlets, interactive mapping tools make it easy to create and publish custom maps in both formal and more casual outlets such as social media. This article defines GIS Lite, describes three commercial products currently licensed by institutions, and discusses issues that arise from their varied functionality and license restrictions.

INTRODUCTION

News outlets from newspapers to television to Internet these days are filled with maps that make it possible for readers to visualize complex social data. Presidential election results, employment rates, and the plethora of data arising from the Census of Population are just a small sampling of social data mapped and consumed daily. The sharp rise in published maps in recent years has increased consumer awareness of the effectiveness of presenting data in map format and has raised expectations for finding, making and using customized maps. Not just in news media, but in academia also, researchers and students have high interest in being able to make and use maps in their work.

Just a few years ago even the simplest maps had to be custom made by specialists. Researchers and publishers had to seek out highly trained experts to make maps on their behalf. As a result, custom maps were generally only to be found in formal publications. The situation has changed partly because geographic information system (GIS) software for geographic analysis and map making is more readily available than in years past. It does, however, remain specialized and wants considerable training for users to be proficient at even a basic level.¹ This gap between supply and demand has been partly filled, especially in the last five years, by the growth of Internet-based “GIS Lite” tools. While some basic tools are freely available on the Internet, several tools are subscription-based and are licensed by libraries, schools and businesses for use. College and university libraries especially are quickly becoming a major resource for data visualization and mapping tools.

The aim of this article is to describe several data-rich GIS Lite tools available in the library market and how these products have met or failed to meet the needs of several real-life college class

Kathleen W. Weessies (weessie2@msu.edu), a LITA member, is Geosciences Librarian and Head of the Map Library, Michigan State University, Lansing, Michigan. **Daniel S. Dotson** (dotson.77@osu.edu) is Mathematical Sciences Librarian and Science Education Specialist, Associate Professor, Ohio State University Libraries, Columbus, Ohio.

situations. This is followed by a discussion of issues arising from user needs and restrictions posed by licensing and copyright.

WHAT IS GIS LITE?

Students and faculty across the academic spectrum often discover that their topic has a geographic element to it and a map would enhance their work (paper, presentation, project, poster, article, book, thesis or dissertation, etc.). If their research involves data analysis, geospatial tools will draw attention to spatial patterns in the data that might not otherwise be apparent. Every scholar with such needs must make a cost/benefit decision concerning GIS: is his or her need greater than the cost in time and effort (sometimes money) necessary to learn or hire skills to produce map products? A full functioning GIS, being a specialized system of software designed to work with geospatially referenced datasets, is designed to address all the problems above. The data may be analyzed and output into customized maps exactly to the researcher's need. The traditional low-end solution available to non-experts, on the other hand, is coloring a blank outline map, either with hand-held tools (markers, colored pencils, etc.) or on a computer using a graphic editing program. The profusion of web mapping options dangles tantalizingly with possibility, and occasionally (and increasingly) is able to provide an output that illustrates a useful point of users' research in a professional enough manner to fill a need.

In recent years the web has blossomed with map applications collectively called the "GeoWeb" or "geospatial web." GeoWeb or geospatial web refers to the "emerging distributed global GIS, which is a widespread distributed collaboration of knowledge and discovery."² Some GeoWeb applications are well known street map resources such as Google Maps and MapQuest. Others are designed to deliver data from an organization, such as the National Hazards Support System (<http://nhss.cr.usgs.gov>), National Pipeline Mapping System (<http://www.npms.phmsa.dot.gov/PublicViewer>), and the Broadband Map (<http://www.broadbandmap.gov>). A few tools focus on map creation and output such as ArcGIS Online (<http://www.arcgis.com/home/webmap/viewer.html>) and Scribble Maps (<http://www.scribblemaps.com>). The newest subgenre of the GeoWeb consists of participatory mapping sites such as OpenStreet Map (<http://www.openstreetmap.org>), Did You Feel It? (<http://earthquake.usgs.gov/earthquake.usgs.gov/earthquakes/dyfi>), and Ushahidi (<http://community.ushahidi.com/deployments>).

The GeoWeb literature is small but growing.³ Elwood reviewed published research on the geographic web.⁴ The GeoWeb literature tends to focus on creation of mappable data and delivery of GeoWeb services.⁵ In these the map consumer only appears as a contributor of data. Very little has been written about users' needs from the GeoWeb.

The term GIS Lite has arisen among map and GIS librarians to describe a subset of GeoWeb applications. GIS Lite is useful to library patrons lacking specialized GIS training but who wish to conduct some GIS and map-making activities on a lower learning curve. For the purpose of this article, GIS Lite will refer to applications, usually web-based, which allow users to manipulate geospatial data and create map outputs without programming skills or training in full GIS software.

While many GeoWeb applications allow only low-level output options, GIS Lite will provide an output intended to be used in activities or rolled into a GIS for further geospatial processing.

In libraries, GIS Lite is closely allied with data and statistics resources. Data and statistics librarianship have already been discussed as disciplines in the literature such as by Hogenboom⁶ and Gray.⁷ New technologies and access to deeper data resources such as the ones presented here have raised the bar for librarians' responsibilities for curating, serving, and aiding patrons in its use. Rather than be passive shepherds of information resources, librarians are now active participants and even information partners. Librarians with map and GIS skills similarly can directly enhance the quality of student scholarship across academic disciplines.⁸ The GIS Lite resources, however, need not remain specialized tools of map and GIS librarians. Librarians working in disciplines across the academic spectrum may incorporate them into their arsenal of tools to meet patron needs.

DATA VISUALIZATION TOOLS

A growing number of academic libraries have licensed access to online data providers. The following data tools contain enough GIS Lite functionality to aid patrons in visualizing and manipulating data (primarily social data) and creating customized map outputs. Three of the more powerful commercial products described here are Social Explorer, SimplyMap, and ProQuest Statistical Datasets.

Social Explorer

Licensed by Oxford University Press, Social Explorer provides selected data from the US Decennial Census 1790 to 2010, plus American Community Survey 2006 through 2010.⁹ The interface enables either retrieval of tabular data or visualization of data in an interactive map. As the user selects options through pull-down menus, the map automatically refreshes to reflect the chosen year and population statistics. The level of geography depicted defaults to county level data. If a user zooms in to an area smaller than a county, then data refreshes to smaller geographies such as census tracts if they are available at that level for that year. Output is in the form of graphic files suitable for sharing in a computer presentation (see figure 1). One advantage of Social Explorer is that it utilizes historic boundaries as they existed for states, territories, counties, and census tracts for each given year. Social Explorer utilizes data and boundary files generated by the National Historical GIS (NHGIS) based at the University of Minnesota in collaboration with other partners. The creation of these historical boundaries was a significant undertaking and accomplishment.¹⁰ Custom tables of data and the historic geographic boundaries may also be retrieved and downloaded for use from an affiliated engine through the NHGIS website (<http://www.nhgis.org>). A disadvantage of this product is that the tool, while robust, does not completely replicate all the data available in the original paper census volumes. Also, historical boundaries have not been created for city or township-level data. The final map layout is not customizable either in the location of title and legend or in the data intervals.

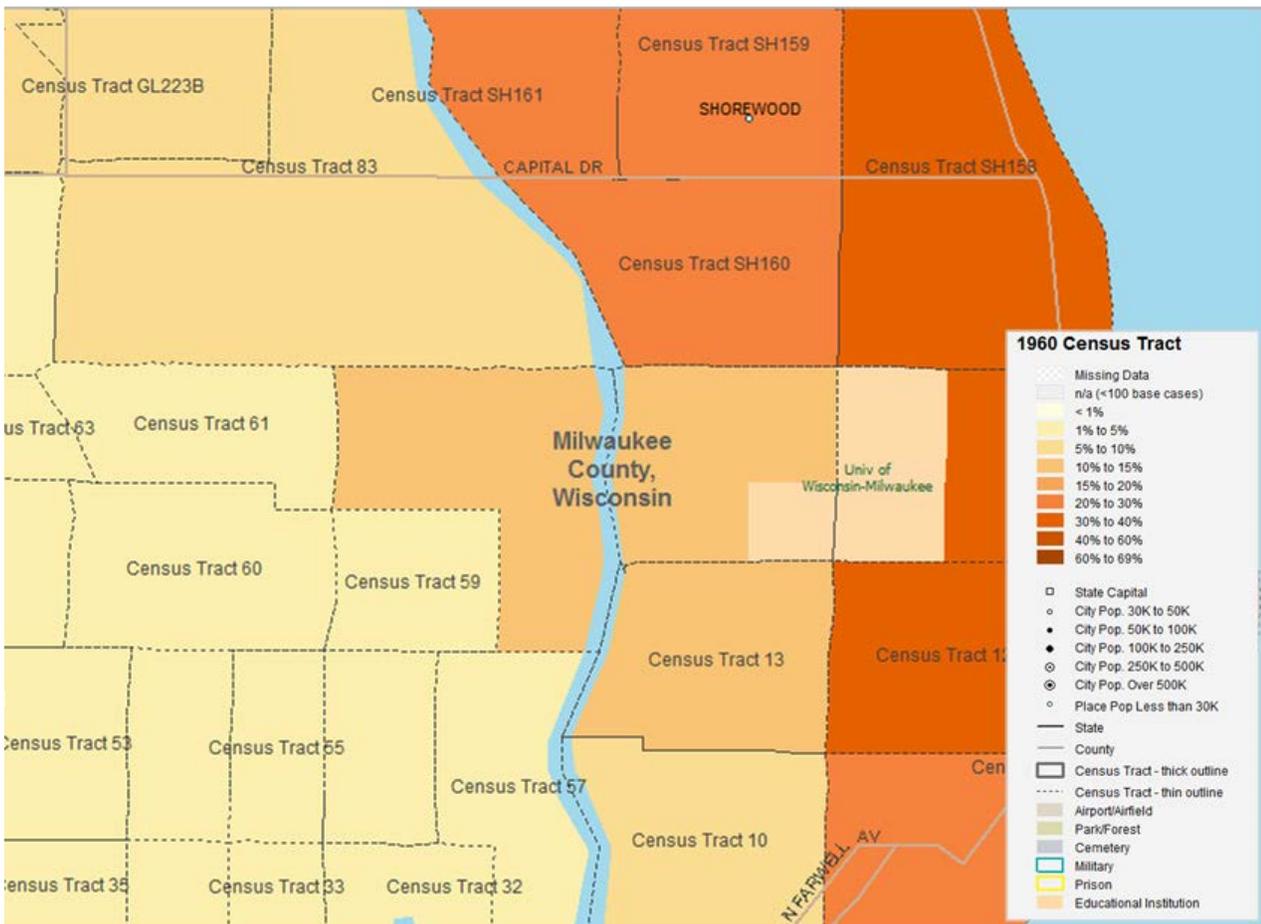


Figure 1: Map Depicting Population Having Four or More Years of College, 1960 (Source: Social Explorer, 2012; image used with permission)

SimplyMap

SimplyMap (<http://geographicresearch.com/simplymap>) is a product of Geographic Research. This powerful interface brings together public and licensed proprietary data to offer a broad array of 75,000 data variables in the United States. US Census Data are available 1980–2010 normalized to the user’s choice of either year 2000 or year 2010 geographies. Numerous other licensed datasets primarily focus on demographics and consumer behavior, which makes it popular as a marketing research tool. Each user establishes a personal login which allows created maps and tables to persist from session to session. Upon creating a map view, the user may adjust the smaller geographic unit at which the theme data is displayed and also may adjust the data intervals as desired. The user creates a layout, adjusting the location of the map legend and title before exporting as a graphic or PDF (see figure 2). Data are also exportable as GIS-friendly shapefiles.

The great advantage of this product is the ability to customize the data intervals. This makes it possible to filter the data and display specific thresholds meaningful to the user. For instance if a user needs to illustrate places where an activity or characteristic is shared by “over half” of the population, then one may change the map to display two data categories: one for places where up to 50 percent of the population shares the characteristic and a second category for places where more than 50 percent of the population shares the characteristic. Another potential advantage is that all local data have been allocated pro rata so that all variables, regardless of their original granularity, may be expressed by county boundaries, by zip code boundaries, or by census tract. A disadvantage of the product is the lack of historical boundaries to match historical data.

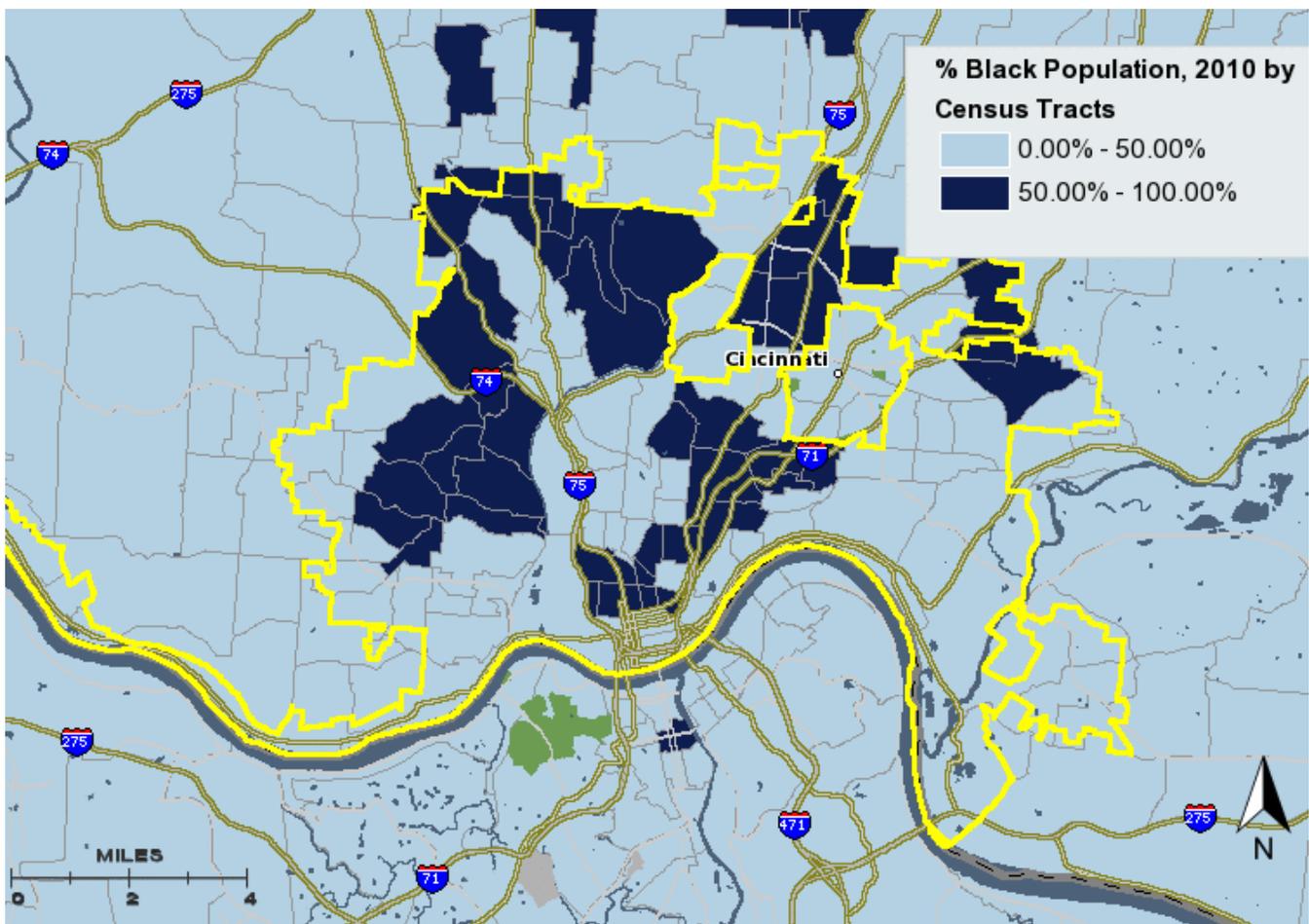


Figure 2. Map Depicting Census Tracts That Have More Than 50% Black Population (Yellow Line Indicates Cincinnati City Boundary) (Source: SimplyMap, 2012; image used with permission)

ProQuest Statistical Datasets

Statistical Datasets was developed by Conquest Systems Inc. and is licensed by ProQuest. This product also mingles a broad array of several thousand public and licensed proprietary datasets, including some international data, in one interface. The user may retrieve data and view it in tabular or chart form. If the data have a geographic element, then the user may switch the view to a map interface. The resulting map may be exported as an image. The data may also be exported to a GIS-friendly shapefile format. This product offers more robust data manipulation than the other products, in that the user may perform calculations between any of the data tables and create a chart or map of the created data element (see figure 3). Statistical Datasets, however, has more simplistic map layout capabilities than the other products.

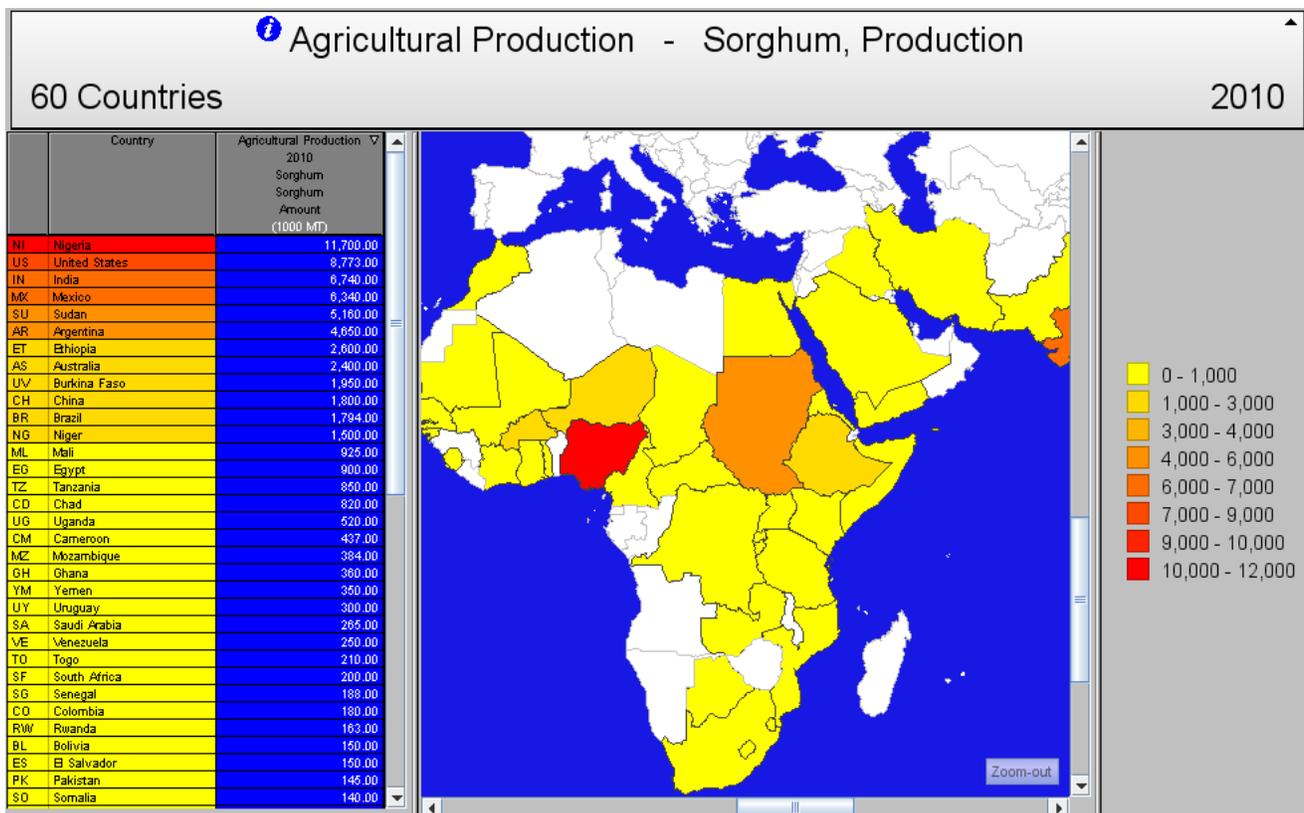


Figure 3. Map of Sorghum Production, by Country, in 2010 (Source: ProQuest Statistical Datasets, 2012; image used with permission)

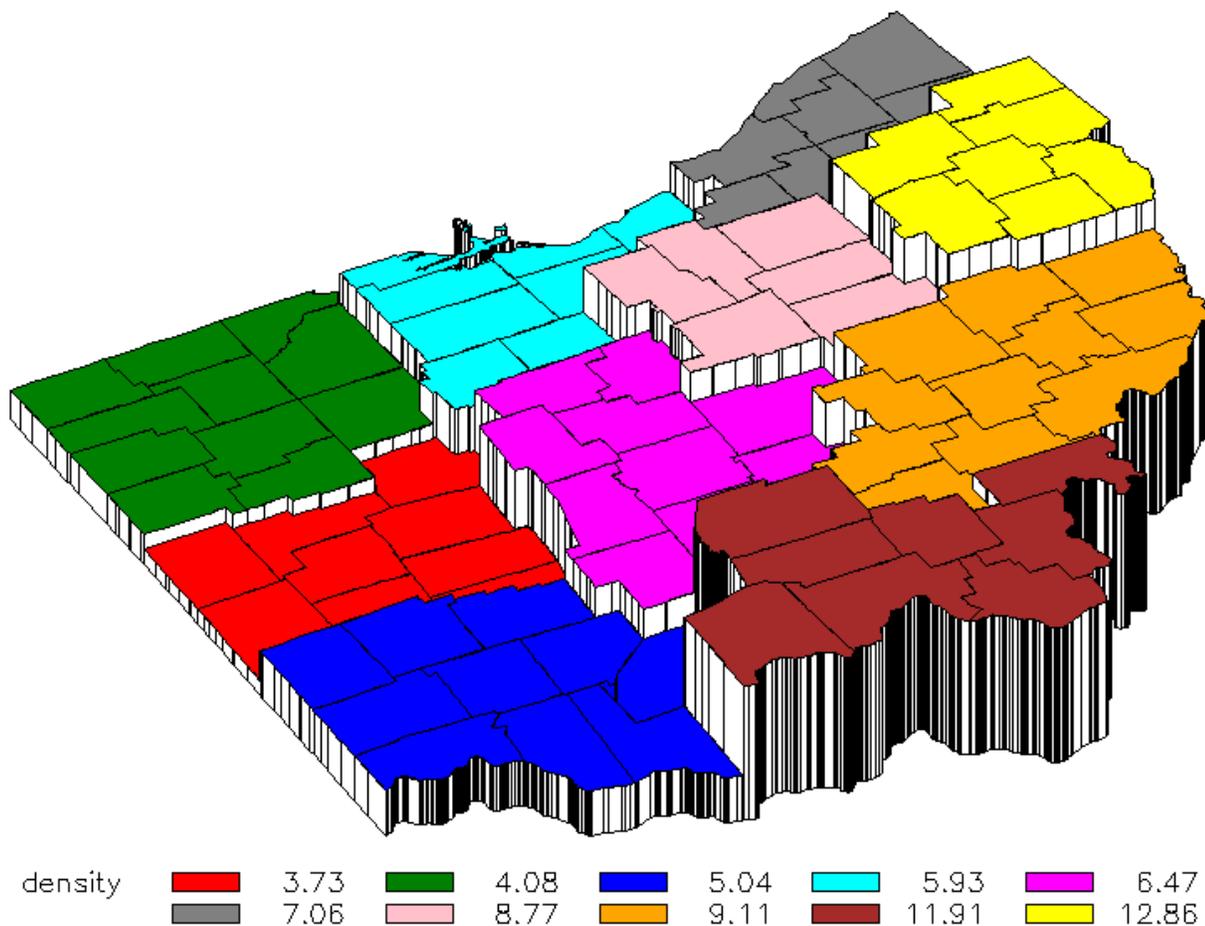
CASE STUDIES

The following three case studies are of college classroom situations in which students utilized maps or map making as part of the assigned course work. The above mapping options are assessed for how well they met the assignment needs.

Case Study 1

An upper level statistics course at The Ohio State University requires students to create maps using SAS (<http://www.sas.com>). While many may not associate the veteran statistical software package with creating maps, this course uses it along with SAS/GRAPH to combine statistical data with a map. The project requires data articulated at the county level in Ohio, which the students then combine into multi-county regions. The end result is a map with regions labeled and rendered in 3D according to the data values. An example of the type of map that could be produced from such data using SAS can be seen in figure 4.

RABBIT DENSITY* FOR SUMMER 98, RURAL MAIL CARRIER SURVEY



*Density is the average number of rabbits counted per 1000 route miles

Figure 4. Map of Observed Rabbit Density in Ohio using SAS, SAS/GRAPH, and Mail Carrier Survey Data,1998 (image used with permission)

While the data are provided in this course, students could potentially seek help from the library in a traditional way to find numerical data expressed at a county level. The librarian would guide

patrons through appropriate avenues to locate data such as to the three products listed above. All three options contain numerous data variables for Ohio at the county level. Because the students are further processing the data elsewhere (in this case SAS), the output options of the three products are less important. Ultimately the availability of data on a desired subject would be the primary determinant for choosing one of the three GIS Lite options discussed here. Social Explorer will export the data in tabular form which can then be ingested into SAS. SimplyMap and ProQuest Statistical Datasets would both be a bit easier, though, because both packages allow the user to export the data as shapefiles which are directly imported into SAS/GRAPH as both boundary files and joined tabular data.

Case Study 2

A first year writing class at Michigan State University has a theme of the American ethnic and racial experience. Assignments all relate to a student's chosen ethnic group and geographic location from approximately 1880 to 1930. Assignments build upon each other to culminate in a final semester paper. Students with ancestors living in the United States at that time are encouraged to examine their own family's ethnicity and how they fit in their geographic context. Otherwise, students may choose any ethnic group and place of interest.

Maps are a required element in the assignments. Maps that display historical census data help students place the subject ethnic group into the larger county, state, and national context over the time frame. The students can see, for instance, if their subject household was part of an ethnic cluster or an outlier to ethnic clusters. The parameters for finding data and maps are generous and open to each student's interpretation. The wish is for students to find social statistics and maps that are insightful to their topic and will help them tell their story.

Of the three statistical resources considered above, currently the only useful one is Social Explorer because it covers the time period studied by the class. The students may map several social indicators at the county level across several decades and compare their local area to the region and the nation. Also they may save their maps and include them in their papers (properly credited).

Case Study 3

"The Ghetto" is an elective Geography class restricted to upperclassman at Michigan State University. In the semester project, students analyze the spatial organization and demographic variables of "ghetto" neighborhoods in a chosen city. A ghetto is defined as neighborhoods that have a 50 percent or higher concentration of a definable ethnic group. Since black and white are the only two races consistently reported at the Census Tract level for all the years covered by the class (1960 through 2010) the students necessarily use that data for their projects.

Data needs for the class are focused and deep. The students specifically need to visualize US census data from 1960 through 2010 at the census tract level within the city limits for several social indicators. Indicators include median income, median housing value, median rent, educational attainment, income, and rate of unemployment. The instructor has traditionally required use of the paper census volumes and students created hand-made maps that highlight

tracts in the subject city that conformed to the ghetto definition and those that did not for each of the census years covered. Computer-retrieved data and computer-generated maps would be acceptable, but at the time of this writing no GIS Lite product is able to make all the maps that meet the specific requirements of this class. Social Explorer covers all of the date range and provides data down to the tract level. However it does not provide an outline of the city limits and does not provide all the data variables required in the assignment. SimplyMap will only work for 2000 through 2010 because tract boundaries are only available for those two years even though the data go back to 1980. SimplyMap does provide two excellent features though: it is the only product that allows an overlay of the (modern) city boundary on top of the census tract map, and it is the only product that allows manipulation of the data intervals. Students may choose to break the data at the needed 50 percent mark, while the other products utilize fixed data intervals not useful to this class. ProQuest Statistical Datasets can compute the data into two categories to create the necessary data intervals; however Census data are only available beginning with Census 2000.

MAP PRODUCTS FOR USER NEEDS

These three real-life class scenarios illustrate how the rich and seemingly duplicative resources of the library can range from perfectly suitable to perfectly useless depending on each project's exact needs. The appropriateness of any given tool can only be assessed fairly if the librarian is familiar with all the "ins and outs" of every product. The GeoWeb and GIS Lite tools mentioned throughout this article are summarized in table 1. The suitability of GIS Lite tools will be further affected by the following issues.

Historical Boundaries

The range and granularity of data tools are subject to factors sometimes at odds with what a researcher would wish to have. At this time, for instance, many historical resources provide data only as detailed as the county level. County level data are available largely due to the efforts of the NHGIS mentioned above and the Newberry Library's Atlas of County Boundaries Project (<http://publications.newberry.org/ahcbp>). Far fewer resources provide historical data at smaller geographies such as city, township, or census tract levels. This is because the smaller the geographies get, the exponentially more there are to create and for map interfaces to process. From the well-known resource *City and County Data Book*,¹¹ it is easy enough to retrieve US city data. The historical boundaries of every city in the United States, however, have not been created. This is because city boundaries are much more dynamic than county boundaries and there is no centralized authoritative source for their changes over time. Two of the three case studies presented here utilized historic data. This isn't necessarily a representative proportion of user needs; librarians should assess data resources in light of their own patrons' needs.

Normalization

Two equally valid data needs concerning any kind of time series data concern changing geographic boundaries. Census tracts, for instance, provide geographic detail roughly at the neighborhood level designed by the Bureau of Census to encompass approximately 2,500 to 8,000

people.¹² Because people move around and the density of population changes from decade to decade, so the configuration and numbering of tracts change over time. Some scholars will wish to see the data values in the tracts as they were drawn at the time of issue. In this situation, a neighborhood of interest might belong to different tracts over the years or even be split between two or more tracts. Other scholars focused on a particular neighborhood may wish to see many decades of census data re-cast into stable tracts in order to be directly comparable. Data providers will take one approach or the other on this issue, and librarians will do well to be aware of their choice.

License Restrictions

A third issue affecting use of these products is the ability to use derived map images, not only in formal outlets such as professional presentations, articles, books, and dissertations, but also informal outlets such as blogs and tweets. For the most part GIS Lite vendors are willing—even pleased—to see their products promoted in the literature and in social media. The vendors uniformly wish any such use to be properly credited. The license that every institution signs when acquiring these products will specify allowed and disallowed activities. The license, fixated on disallowing abuse or resale or other commercialization of the data, might leave a chilling effect on users wishing to use the images in their work. If a user is in any doubt as to the suitability of an intended use of a map, he or she should be encouraged to contact the vendor to seek permission for its use.

As data resources grow and become more readily usable, the possibility for scholarly inquiry grows. Librarians with familiarity with GIS Lite tools may partner with their patrons and guide them to the best resources.

Table 1: A Selection of GeoWeb and GIS Lite Tools and Their Output Options

| Tool Name | URL | Free or Fee | Electronic Output Options* |
|---|---|--|---|
| GeoWeb Tools | | | |
| Atlas of Historical County Boundaries | http://publications.newberry.org/ahcbp/ | Free | Spatial data as Shapefile, KMZ; Image as PDF |
| Did You Feel It? | http://earthquake.usgs.gov/earthquakes/dyfi/ | Free | Tabular data as TXT, XML. Image as JPG, PDF, PS |
| Google Maps | https://maps.google.com/ | Free | None |
| MapQuest | http://www.mapquest.com | Free | None |
| National Broadband Map | http://www.broadbandmap.gov/ | Free | Image as PNG |
| National Hazards Support Systems (USGS) | http://nhss.cr.usgs.gov/ | Free | Image as PDF, PNG |
| National Pipeline Mapping System | https://www.npms.phmsa.dot.gov/PublicViewer/ | Free | Image as JSF |
| OpenStreetMap | http://www.openstreetmap.org/ | Free | Tabular data as XML; Image as PNG, JPG, SVG, PDF |
| Ushahidi Community - Deployments | http://community.ushahidi.com/deployments/ | Free | Image as JPG |
| GIS Lite Tools | | | |
| ArcGIS Online | http://www.arcgis.com | Limited free options; access is part of institutional site license | Spatial data as ArcGIS 10; Image as PNG (in ArcExplorer) |
| ProQuest Statistical Datasets | http://cisupa.proquest.com/ws_display.asp?filter=Statistical%20Datasets%20Overview | Fee | Tabular data as Excel, PDF, Delimited text, SAS, XML; Spatial data as Shapefile; Image may be copied to clipboard |
| SAS/GRAPH | http://www.sas.com/technologies/bi/query_reporting/graph/index.html | Fee | Image as PDF, PNG, PS, EMF, PCL |
| Scribble Maps | http://www.scribblemaps.com/ | Free | Spatial data as KML, GPX; Image as JPG |
| SimplyMap | http://geographicresearch.com/simplymap | Fee | Tabular data as Excel, CSV, DBF, Spatial data as Shapefile; Image as PDF, GIF |

* Does not include taking a screen shot of the monitor or making a durable URL to the page

REFERENCES

1. National Research Council, Division on Earth and Life Studies, Board on Earth Sciences and Resources, Geographical Sciences Committee, *Learning to Think Spatially* (Washington, D.C.: f Academies Press, 2006): 9.
2. Pinde Fu and Jiulin Sun, *Web GIS: Principles and Applications* (Redlands, CA: ESRI Press, 2011): 15.
3. For good overviews of the GeoWeb, see Muki Haklay, Alex Singleton and Chris Parker, "Web mapping 2.0: The Neogeography of the GeoWeb," *Geography Compass* 2, no. 6 (2008): 2011-2039, <http://dx.doi.org/10.1111/j.1749-8198.2008.00167.x>; Jeremy W Crampton, "Cartography: Maps 2.0," *Progress in Human Geography* 33, no. 1 (2009): 91-100, <http://dx.doi.org/10.1177/0309132508094074>.
4. Sarah Elwood, "Geographic Information Science: Visualization, Visual Methods, and the GeoWeb," *Progress in Human Geography* 35, no. 3 (2010): 401-408, <http://dx.doi.org/10.1177/0309132510374250>.
5. Songnian Li; Suzana Dragičević, and Bert Veenendaal eds, *Advances in Web-based GIS, Mapping Services and Applications* (Boca Raton, FL: CRC Press, 2011).
6. Hogenboom, Karen, Carissa Phillips, and Merinda Hensley, "Show Me the Data! Partnering with Instructors to Teach Data Literacy," in *Declaration of Interdependence: The Proceedings of the ACRL 2011 Conference, March 30-April 2, 2011, Philadelphia, PA*, ed. Dawn M. Mueller. (Chicago: Association of College and Research Libraries, 2011), 410-417, http://www.ala.org/acrl/files/conferences/confsandpreconfs/national/2011/papers/show_me_the_data.pdf.
7. Ann S. Gray, "Data and Statistical Literacy for Librarians," *IASSIST Quarterly* 28 no. 2/3 (2004): 24-29, <http://www.iassistdata.org/content/data-and-statistical-literacy-librarians>.
8. Kathy Weimer, Paige Andrew, and Tracey Hughes, *Map, GIS and Cataloging / Metadata Librarian Core Competencies* (Chicago: American Library Association Map and Geography Round Table, 2008), <http://www.ala.org/magirt/files/publicationsab/MAGERTCoreComp2008.pdf>.
9. Social Explorer. <http://www.socialexplorer.com/pub/home/home.aspx>.
10. Catherine Fitch and Steven Ruggles, Building the National Historical Geographic Information System *Historical Methods* 36, no. 1 (2003): 41-50, <http://dx.doi.org/10.1080/01615440309601214>.
11. U. S. Bureau of Census. *County and City Data Book*, <http://www.census.gov/prod/www/abs/ccdb.html>.

12. Census Tracts and Block Numbering Areas.

http://www.census.gov/geo/www/cen_tract.html.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Michael Fligner, Dr. Clarence Hooker, and Dr. Joe Darden for permission to use their courses as case studies.